Efficient generalized spherical CNNs

arXiv:2010.11661

Oliver Cobb, Chris Wallis, Augustine Mavor-Parker, Augustin Marignier, Matthew Price, Mayeul d'Avezac & Jason D. McEwen*

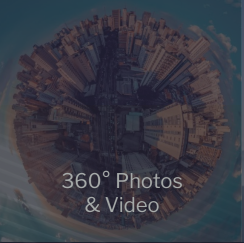
* Corresponding author: jason.mcewen@kagenova.com, www.jasonmcewen.org

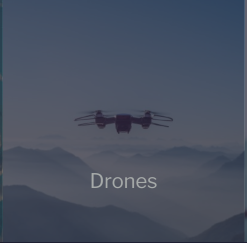
Kagenova Limited

www.kagenova.com

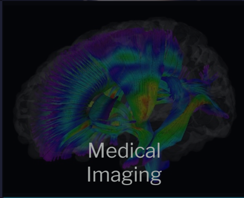
International Conference on Learning Representations (ICLR) 2021

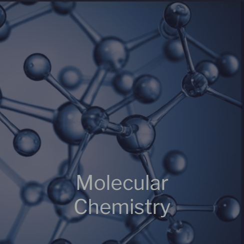
Generalized spherical CNNs

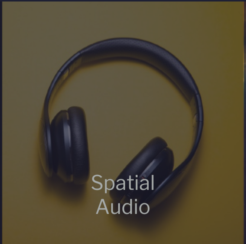


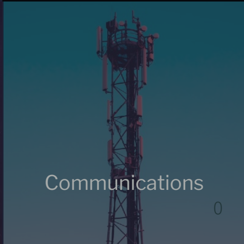


Data on the sphere arises in many applications



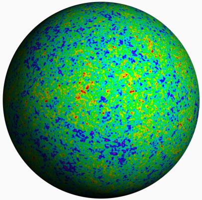






Cosmology and virtual reality

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



Cosmic microwave background

360° virtual reality

Construct CNNs natively on the sphere and encode rotational equivariance.

Generalized spherical CNNs

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

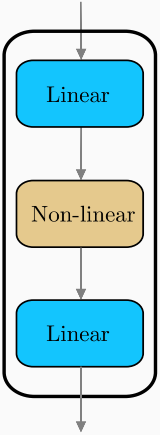
$$\mathcal{A}^{(s)} = (\mathcal{L}_1, \mathcal{N}, \mathcal{L}_2),$$

such that

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

where

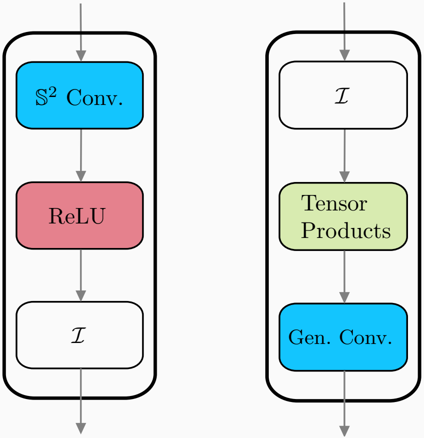
- $\mathcal{L}_1, \mathcal{L}_2 : \boldsymbol{\mathcal{F}}_L \rightarrow \boldsymbol{\mathcal{F}}_L$ are spherical convolution operators,
- $\mathcal{N} : \boldsymbol{\mathcal{F}}_L \rightarrow \boldsymbol{\mathcal{F}}_L$ is a non-linear, spherical activation operator.



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)

Efficient generalized spherical CNNs

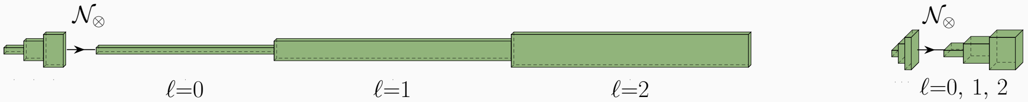
Contributions to improve efficiency

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 $\mathbb{P}_L^\ell$:

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

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

Optimized 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 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).

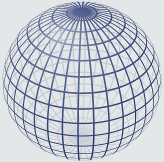
Computational complexity significantly reduced from $\mathcal{O}(L^5)$ to $\mathcal{O}(L^3 \log L)$, where L denotes resolution (bandlimit).

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.

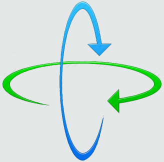
Novel sampling theorem on sphere (McEwen & Wiaux 2011)



SSHT: Spin spherical harmonic transforms

www.spinsht.org

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

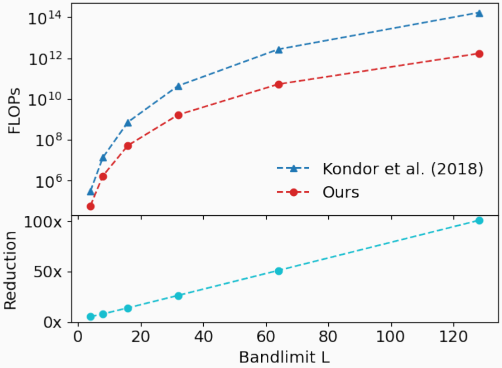


SO3: Fast Wigner transforms on rotation group

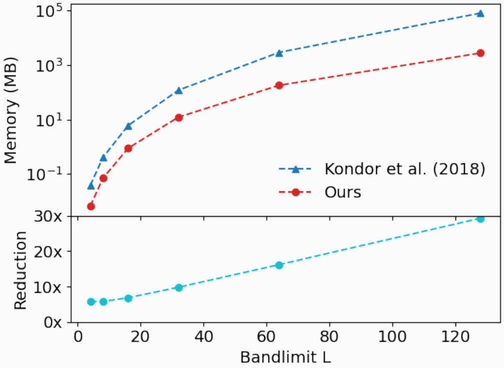
www.sothree.org

Numerical results

Computational cost and memory requirements



Computational cost



Memory requirements

Rotational equivariance

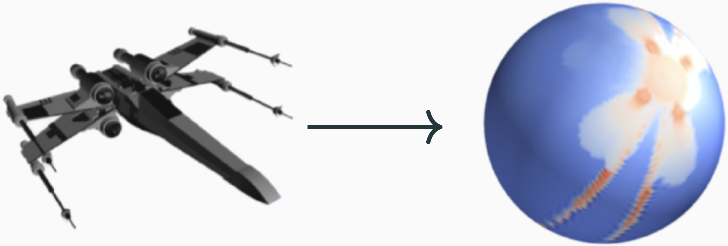
Equivariance errors

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

* Floating point precision.

3D shape classification: problem

Classify 3D meshes and perform shape retrieval.



[Image credit: Esteves et al. 2018]

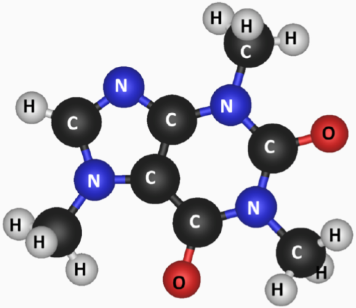
3D shape classification: results

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**

Summary

Efficient generalized spherical CNNs (Cobb et al. 2021; 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 that are strictly rotationally equivariant to be used as primary building blocks.
- State-of-the-art performance, both in terms of accuracy and parameter efficiency.

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