# Geometric deep learning on the sphere

## Efficient generalised spherical CNNs

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

28 October 2020

# Mentimeter

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

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

# Outline

1. Symmetry in deep learning
2. Spherical CNNs
3. Efficient generalised spherical CNNs
4. Numerical experiments

# 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).

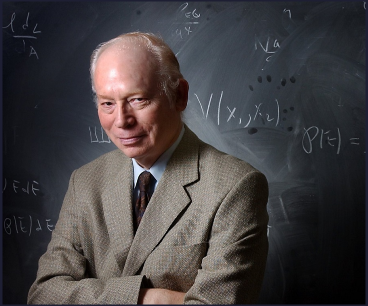
# Physics and deep learning

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

# 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.

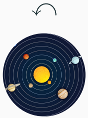
Spatial translation

Rotation

# Symmetry (invariance) to continuous transformation

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

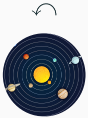
Spatial translation



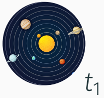
Rotation

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

Spatial translation



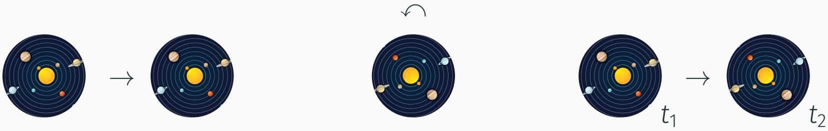
Rotation




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

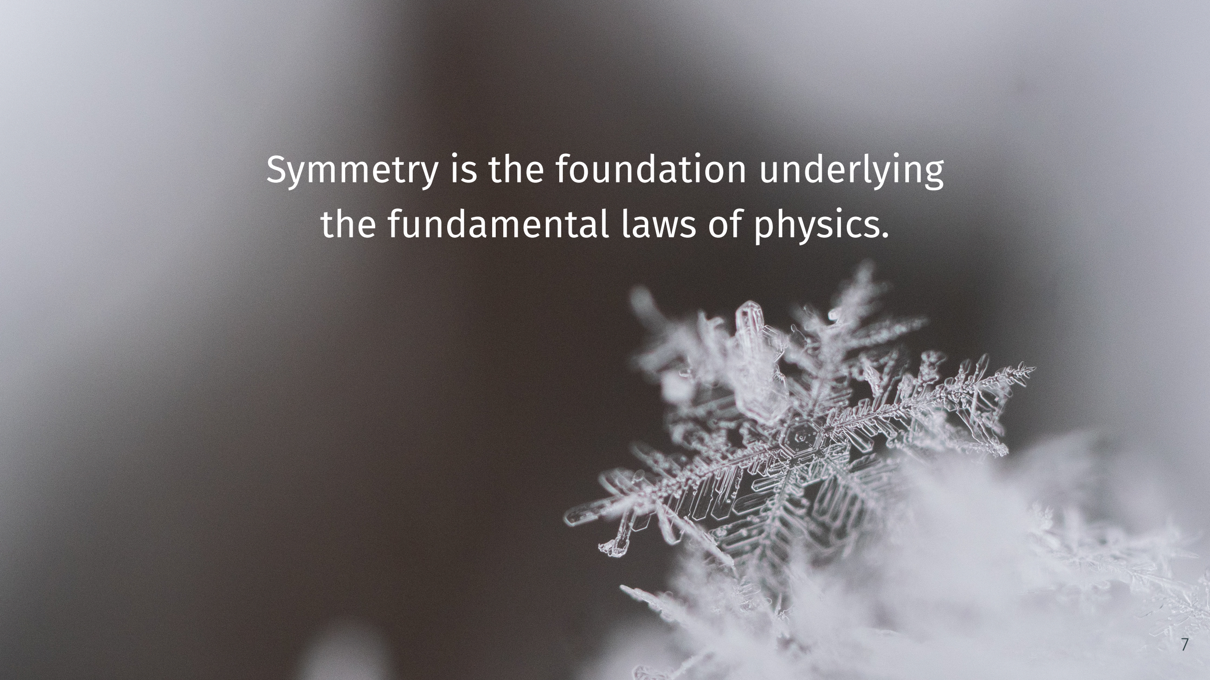
*For every continuous symmetry of the universe, there exists a conserved quantity.*

Symmetries at the heart of physics:

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

(Energy not conserved in general relativity since time translation broken.)

