Mathematical Foundations of AI
1/3 September 2026
University of Oxford, UK

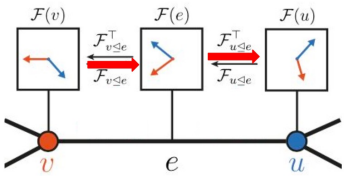
Equivariant Sheaf Neural Networks

Learning Geometric Transport on Graphs

Alessio Borgi*, **Mario Severino***, Fabrizio Silvestri, Pietro Liò

Introduction & Background

- Many physical systems are graphs embedded in the Euclidean space.
- **Existing Equivariant GNN** preserve $E(n)$ symmetry but are often **isotropic**, struggling with shear, torsion, vorticity and directional interactions.
- Standard **Sheaf Neural Networks** model anisotropic transport but do not preserve $E(n)$ equivariance.



How to move a vector x from v to u?

Moving $\mathbf{x} \in \mathcal{F}(v) \to \mathcal{F}(u)$ is done via maps composition!

$$\mathcal{F}^{\top}_{u \trianglelefteq e}\, \mathcal{F}_{v \trianglelefteq e}\, \mathbf{x} \in \mathcal{F}(u)$$

What is a Sheaf Neural Network?

- A cellular sheaf equips a graph with local feature spaces $\mathcal{F}(v), \mathcal{F}(e)$ (**stalks**) and edge-dependent **transport maps**.
- The **disagreement** between neighbouring nodes is measured after transporting their representations into a common space.

$$(\delta_{\mathcal{F}} h)_{j \to i} = \mathcal{F}_{i \trianglelefteq e} h_i - \mathcal{F}_{j \trianglelefteq e} h_j$$

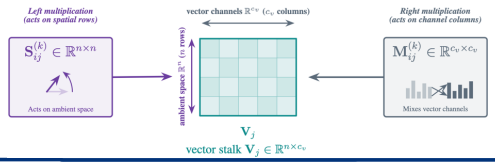
- The **Sheaf Laplacian** is:

$$L_{\mathcal{F}}(\mathbf{x})_u = \sum_{u, v \trianglelefteq e} \mathcal{F}^{\top}_{u \trianglelefteq e} (\mathcal{F}_{u \trianglelefteq e} \mathbf{x}_u - \mathcal{F}_{v \trianglelefteq e} \mathbf{x}_v)$$

Equivariant Sheaf Neural Networks

- Each node stalk contains **invariant scalar features** and **$O(n)$-covariant vector features**.
- ESNN factorises every edge map into **covariant spatial transport** and **invariant channel mixing**.
- **Restriction maps** are constrained to **preserve rotations, reflections,** and **translations**.

$$\mathcal{T}_{i \leftarrow j}(\mathbf{V}_j) = \sum_{k=1}^{K} \mathbf{S}_{ij}^{(k)} \mathbf{V}_j \mathbf{M}_{ij}^{(k)}$$



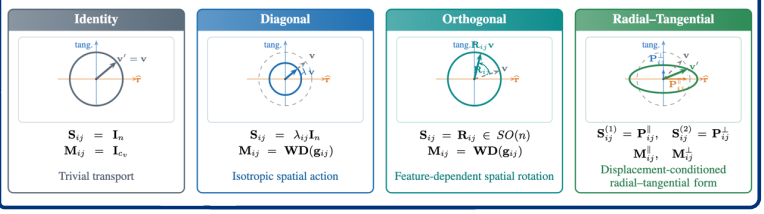
Research Question

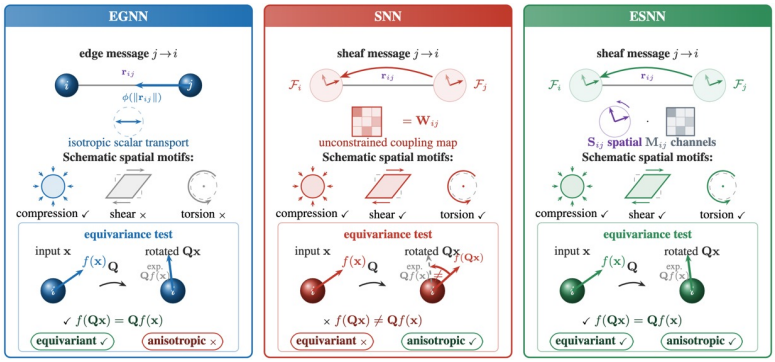
Can we combine anisotropic transport of SheafNNs with $E(n)$-equivariance of geometric GNNs?

