EQUIVARIANT SHEAF NEURAL NETWORKS: LEARNING GEOMETRIC TRANSPORT ON GRAPHS

Anonymous authors
Paper under double-blind review

ABSTRACT

Equivariant graph neural networks constrain how geometric features transform, but efficient first-order architectures remain limited in how vector information can be transported across graph edges. We introduce ESNN, an Equivariant Sheaf Neural Network that learns directed, matrix-valued transport between neighboring vector features while preserving exact Euclidean equivariance. Rather than increasing representation order, ESNN retains first-order scalar and vector features and places additional geometric expressivity in the edge interactions themselves. Its transport mechanisms range from isotropic transformations to feature-conditioned rotations and anisotropic radial–tangential actions, allowing vector messages to be transformed according to the local geometry before aggregation. We also equip ESNN with a controlled symmetry-relaxation mechanism for systems with a preferred ambient direction, preserving exact equivariance to the remaining symmetry while recovering full $E(n)$-equivariance as a special case. Across diverse geometric learning problems spanning physical dynamics, point clouds, and molecular systems, ESNN improves predictive performance over first-order equivariant baselines, learns a latent preferred direction from data, and remains robust to unseen rotations. These results establish edge-wise geometric transport as a source of equivariant expressivity that does not require higher-order representations.

1 INTRODUCTION

Many physical systems, from molecular dynamics to fluid mechanics, are naturally represented as geometric graphs embedded in Euclidean space (Satorras et al., 2021; Brandstetter et al., 2022; Du et al., 2023; Wang et al., 2024b; Aykent & Xia, 2025). Their states may contain invariant scalar attributes together with vectors or tensors that transform with the geometry (Schütt et al., 2021; Simeon & De Fabritiis, 2023; Liao et al., 2024). Equivariant neural networks encode these transformation laws directly, ensuring predictable behavior under translations, rotations, and, when appropriate, reflections. Equivariance alone, however, does not determine how geometric information should be transformed as it passes between neighboring nodes. Many interactions are intrinsically directional: longitudinal and transverse responses, shear, and anisotropic propagation depend on local orientation rather than on distance alone. Existing geometric architectures capture this structure in different ways. Cartesian models work efficiently with invariant scalars and first-order vectors (Satorras et al., 2021; Schütt et al., 2021), while steerable models enrich the propagated representations through higher-order features, spherical harmonics, and tensor products (Thomas et al., 2018; Batatia et al., 2025; 2022). We instead retain a first-order scalar–vector representation and increase expressivity through richer geometric transport of vector features. This perspective is closely related to Sheaf Neural Networks (SNNs), where interactions between local feature spaces are mediated by matrix-valued maps (Hansen & Gebhart, 2020; Bodnar et al., 2022). Generic sheaf maps, however, are not inherently constrained to respect the transformation laws of geometric vector features. ESNN addresses this by constraining matrix-valued edge transport to satisfy Euclidean equivariance. Each transport separates an $O(n)$-covariant spatial action from invariant mixing across vector channels, allowing the geometric transformation applied to a message to depend on the local interaction while retaining scalar–vector representations throughout the network. We instantiate this framework through a family of transport mechanisms, from identity and isotropic cases to feature-conditioned rotations and anisotropic radial–tangential transport. ESNN also accommodates systems with a preferred ambient direction, such as those governed by gravity or imposed flow. This direction may be specified from prior knowledge or learned from data. The resulting model remains exactly equivariant to the