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

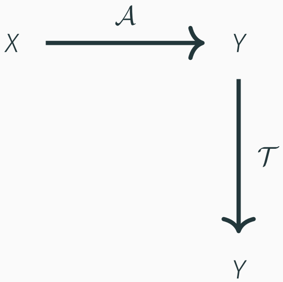
**Equivariance**

*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

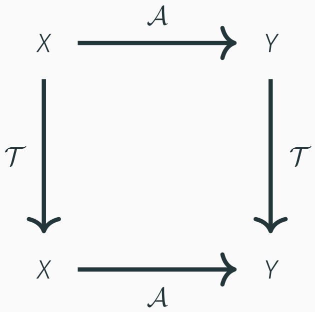
## Equivariance

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

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

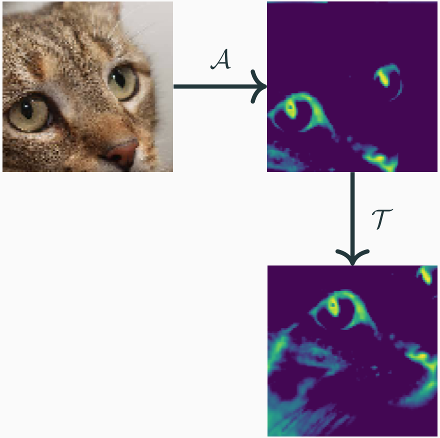
*for all possible signals $f$.*

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



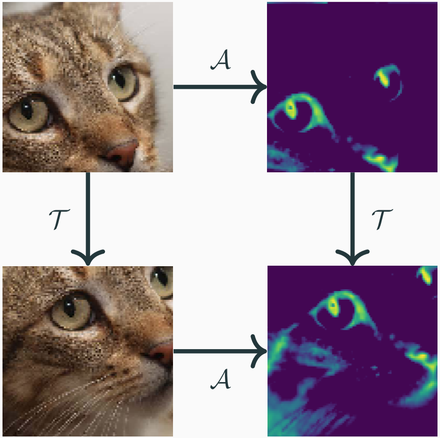
# Planar (Euclidean) CNNs exhibit translational equivariance

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

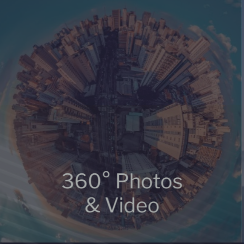
Still a cat

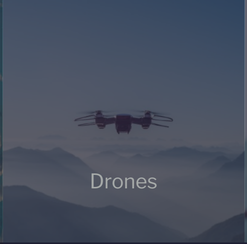
# Spherical CNNs

Data on the sphere is prevalent

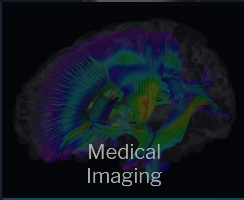
Data on the sphere is prevalent

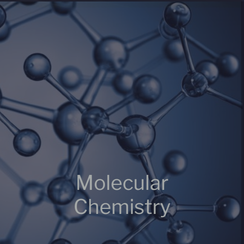
Encode symmetries of the sphere and rotations

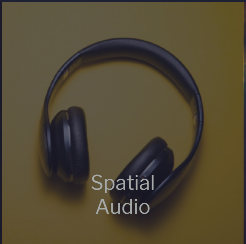


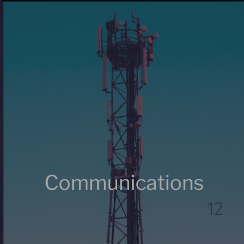



# Data on the sphere arises in many applications



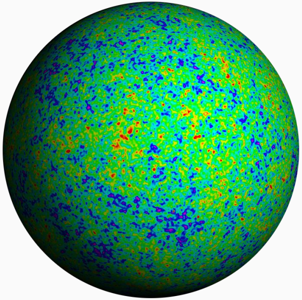






# Cosmology and virtual reality

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

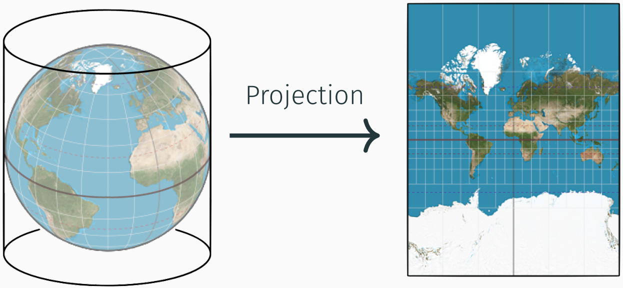


Cosmic microwave background

360° virtual reality

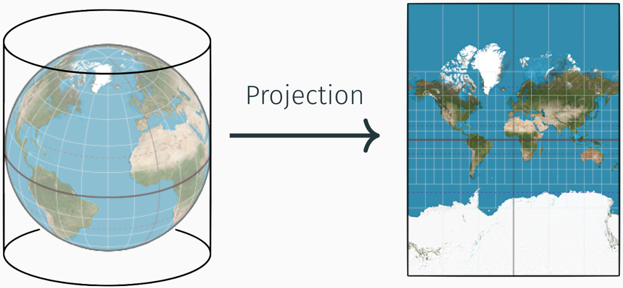
# Why not standard (Euclidean) deep learning approaches?

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.



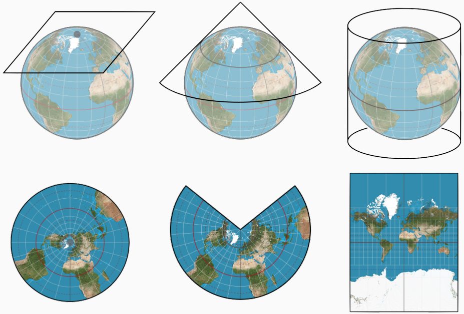

Greenland appears to be a similar size to Africa in the projected planar map, whereas it is over 10 times smaller.