Transport Types

- Increasingly flexible constraints recover existing equivariant models and culminate in maximally expressive first-order anisotropic transport.

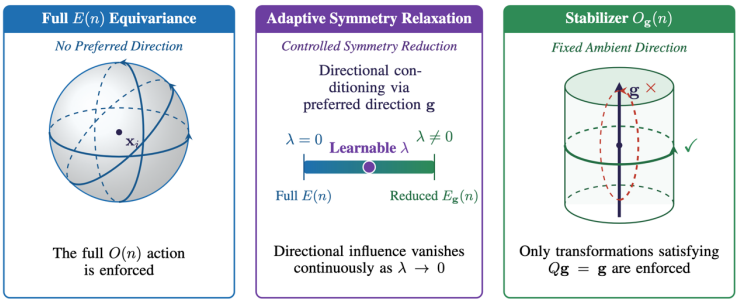
Principal ESNN spatial transport families





- **Equivariant GNNs** preserve $E(n)$-symmetry but usually rely on **isotropic radial interactions**.
- Standard **Sheaf Neural Networks** support **anisotropic transport**, but unconstrained restriction maps may violate equivariance.
- ESNN combines both properties by constraining spatial transport to be O(n)-covariant.
- ESNN messages can model compression, shear, torsion, and direction-dependent interactions while remaining equivariant.

- Full $E(n)$-equivariance is appropriate for isolated systems with no preferred direction.
- External fields, such as **gravity or uniform flow**, introduce a distinguished spatial axis.
- ESNN learns a relaxation parameter λ, preserving full equivariance when $\lambda \approx 0$.
- When the data require it, the model relaxes to cylindrical $O_g(n)$-subequivariance, preserving only transformations that leave the field direction g fixed.



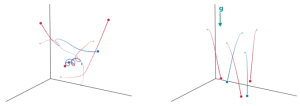
Experimental Results

(a) *N-body*

Method	MSE ↓
Linear	0.0819
SE(3) Transformer	0.0244
Tensor Field Network	0.0155
Graph Neural Network	0.0107
Radial Field	0.0104
EGNN	0.0071
ESNN-Id	0.0060
ESNN-Diag	0.0054
ESNN-Ortho	**0.0051**
ESNN-RadTan	**0.0052**

(b) *N-body + Gravity* – *Soft Symmetry*

Mode	Transport	MSE ↓	$\max_\ell \lvert\lambda_\ell^{(t)}\rvert \lVert g \rVert_2$	A_g ↑
None	ESNN-Diag	0.107160 ± 0.034566	–	–
	ESNN-Ortho	0.125181 ± 0.048970	–	–
	ESNN-RadTan	0.101860 ± 0.032444	–	–
Learned	ESNN-Diag	0.020548 ± 0.001641	0.267400 ± 0.070406	1.000
	ESNN-Ortho	0.019687 ± 0.001452	0.400903 ± 0.072151	1.000
	ESNN-RadTan	0.021276 ± 0.003345	2.055211 ± 0.633743	1.000
Fixed	ESNN-Diag	0.020838 ± 0.002625	0.346428 ± 0.035724	1.000
	ESNN-Ortho	0.019823 ± 0.001651	0.609956 ± 0.177071	1.000
	ESNN-RadTan	0.023381 ± 0.005414	1.379156 ± 0.301397	1.000



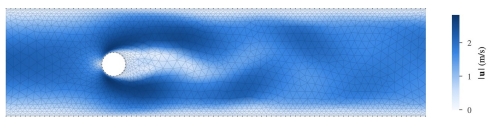
QM9 – Molecular Property Prediction Benchmark