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

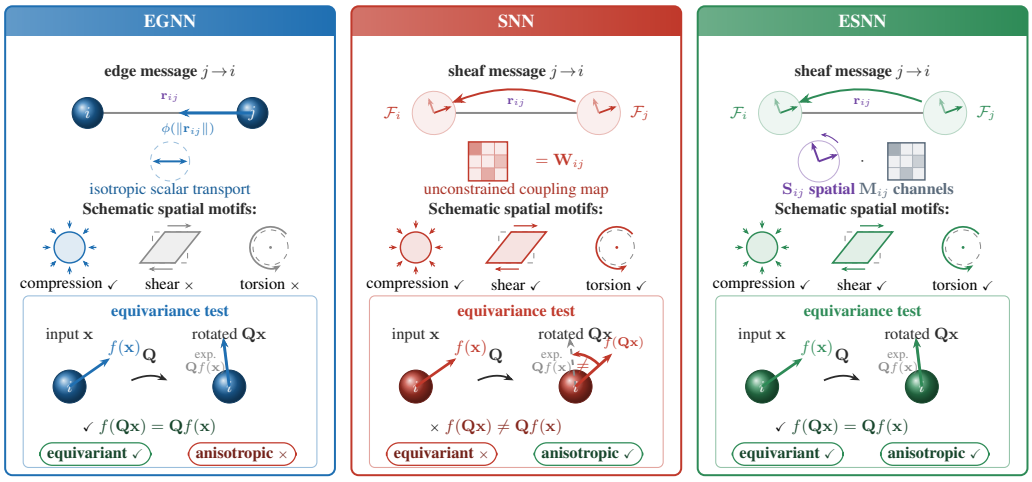

Figure 1: *Edge-wise geometric transport in EGNN-style models, generic sheaf networks, and ESNNs.* EGNN-style models couple invariant weights with relative displacements, while sheaf networks allow matrix-valued maps between local feature spaces. ESNN combines matrix-valued transport with exact equivariance through $O(n)$-covariant spatial actions and invariant channel mixing.

subgroup that preserves the preferred direction, while full $E(n)$-equivariance is recovered when the directional conditioning is inactive. Figure 1 contrasts EGNN-style isotropic message passing, unconstrained sheaf transport, and ESNN's equivariant matrix-valued transport.

Our main contributions are summarized as follows: **(i)** We introduce *Equivariant Sheaf Neural Networks*, a first-order Cartesian framework that learns matrix-valued transport of vector features across graph edges while preserving $E(n)$-equivariance. **(ii)** We develop a family of equivariant transport mechanisms that progressively enrich how vector features interact across edges, from identity and isotropic transport to feature-conditioned orthogonal transformations and anisotropic radial–tangential transport. **(iii)** We extend ESNN to systems with a preferred ambient direction, while retaining exact equivariance to the corresponding stabilizer subgroup, and evaluate the framework across particle dynamics, mesh-based physical simulation, point-cloud classification, and molecular property prediction.

2 Related Work and Background

Related Work. ESNN combines ideas from equivariant message passing and sheaf-based graph learning. Cartesian architectures such as EGNN (Satorras et al., 2021) and scalar–vector models such as PaiNN and GVP (Schütt et al., 2021; Jing et al., 2021) achieve efficient $E(n)$-equivariance using invariant scalars together with low-order covariant features. Steerable architectures instead propagate higher-order representations and couple them through equivariant tensor products (Thomas et al., 2018; Fuchs et al., 2020; Brandstetter et al., 2022; Batatia et al., 2022; Liao et al., 2024). ESNN remains in the first-order scalar–vector setting, but increases geometric expressivity by enriching the transport applied to vector features rather than by introducing higher-order representations. The transport viewpoint is closely related to SNNs, where scalar edge weights are replaced by linear maps between local feature spaces (Hansen & Gebhart, 2020; Bodnar et al., 2022). Subsequent work has explored more structured forms of transport, including orthogonal connection maps for geometric propagation (Barbero et al., 2022a) and directed and asymmetric sheaf propagation (Ribeiro et al., 2026; Fiorini et al., 2026). Closely related copresheaf formulations place similar emphasis on directed maps and their composition across the graph (Hajij et al., 2025). ESNN adopts this matrix-valued view of message transport, but constrains the learned maps to respect the Euclidean transformation laws of geometric vector features. A complementary line of work studies settings in which external structure reduces the symmetry of the problem. Subequivariant models enforce the stabilizer of a prescribed direction (Han et al., 2022), while relaxed-equivariant approaches learn controlled departures from an underlying symmetry (Wang et al., 2022; Hofgard et al., 2024). ESNN incorporates a preferred direction within the transport framework, retaining exact stabilizer equivariance and recovering full $E(n)$-equivariance when the directional conditioning is inactive. Further discussion of related work is provided in Appendix A.

2.1 GEOMETRIC BACKGROUND

We consider a physical system represented by a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where each node $i \in \mathcal{V}$ has coordinates $\mathbf{x}_i \in \mathbb{R}^n$. We seek message-passing operators that respect Euclidean symmetries while allowing interactions to depend on local geometry. We first specify the relevant feature transformations and then introduce the cellular-sheaf viewpoint that motivates the transport construction of Section 3.