# 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.

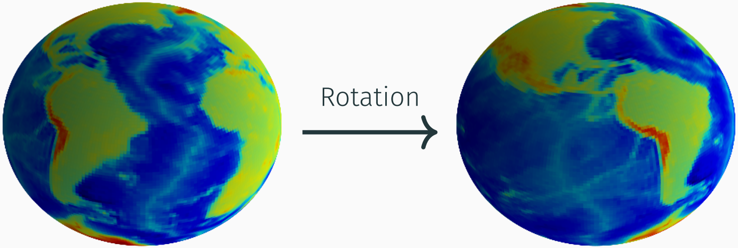


No projection of the sphere to the plane can preserve both shapes and areas
$\Rightarrow$ distortions are unavoidable.

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

# Rotational equivariance

On the sphere, the analog of translations are rotations.

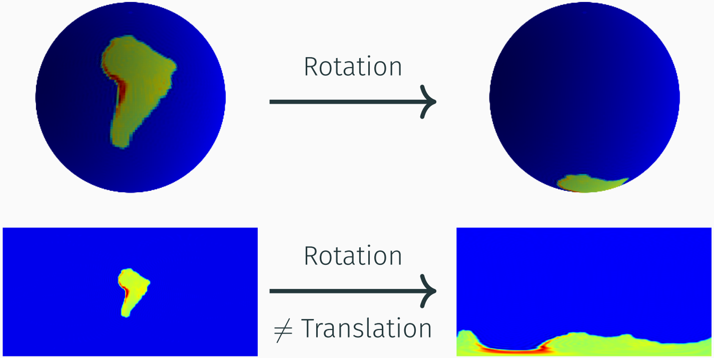



Would like spherical CNNs to exhibit rotational equivariance.

(Just as planar CNNs exhibit translational equivariance.)

# Rotational equivariance

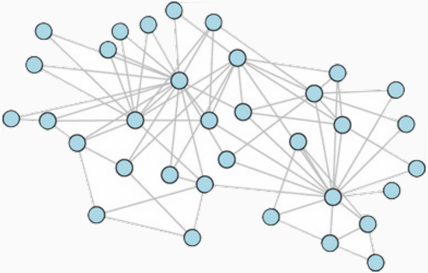
Rotational equivariance not captured by translational equivariance of planar CNNs.




(Harmonic networks constructed to capture rotational equivariance for planar images by Worral et al. 2017.)

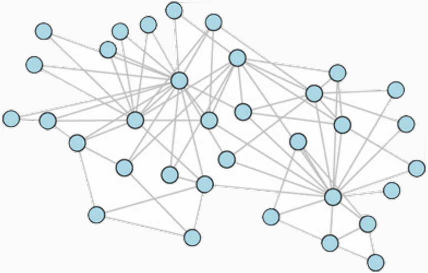
# Why not graph-based geometric deep learning?

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.



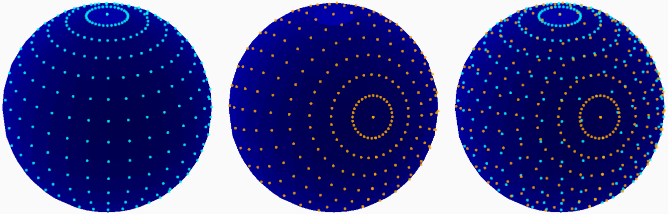
Again, **breaks symmetries and geometric properties** of sphere.

Cannot capture rotational equivariance.

# 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.

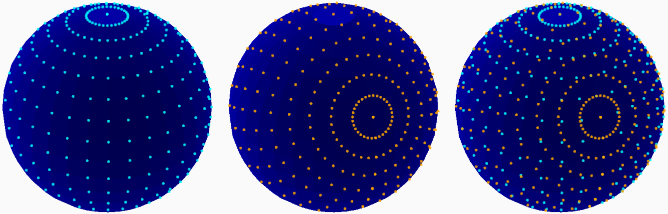


Capturing strict equivariance with operations defined directly in discretised (pixel) space **not possible** due to structure of sphere.

# Capturing rotational equivariance in spherical CNNs

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

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

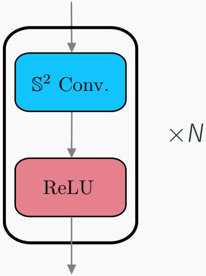


Capturing strict equivariance with operations defined directly in discretised (pixel) space **not possible** due to structure of sphere.

Instead, consider Fourier approach $\rightarrow$ access to underlying continuous representations.

# Spherical CNN

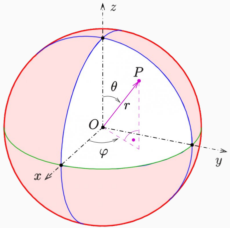
Spherical CNNs constructed by analog of Euclidean CNNs but using convolution on the sphere and with pointwise non-linear activations functions, e.g. ReLU (Cohen et al. 2018; Esteves et al. 2018).



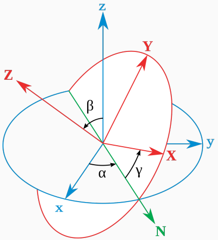

(Alternative, real space constructions have also been developed but do not exhibit rotational equivariance so not considered further; e.g. Boomsma & Frellsen 2017, Jiang et al. 2019, Perraudin et al. 2019.)