Method / Units	α m_0^3	$\Delta\epsilon$ meV	ϵ_{HOMO} meV	ϵ_{LUMO} meV	μ mD	C_v $\text{cal mol}^{-1}\,\text{K}^{-1}$	G meV	H meV	$\langle R^2 \rangle$ m_0^2	U meV	U_0 meV	ZPVE meV
Invariant models												
Cormorant	85	61	34	38	38	26	20	21	961	21	22	2.03
NMP	92	69	43	38	30	40	19	17	180	20	20	1.50
DimeNet++†	44	32.6	24.6	19.5	29.7	23	7.56	6.53	331	6.28	6.32	1.21
ComENet†	45	32.4	23.1	19.8	24.5	22	7.98	6.86	259	6.82	6.69	1.20
SphereNet†	46	31.1	22.8	18.9	24.5	22	7.78	6.33	268	6.36	6.26	1.12
Scalarization-based models												
ClofNet	63	53	33	25	40	27	9	9	610	9	8	1.23
EGNN	71	48	29	25	29	31	12	12	106	12	11	1.55
PaiNN†	45	45.7	27.6	20.4	12.0	24	7.35	5.98	66	5.83	5.85	1.28
LEFTNet	48	40	24	18	12	23	7	6	109	7	6	1.33
EQGAT	53	32	20	16	11	24	23	24	382	25	25	2.00
ET	59	36.1	20.3	17.5	11	26	7.62	6.16	33	6.38	6.15	1.84
Geoformer	40	33.8	18.4	15.4	10	22	6.13	4.39	28	4.41	4.43	1.28
SaVeNet-B†	39	24.8	18.4	16.3	9.3	23	6.64	5.43	58	5.48	5.43	1.18
Equivariant Sheaf Neural Networks												
ESNN-Id	63.08	65.35	43.13	28.52	58.81	37.83	15.76	18.71	749.44	–	10.54	1.54
ESNN-Diag	62.16	43.69	27.44	21.43	17.15	**31.57**	**11.48**	**12.71**	134.99	11.77	13.57	1.62
ESNN-Ortho	79.42	43.35	**25.10**	23.17	21.77	36.50	17.76	26.07	323.52	18.59	25.25	1.94
ESNN-RadTan	**60.32**	**39.01**	25.16	**19.17**	**16.08**	**31.57**	14.77	13.86	**119.11**	**10.63**	**10.34**	**1.45**

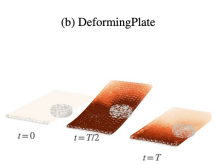
MeshGraphNets – Mesh-based Physical Dynamics

Method	CylinderFlow 1-Step	CylinderFlow 50-Step	CylinderFlow Full	DeformingPlate 1-Step	DeformingPlate 50-Step	DeformingPlate Full	Airfoil 1-Step	Airfoil 50-Step	Airfoil Full
MeshGraphNets	2.34 ± 0.12	6.3 ± 0.7	40.88 ± 7.2	0.25 ± 0.05	1.8 ± 0.5	15.1 ± 4.0	314 ± 36	582 ± 37	11529 ± 1203
ESNN-Id	3.96	15.8	69.6	0.15	1.3	7.7	5529	9910	17974
ESNN-Diag	2.82	10.3	37.04	0.17	1.4	9.2	5130	9616	11598
ESNN-Ortho	2.40	7.7	**35.94**	0.13	1.1	8.1	3014	4722	10328
ESNN-RadTan	3.08	10.9	49.27	0.08	1.0	**5.8**	2584	3346	**7787**

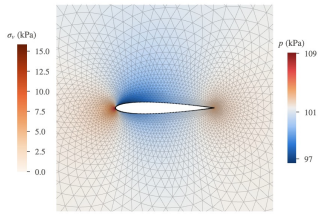
(a) CylinderFlow

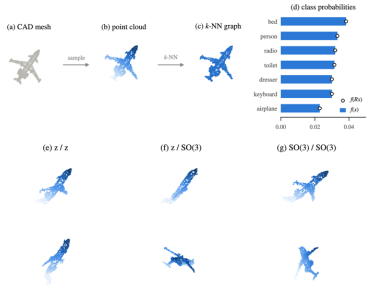


(b) DeformingPlate



(c) Airfoil



ModelNet40 – Classification under Rotations

Method	z/z	z/SO(3)	SO(3)/SO(3)	Δ_{OOD} ↓
PointNet	85.9	19.6	74.7	66.3
RS-CNN	90.3	48.7	82.6	41.6
DGCNN	90.3	33.8	88.6	56.5
RI-Conv	86.5	86.4	86.4	0.1
GC-Conv	89.0	89.1	89.2	0.1
Luo et al. DGCNN	88.4	88.4	88.9	0.0
LGR-Net	90.9	90.9	91.1	0.0
Li et al. (w/ TTA)	91.6	91.6	91.6	0.0
CRIN	91.8	91.8	91.8	0.0
TFN	88.5	85.3	87.6	3.2
VN-PointNet	77.5	77.5	77.2	0.0
VN-DGCNN	89.5	89.5	90.2	0.0
ESNN-Id	84.684	85.737	85.575	1
ESNN-Diag	84.603	84.319	84.400	0.2
ESNN-Ortho	84.927	85.170	84.684	0.2
ESNN-RadTan	85.373	84.643	86.264	0.7