Equivariance. The Euclidean group $E(n) = O(n) \ltimes \mathbb{R}^n$ acts on coordinates and relative displacements as:

$$\mathbf{x}_i \mapsto Q\mathbf{x}_i + \mathbf{t}, \qquad \mathbf{r}_{ij} = \mathbf{x}_i - \mathbf{x}_j \mapsto Q\mathbf{r}_{ij}, \qquad Q \in O(n), \qquad \mathbf{t} \in \mathbb{R}^n \tag{1}$$

Relative displacements are therefore translation invariant, and operations built from relative geometry only need to preserve the corresponding $O(n)$ transformation laws.

Feature Representation. Each node carries invariant scalar features $\mathbf{s}_i \in \mathbb{R}^{c_s}$ and covariant vector features $\mathbf{V}_i \in \mathbb{R}^{n \times c_v}$, collected as:

$$\mathbf{h}_i = (\mathbf{s}_i, \mathbf{V}_i), \qquad \mathbf{s}_i \mapsto \mathbf{s}_i, \qquad \mathbf{V}_i \mapsto Q\mathbf{V}_i \tag{2}$$

The corresponding local feature space is $\mathcal{F}(i) = \mathbb{R}^{c_s} \oplus (\mathbb{R}^n \otimes \mathbb{R}^{c_v})$, which we interpret below as the node *stalk*. ESNN transports the vector component geometrically, while scalar features are propagated through invariant operations.

Cellular Sheaves and Transport. Standard message passing implicitly assumes that neighboring features can be compared and aggregated in a common space. Cellular sheaves generalize this picture by assigning local vector spaces to nodes and edges together with linear restriction maps (Hansen & Gebhart, 2020; Bodnar et al., 2022). For an edge $e = \{i, j\}$, a cellular sheaf provides an edge stalk $\mathcal{F}(e)$ and maps $\rho_{i \to e} : \mathcal{F}(i) \to \mathcal{F}(e)$ and $\rho_{j \to e} : \mathcal{F}(j) \to \mathcal{F}(e)$, which express the two node features in a common edge space. These maps define the cellular coboundary and the associated sheaf Laplacian $\mathbf{L}_{\mathcal{F}} = \delta_{\mathcal{F}}^* \delta_{\mathcal{F}}$. For an edge $e = \{i, j\}$, its off-diagonal term induces, up to the conventional minus sign, the node-to-node map:

$$\rho_{i \to e}^* \rho_{j \to e} : \mathcal{F}(j) \to \mathcal{F}(i) \tag{3}$$

This induced coupling provides the starting point for the ESNN construction. Rather than introducing separate incidence maps through an intermediate edge stalk, ESNN directly parameterizes transport between neighboring node features. For geometric vector channels, these maps must additionally respect the ambient $O(n)$ action, which motivates the spatial–channel formulation developed in Section 3. The connection to classical sheaf diffusion and directed transport is developed further in Appendix C.

3 EQUIVARIANT SPATIAL TRANSPORT

The vector features introduced in Section 2.1 naturally decompose into two factors: a spatial component $\mathbb{R}^n$, which carries the $O(n)$ action, and a channel component $\mathbb{R}^{c_v}$, which remains invariant under the group. ESNN exploits this structure by factorizing each edge transport into a spatial action on geometric vectors and a channel action across vector features. Let $\mathcal{C}_{ij}$ denote the local context for the directed interaction $j \to i$, including the relative displacement $\mathbf{r}_{ij}$ and, when used, covariant features at the incident nodes. We write $Q \cdot \mathcal{C}_{ij}$ for the simultaneous action of $Q \in O(n)$ on all covariant quantities in this context.