# Signals on the sphere and rotation group

Consider signals $f \in L^2(\Omega)$ on the sphere ($\Omega = \mathbb{S}^2$) or rotation group ($\Omega = \mathrm{SO}(3)$).




Sphere $\Omega = \mathbb{S}^2$
with coordinates $(\theta, \phi) \in \mathbb{S}^2$



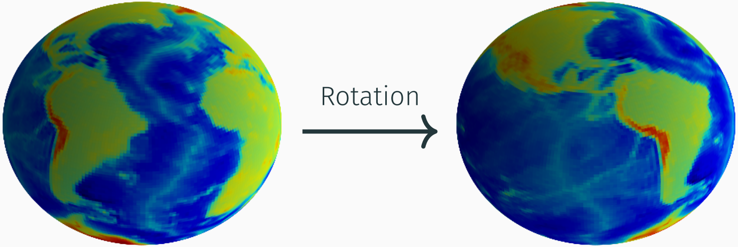

Rotation group $\Omega = \mathrm{SO}(3)$
with coordinates $\rho = (\alpha, \beta, \gamma) \in \mathrm{SO}(3)$

# 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 ($\Omega = \mathrm{SO}(3)$) can be rotated by $\rho \in \mathrm{SO}(3)$ by

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



# Convolution of signals

## Convolution of signals in spatial domain

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

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

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

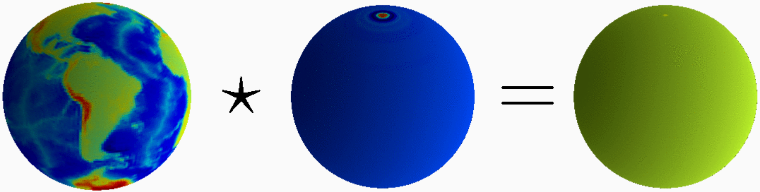
# Convolution of signals

## Convolution of signals in spatial domain

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

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

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



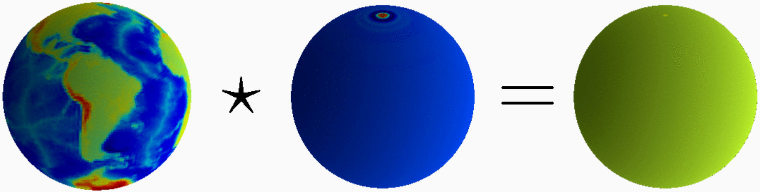
# Convolution of signals

## Convolution of signals in spatial domain

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

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

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



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

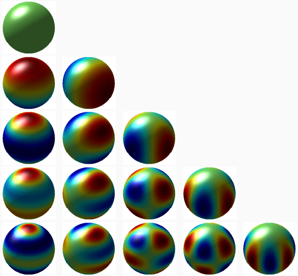
# Fourier transforms on the sphere and rotation group

## Fourier transform on sphere: spherical harmonic transform

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

$$f(\omega) = \sum_{\ell m} \hat{f}^{\ell}_{m} \, Y^{\ell}_{m}(\omega), \qquad \text{for } \omega \in \mathbb{S}^2.$$

where $\hat{f}^{\ell}_{m} = \langle f, Y^{\ell}_{m} \rangle$ (also denoted $\hat{f}^{\ell}$).



Spherical harmonics

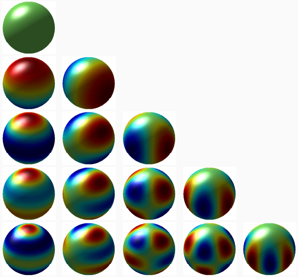
# Fourier transforms on the sphere and rotation group

## Fourier transform on sphere: spherical harmonic transform

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

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

where $\hat{f}^{\ell}_{m} = \langle f, Y^{\ell}_{m} \rangle$ (also denoted $\hat{f}^{\ell}$).



Spherical harmonics

## Fourier transform on rotation group: Wigner transform

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

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

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

# Harmonic computations

## Rotation of signals in harmonic domain

The rotation $f \mapsto \mathcal{R}_\rho f$ of a signal $f \in L^2(\Omega)$ can be described in harmonic space by

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

# Harmonic computations

## Rotation of signals in harmonic domain

The rotation $f \mapsto \mathcal{R}_\rho f$ 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

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

# Harmonic computations

## Rotation of signals in harmonic domain

The rotation $f \mapsto \mathcal{R}_\rho f$ 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

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

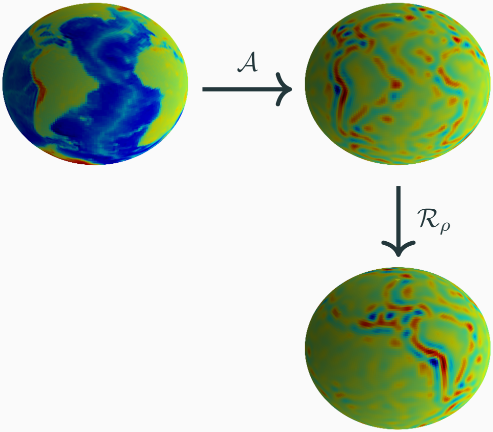
By computing convolutions in harmonic space, discretisation effects are eliminated.

Furthermore, fast harmonic transform algorithms can be leveraged (as we will see later).

# Convolution is rotationally equivariant

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

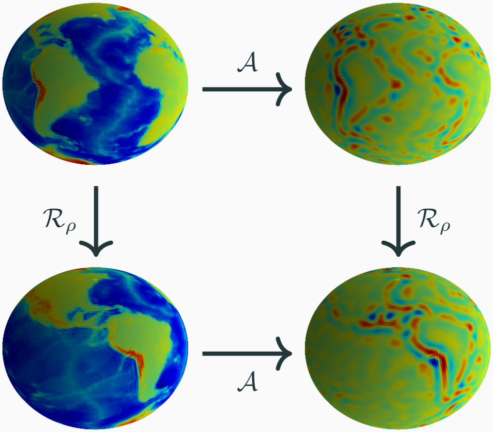
$$((\mathcal{R}_\rho f) \star \psi)(\rho') = (\mathcal{R}_\rho (f \star \psi))(\rho').$$



# Convolution is rotationally equivariant

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

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



# 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

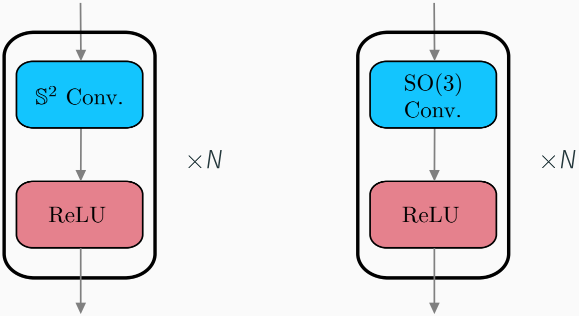
| Layer | Mean Relative Error* |
|---|---|
| $\mathbb{S}^2$ to $\mathbb{S}^2$ conv. | $4.4 \times 10^{-7}$ |
| $\mathbb{S}^2$ to SO(3) conv. | $5.3 \times 10^{-7}$ |
| SO(3) to SO(3) conv. | $9.3 \times 10^{-7}$ |
| $\mathbb{S}^2$ ReLU | $3.4 \times 10^{-1}$ |
| SO(3) ReLU | $3.7 \times 10^{-1}$ |

* Floating point precision.

# Spherical CNNs

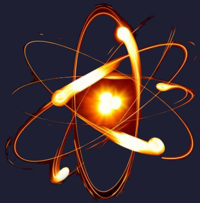
Approach taken by Cohen et al. 2018 and Esteves et al. 2018.

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.

# Generalised signals

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

# Generalised signals

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

**Generalised signals**

Consider set of variable length vectors of the form

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

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

# Generalised signals

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

## Generalised signals

Consider set of variable length vectors of the form

$$f = \{\hat{f}^\ell_t \in \mathbb{C}^{2\ell+1} : \ell = 0, .., L-1;\ t = 1, ..., \tau^\ell_f\},$$

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

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

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

# Rotation of generalised signals

## Rotation of generalised signals

The rotation $f \mapsto \mathcal{R}_\rho f$ of a signal $f \in \boldsymbol{\mathcal{F}}_L$ 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 $\boldsymbol{\mathcal{F}}_L$.

# Convolution of generalised signals

## Convolution of generalised signals

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

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

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

Do not force the filter $\psi$ 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).

# 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).

$D^\ell : \mathrm{SO}(3) \to \mathrm{GL}(\mathbb{C}^{2\ell+1})$ is an irreducible group representation of SO(3) on $\mathbb{C}^{2\ell+1}$ (since it is a group homomorphism from SO(3) to the general linear group $\mathrm{GL}(\mathbb{C}^{2\ell_1+1})$).

## 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**.

# 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 $\ell$ there is a block equal to $D^{\ell}(\rho)$.

# 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 $\ell$ 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

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

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

# Why is the tensor product decomposition useful?

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

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

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

# Non-linearly transforming generalised signals

Simply compute $(C^{\ell_1,\ell_2,\ell})^\top(\hat{f}^{\ell_1} \otimes \hat{f}^{\ell_2})$ for all pairs of input fragments and collect them into a generalised signal (Kondor et al. 2018).

# Non-linearly transforming generalised signals

Simply compute $(C^{\ell_1,\ell_2,\ell})^\top(\hat{f}^{\ell_1} \otimes \hat{f}^{\ell_2})$ 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 $\mathcal{N}_\otimes : \mathcal{F}_L \to \mathcal{F}_L$ defined as

$$\mathcal{N}_\otimes(f) = \{(C^{\ell_1,\ell_2,\ell})^\top(\hat{f}^{\ell_1}_{t_1} \otimes \hat{f}^{\ell_2}_{t_2}) \ : \ \ell = 0, ..., L-1;\ (\ell_1, \ell_2) \in \mathbb{P}^\ell_L;\ t_1 = 0, ..., \tau_f^{\ell_1};\ t_2 = 0, ..., \tau_f^{\ell_2}\},$$

where for each degree the set

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

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

# Non-linearly transforming generalised signals

Equivariance errors

| Layer | Mean Relative Error* |
| --- | --- |
| Tensor-product activation $\rightarrow$ Generalized conv. | $5.0 \times 10^{-7}$ |
| $\mathbb{S}^2$ ReLU | $3.4 \times 10^{-1}$ |
| SO(3) ReLU | $3.7 \times 10^{-1}$ |

* Floating point precision.