Definition 3.1 (*Equivariant Transport Map*)**.** For a directed interaction $j \to i$ and a fixed local edge context, ESNN defines a transport map $\mathcal{T}_{i \leftarrow j} : \mathbb{R}^{n \times c_v} \longrightarrow \mathbb{R}^{n \times c_v}$ of the form:

$$\mathcal{T}_{i \leftarrow j}(\mathbf{V}_j) = \sum_{k=1}^{K} \mathbf{S}_{ij}^{(k)} \mathbf{V}_j \mathbf{M}_{ij}^{(k)} \tag{4}$$

where K denotes the number of components in the transport expansion. Each term combines two actions:

- The *spatial operator* $\mathbf{S}_{ij}^{(k)} \in \mathbb{R}^{n \times n}$ acts on the spatial axis and transforms covariantly:

$$\mathbf{S}_{ij}^{(k)}(Q \cdot \mathcal{C}_{ij}) = Q\, \mathbf{S}_{ij}^{(k)}(\mathcal{C}_{ij}) Q^\top \qquad \forall Q \in O(n) \tag{5}$$

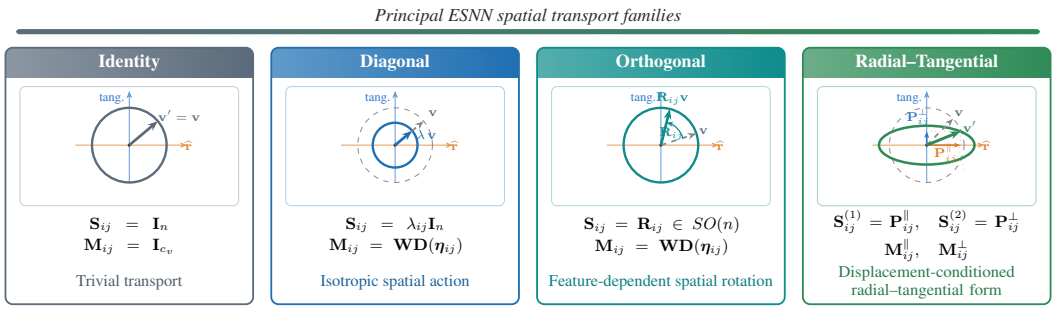

Figure 2: *ESNN transport families.* Identity leaves the vector representation unchanged; Diagonal applies an edge-dependent isotropic spatial scaling; Orthogonal learns a feature-conditioned spatial rotation; and Radial–Tangential acts independently on components parallel and orthogonal to the relative displacement.

- The *channel operator* $\mathbf{M}_{ij}^{(k)} \in \mathbb{R}^{c_v \times c_v}$ mixes vector channels and is constructed from $O(n)$-invariant edge features:

$$\mathbf{M}_{ij}^{(k)}(Q \cdot \mathcal{C}_{ij}) = \mathbf{M}_{ij}^{(k)}(\mathcal{C}_{ij}) \qquad \forall Q \in O(n) \tag{6}$$

These two transformation laws are sufficient to make the resulting transport equivariant.

Proposition 3.2 (Equivariance of the Transport Map)**.** *Suppose that, for every component k, the spatial operator satisfies:*

$$\mathbf{S}_{ij}^{(k)}(Q \cdot \mathcal{C}_{ij}) = Q\mathbf{S}_{ij}^{(k)}(\mathcal{C}_{ij})Q^{\top} \tag{7}$$

and the channel operator $\mathbf{M}_{ij}^{(k)}$ is unchanged under the $O(n)$ action. Then the transport map in Equation 4 is $O(n)$-equivariant:

$$\mathcal{T}'_{i \leftarrow j}(Q\mathbf{V}_j) = Q\,\mathcal{T}_{i \leftarrow j}(\mathbf{V}_j) \tag{8}$$

where $\mathcal{T}'_{i \leftarrow j}$ denotes the transport evaluated from the transformed geometric inputs. (Proof is provided in Appendix D.1).

The transport associated with each edge is conditioned on the local geometric context, which may include both invariant quantities and covariant features. This allows the operator in Equation 4 to adapt to the current interaction while preserving the scalar–vector representation of the network. In ESNN, the invariant channel maps are parameterized as:

$$\mathbf{M}_{ij}^{(k)} = \mathbf{W}^{(k)}\mathbf{D}\left(\boldsymbol{\eta}_{ij}^{(k)}\right) \tag{9}$$

where $\mathbf{W}^{(k)}$ is shared across edges and $\boldsymbol{\eta}_{ij}^{(k)} \in \mathbb{R}^{c_v}$ is an edge-dependent gating vector predicted from invariant edge features, with $\mathbf{D}(\boldsymbol{\eta}_{ij}^{(k)})$ denoting the corresponding diagonal matrix. The shared transformation mixes vector channels, while the gate adapts this mixing to the local interaction. In this way, Equation 4 separates channel mixing from geometric transformation: the channel maps act across vector features, while the spatial operators determine how those features are transformed in the ambient space. Section 4 introduces the resulting spatial transport families, with further details on the decomposition provided in Appendix B.