# Generalised spherical CNNs

Consider the $s$-th layer of a generalised spherical CNN to take the form of a triple (Cobb et al. 2020)

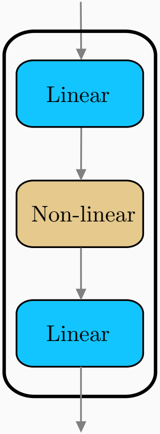
$$\mathcal{A}^{(s)} = (\mathcal{L}_1, \mathcal{N}, \mathcal{L}_2),$$

such that

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

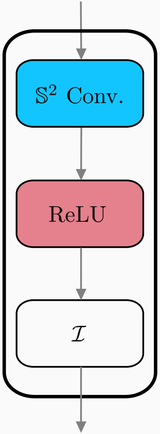
where

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

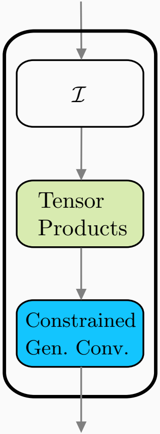


# Generalised spherical CNNs

- Encompass other frameworks as special cases.



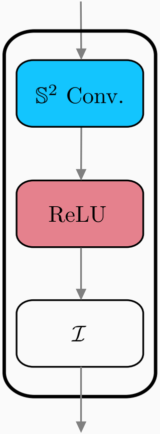

Cohen et al. (2018), Esteves et al. (2018)



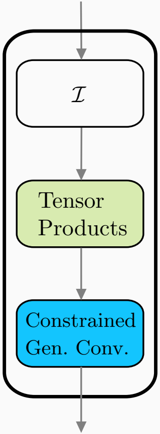

Kondor et al. (2018)

# Generalised spherical CNNs

- Encompass other frameworks as special cases.
- General framework supports hybrids models.




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

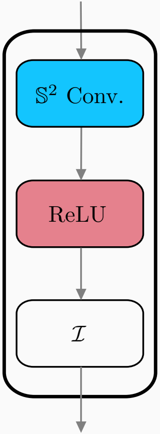



Kondor et al. (2018)

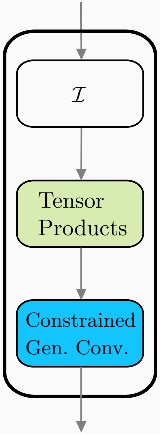
# Generalised spherical CNNs

- Encompass other frameworks as special cases.
- General framework supports hybrids models.

Construct more efficient layers...




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




Kondor et al. (2018)

# Computational cost of strictly equivariant layers

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

# Computational cost of strictly equivariant layers

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

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

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

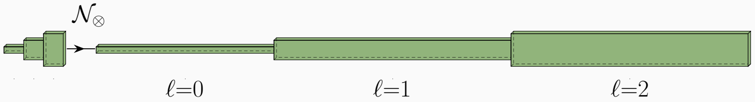
where $C$ is **representational capacity** and $L$ **spatial resolution** (bandlimit):

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

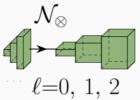
# Channel-wise structure

Let network activations take form $(f_1, \ldots, f_K) \in \mathcal{F}_L^K$ where $f_i \in \mathcal{F}_L$ all same type $\tau_f$.

Apply $\mathcal{N}_\otimes$ to each of $K$ channels separately $\Rightarrow$ **cost is reduced by $K$ times** relative to single channel with the same total number of fragments.




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

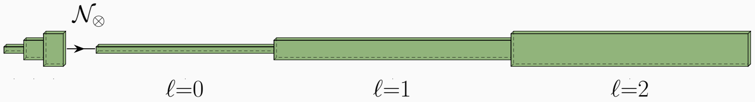



Ours (Cobb et al. 2020)

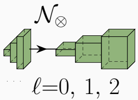
# Channel-wise structure

Let network activations take form $(f_1, \ldots, f_K) \in \mathcal{F}_L^K$ where $f_i \in \mathcal{F}_L$ all same type $\tau_f$.

Apply $\mathcal{N}_\otimes$ to each of $K$ channels separately $\Rightarrow$ **cost is reduced by $K$ times** relative to single channel with the same total number of fragments.




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




Ours (Cobb et al. 2020)

Using $C = K$ to control representational capacity, reduce $\mathcal{O}(C^2)$ to $\mathcal{O}(C)$.

# Constrained generalised convolution

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

1. **Uniform convolution**: linear projection uniformly across channels to project down onto the desired type (interpreted as learned extension of $\mathcal{N}_{\otimes}$ to undo the drastic 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.

# Constrained generalised convolution

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

1. **Uniform convolution**: linear projection uniformly across channels to project down onto the desired type (interpreted as learned extension of $\mathcal{N}_{\otimes}$ to undo the drastic 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 dramatically improved**.