4 Equivariant Transport Maps

We now instantiate the spatial–channel decomposition of Section 3 through four transport families, illustrated in Figure 2. They differ in the spatial action applied to a vector message, while their learned coefficients and channel maps are constructed from $O(n)$-invariant edge features.

Identity Transport. The simplest choice leaves the vector message unchanged:

$$\mathcal{T}_{i \leftarrow j}(\mathbf{V}_j) = \mathbf{V}_j \tag{10}$$

corresponding to $\mathbf{S}_{ij} = \mathbf{I}_n$ and $\mathbf{M}_{ij} = \mathbf{I}_{c_v}$. Identity Transport therefore provides the reference case in which vector features are exchanged without an edge-dependent transformation.

Diagonal Transport. A first extension allows edge-dependent modulation while remaining isotropic in physical space. We set $\mathbf{S}_{ij} = \lambda_{ij}\mathbf{I}_n$, where λ_{ij} is predicted from invariant edge features, giving:

$$\mathcal{T}_{i\leftarrow j}(\mathbf{V}_j) = \lambda_{ij}(\mathbf{V}_j\mathbf{W})\mathbf{D}(\boldsymbol{\eta}_{ij}) \tag{11}$$

The same spatial scaling is applied in every direction, while the channel map remains adaptive to the local interaction.

Orthogonal Transport. Diagonal Transport remains spatially isotropic. Orthogonal Transport instead derives a feature-dependent rotation from the vector states at the two endpoints of an interaction. We define:

$$\mathbf{C}_{ij} = \mathbf{V}_j\mathbf{V}_i^\top, \qquad \widetilde{\mathbf{C}}_{ij} = \frac{\mathbf{C}_{ij}}{\|\mathbf{C}_{ij}\|_F + \varepsilon}, \qquad \boldsymbol{\Omega}_{ij} = \widetilde{\mathbf{C}}_{ij} - \widetilde{\mathbf{C}}_{ij}^\top \tag{12}$$

and use an invariant edge coefficient β_{ij} to form:

$$\mathbf{R}_{ij} = \exp(\beta_{ij}\boldsymbol{\Omega}_{ij}) \in SO(n), \qquad \mathcal{T}_{i\leftarrow j}(\mathbf{V}_j) = \mathbf{R}_{ij}(\mathbf{V}_j\mathbf{W})\mathbf{D}(\boldsymbol{\eta}_{ij}) \tag{13}$$

Since $\mathbf{R}_{ij}$ is inferred from the evolving vector features, the spatial transformation adapts to the local feature geometry rather than to relative displacement alone. Under a global orthogonal transformation, $\boldsymbol{\Omega}_{ij}$ and $\mathbf{R}_{ij}$ transform by conjugation, so Proposition 3.2 applies; the derivation is given in Appendix D.2. Here *Orthogonal* refers to the spatial factor $\mathbf{R}_{ij}$, while the full edge map also includes learned channel mixing and gating.

Radial–Tangential Transport. Many geometric interactions distinguish motion along an edge from motion in the transverse directions. For $\mathbf{r}_{ij} \neq \mathbf{0}$, let $\widehat{\mathbf{r}}_{ij} = \mathbf{r}_{ij}/\|\mathbf{r}_{ij}\|$ and define:

$$\mathbf{P}_{ij}^{\parallel} = \widehat{\mathbf{r}}_{ij}\widehat{\mathbf{r}}_{ij}^\top, \qquad \mathbf{P}_{ij}^{\perp} = \mathbf{I}_n - \mathbf{P}_{ij}^{\parallel} \tag{14}$$

Radial–Tangential Transport applies independent channel transformations to the radial and tangential components:

$$\mathbf{M}_{ij}^{\parallel} = \mathbf{W}_{\parallel}\mathbf{D}(\boldsymbol{\eta}_{ij}^{\parallel}), \qquad \mathbf{M}_{ij}^{\perp} = \mathbf{W}_{\perp}\mathbf{D}(\boldsymbol{\eta}_{ij}^{\perp}) \tag{15}$$

and transports:

$$\mathcal{T}_{i\leftarrow j}(\mathbf{V}_j) = \mathbf{P}_{ij}^{\parallel}\mathbf{V}_j\mathbf{M}_{ij}^{\parallel} + \mathbf{P}_{ij}^{\perp}\mathbf{V}_j\mathbf{M}_{ij}^{\perp} \tag{16}$$

Both projectors transform by conjugation under $O(n)$, so the transport satisfies Proposition 3.2; the explicit proof is given in Appendix D.3. The projectors need not be formed explicitly, and their action scales linearly with the ambient dimension n. When relative displacement is the only covariant geometric input, the radial–tangential form is also complete.

Theorem 4.1 (Displacement-Conditioned Equivariant Transport)**.** *Let $n \geq 2$, and consider a linear transport on $\mathbb{R}^n \otimes \mathbb{R}^{c_v}$ whose only covariant geometric conditioning variable is a non-zero relative displacement $\mathbf{r}_{ij}$, with arbitrary additional $O(n)$-invariant scalar conditioning. Every such $O(n)$-equivariant transport can be written as:*

$$\mathcal{T}_{i\leftarrow j}(\mathbf{V}_j) = \mathbf{P}_{ij}^{\parallel}\mathbf{V}_j\mathbf{A}_{ij} + \mathbf{P}_{ij}^{\perp}\mathbf{V}_j\mathbf{B}_{ij} \tag{17}$$

for invariant channel endomorphisms $\mathbf{A}_{ij}, \mathbf{B}_{ij} \in \mathbb{R}^{c_v\times c_v}$. The proof is given in Appendix B.2.

Thus, with relative displacement as the only covariant geometric input, every admissible spatial action decomposes into radial and tangential components, each paired with an invariant channel map. ESNN instantiates this structure through the parameterization in Equation 9. Additional learned covariant features can support richer spatial operators, as described in Appendix B.3.

5 The ESNN Architecture

An ESNN layer combines the equivariant transports of Section 4 with invariant scalar messaging and, when required, an equivariant update of the geometry. At node i, the hidden state consists of invariant scalar features $\mathbf{s}_i \in \mathbb{R}^{c_s}$ and covariant vector features $\mathbf{V}_i \in \mathbb{R}^{n\times c_v}$.

Input representations. Scalar descriptors are embedded in the invariant branch. When geometric vector quantities are available directly from the input, such as velocities, they are used to initialize $\mathbf{V}_i$. For tasks without native vector inputs, we construct task-specific covariant lifts $\mathbf{L}_i$ from the input geometry $\mathbf{X}$ satisfying:

$$\mathbf{L}_i(Q\mathbf{X} + \mathbf{t}) = Q\,\mathbf{L}_i(\mathbf{X}) \tag{18}$$

so that the vector representation obeys the correct transformation law. The lifts used for each task are described in Appendix E.

Edge context and message transport. For each directed interaction $j \to i$, ESNN constructs an invariant edge representation:

$$\mathbf{z}_{ij} = \mathrm{Inv}\big(\mathbf{s}_i, \mathbf{s}_j, \mathbf{V}_i, \mathbf{V}_j, \mathbf{r}_{ij}, \mathbf{e}_{ij}^{\mathrm{attr}}\big) \tag{19}$$

from distances, vector norms, and inner products, together with optional invariant edge attributes. This context determines the gates and coefficients of the transport maps introduced in Section 4. Vector and scalar messages are then given by:

$$\mathbf{m}^V_{i\leftarrow j} = \mathcal{T}_{i\leftarrow j}(\mathbf{V}_j), \qquad \mathbf{m}^{\mathrm{s}}_{i\leftarrow j} = \mathbf{s}_j + \phi_s(\mathbf{z}^{\mathrm{s}}_{ij}) \tag{20}$$

where ϕ_s is a learnable map acting on the invariant edge representation. Geometric information can therefore influence both scalar and vector messages while preserving their respective transformation laws.