# Optimised 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.

Recall, degree mixing set $\mathbb{P}_L^\ell$:

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

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

# Optimised degree mixing sets

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

- Some notion of relationship between $\ell_1$ and $\ell_2$ is captured if there exists a path between the two nodes in $G_L^\ell$.
- Select smallest subgraph such that all relationships are preserved $\Rightarrow$ **minimum spanning tree** (MST). Weight edges by computational cost to minimise overall cost.

# Optimised degree mixing sets

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

- Some notion of relationship between $\ell_1$ and $\ell_2$ is captured if there exists a path between the two nodes in $G_L^\ell$.
- Select smallest subgraph such that all relationships are preserved $\Rightarrow$ **minimum spanning tree** (MST). Weight edges by computational cost to minimise overall cost.

Two optimised degree mixing sets:

1. **MST**: reduce resolution complexity from $\mathcal{O}(L^5)$ to $\mathcal{O}(L^4)$, whilst preserving performance.

# Optimised degree mixing sets

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

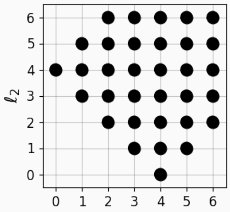
- Some notion of relationship between $\ell_1$ and $\ell_2$ is captured if there exists a path between the two nodes in $G_L^\ell$.
- Select smallest subgraph such that all relationships are preserved $\Rightarrow$ **minimum spanning tree** (MST). Weight edges by computational cost to minimise overall cost.

Two optimised degree mixing sets:

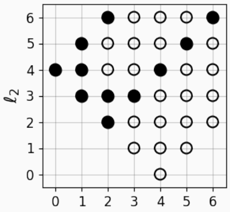
1. **MST**: reduce resolution complexity from $\mathcal{O}(L^5)$ to $\mathcal{O}(L^4)$, whilst preserving performance.
2. **Reduced MST** (logarithmic subsampling): reduce to $\mathcal{O}(L^3 \log L)$, with small but insignificant performance degradation that is offset by computational saving.

# Optimised degree mixing sets

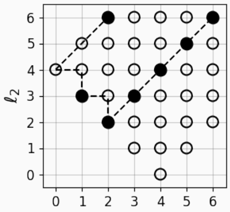
Visualization of the degree mixing set $\mathbb{P}_L^\ell$ for $L = 7$ and $\ell = 4$.




Full $\mathbb{P}_L^\ell$ 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.

# Efficient sampling theory and fast harmonic transforms

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

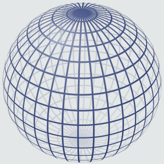
Leverage to access **underlying continuous signal representations**, avoiding discretization artifacts, and **compute fast convolutions**.

# 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**.

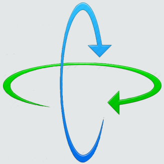
**Novel sampling theorem on sphere (McEwen & Wiaux 2011)**



SSHT: Spin spherical harmonic transforms

www.spinsht.org

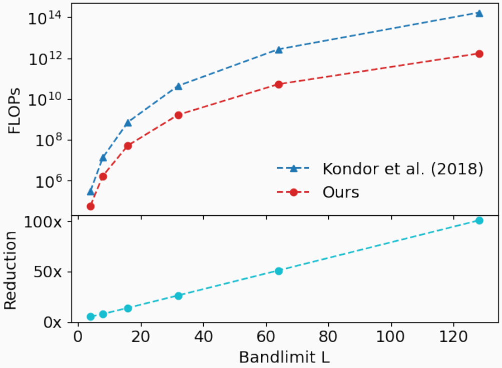
**Novel sampling theorem on rotation group (McEwen et al. 2015)**



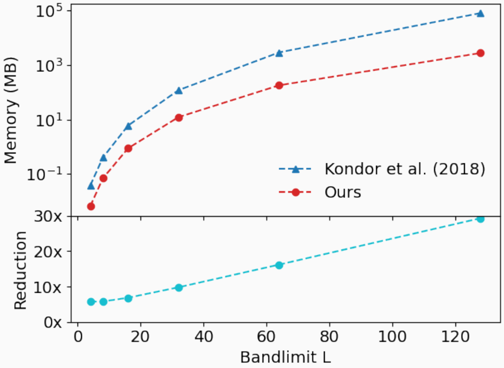
SO3: Fast Wigner transforms on rotation group

www.sothree.org

# Computational cost and memory requirements




Computational cost



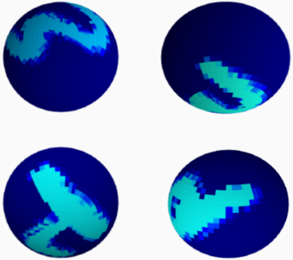

Memory requirements

# Numerical experiments

# Spherical MNIST: problem

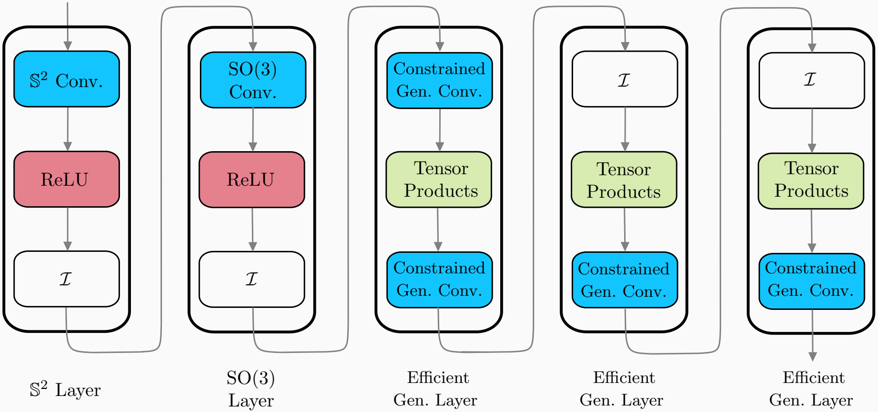
Canonical benchmark problem of classifying MNIST digits projects onto the sphere.

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** |

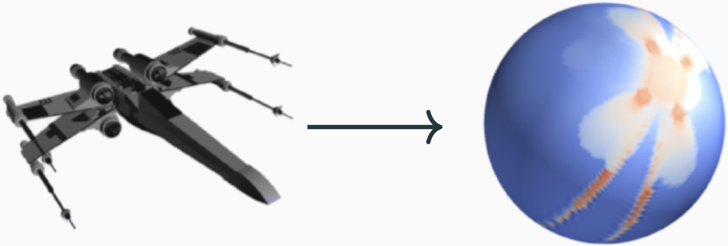
# Spherical MNIST: results

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** |

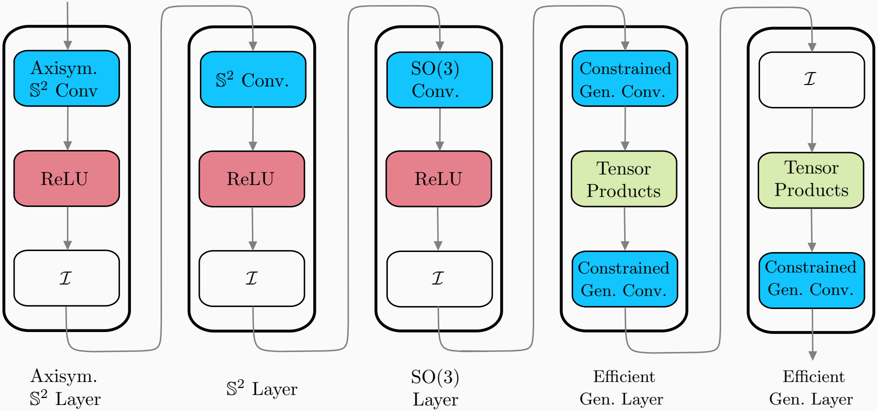
# 3D shape classification: problem

Classify 3D meshes and perform shape retrieval.



[Image credit: Esteves et al. 2018]

# 3D shape classification: architecture



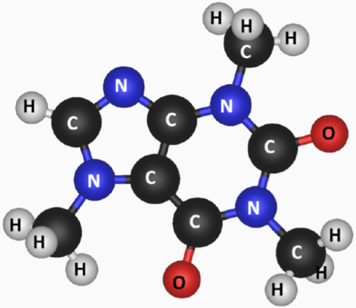
# 3D shape classification: results

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

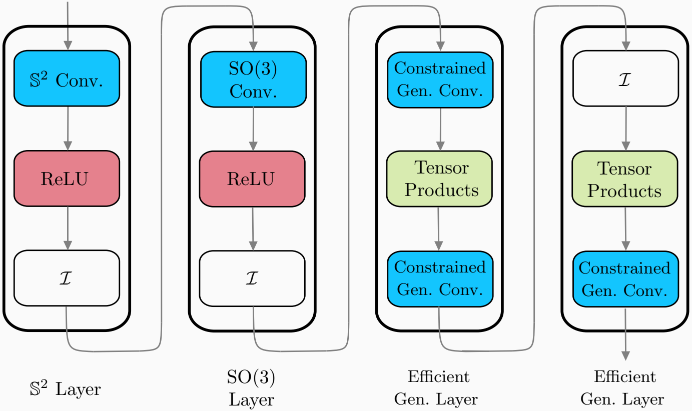
| | P@N | R@N | F1@N | mAP | NDCG | Params |
|---|---|---|---|---|---|---|
| Kondor et al. 2018 | 0.707 | 0.722 | 0.701 | 0.683 | 0.756 | >1M |
| Cohen et al. 2018 | 0.701 | 0.711 | 0.699 | 0.676 | 0.756 | 1.4M |
| Esteves et al. 2018 | 0.717 | **0.737** | - | **0.685** | - | 500k |
| Ours | **0.719** | 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: architecture



# 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** |

# 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. 2020; arXiv:2010.11661)
  - General framework that encompasses others as special cases.
  - Supports hybrid models to leverage strength of alternatives alongside each other.
  - New efficient layers to be used as primary building blocks.
  - State-of-the-art performance, both in terms of accuracy and parameter efficiency.

kagenova

www.kagenova.com

Unlocking the potential of (geometric) deep learning to solve a wide range of problems in virtual reality (VR) and beyond.

We're hiring!

hello@kagenova.com

Questions?