Transport aggregation. Messages are aggregated using invariant normalized weights ν_{ij}, with ν_{ii} denoting the identity self-loop contribution:

$$\mathbf{V}_i^{\mathrm{agg}} = \nu_{ii}\mathbf{V}_i + \sum_{j\in\mathcal{N}(i)} \nu_{ij}\,\mathbf{m}^V_{i\leftarrow j}, \qquad \mathbf{s}_i^{\mathrm{agg}} = \nu_{ii}\mathbf{s}_i + \sum_{j\in\mathcal{N}(i)} \nu_{ij}\,\mathbf{m}^{\mathrm{s}}_{i\leftarrow j} \tag{21}$$

By default, ν_{ij} are degree-normalized edge weights, though invariant attention weights can be used in the same way. Because the two directions of an edge may carry different transport maps, the resulting operator is generally directional and need not be self-adjoint. With symmetric aggregation weights, adjoint transport across opposite directions yields a self-adjoint specialization. Its relation to sheaf and quiver operators is discussed further in Appendix C.

Feature update. The aggregated representations are updated within their respective feature spaces:

$$\mathbf{V}_i' = \mathbf{V}_i + \sigma_V(\mathbf{V}_i^{\mathrm{agg}}\mathbf{W}_V)\,, \qquad \mathbf{s}_i' = \mathbf{s}_i + \sigma_s(\mathbf{W}_s\mathbf{s}_i^{\mathrm{agg}}) \tag{22}$$

where $\mathbf{W}_V$ acts only across vector channels, $\sigma_V(\mathbf{v}) = a(\|\mathbf{v}\|)\mathbf{v}$ is applied channel-wise, and σ_s denotes the scalar nonlinearity. The channel mixing and radial nonlinearity preserve the vector transformation law, while the scalar branch remains invariant.

Coordinate update. For tasks in which the geometry evolves over time, ESNN optionally updates the node coordinates through an invariantly weighted sum of relative displacements:

$$\Delta\mathbf{x}_i = \frac{1}{|\mathcal{N}(i)|}\sum_{j\in\mathcal{N}(i)} \gamma_{ij}\mathbf{r}_{ij}, \qquad \mathbf{x}_i' = \mathbf{x}_i + \Delta\mathbf{x}_i \tag{23}$$

where γ_{ij} is an invariant scalar predicted from node and edge features. This preserves a simple first-order equivariant coordinate update while allowing the latent vector interactions to use the richer transport mechanisms introduced above. Task-specific dynamical updates are described in Appendix E.

These operations preserve the appropriate transformation laws of the scalar, vector, and coordinate features, yielding an $E(n)$-equivariant ESNN layer.

Theorem 5.1 ($E(n)$-Equivariance)**.** *Assume that the graph topology is fixed or constructed from $E(n)$-invariant quantities, that edge attributes are $O(n)$-invariant, and that the transport maps satisfy Proposition 3.2. Assume also that $\mathbf{r}_{ij} \neq \mathbf{0}$ whenever a normalized edge direction is used. In the absence of the symmetry-relaxing inputs introduced in Section 6, the ESNN layer is $E(n)$-equivariant:*

$$(\mathbf{x}_i, \mathbf{V}_i, \mathbf{s}_i) \mapsto (Q\mathbf{x}_i + \mathbf{t}, Q\mathbf{V}_i, \mathbf{s}_i) \quad \Longrightarrow \quad (\mathbf{x}_i', \mathbf{V}_i', \mathbf{s}_i') \mapsto (Q\mathbf{x}_i' + \mathbf{t}, Q\mathbf{V}_i', \mathbf{s}_i') \tag{24}$$

The proof is given in Appendix D.5.

Theorem 5.1 shows that richer edge-wise transport can be introduced without sacrificing the Euclidean equivariance of the architecture.

6 Controlled Symmetry Relaxation

Full $E(n)$-equivariance is appropriate when the ambient space has no distinguished direction. External structure such as gravity or background flow can instead reduce the symmetry of the dynamics (Han et al., 2022; Wang et al., 2024a). ESNN incorporates this directional structure through an additional scalar conditioning while leaving the underlying spatial transport families unchanged.

Let $\mathbf{g} \in \mathbb{R}^n$ denote a global preferred direction, either specified a priori or learned as a model parameter. At layer ℓ, the invariant edge context is augmented with:

$$z_{ij}^{\mathbf{g},(\ell)} = \lambda_{\mathbf{g}}^{(\ell)} \langle \mathbf{r}_{ij}, \mathbf{g} \rangle , \qquad \mathbf{z}_{ij}^{(\ell)} \leftarrow \left[\mathbf{z}_{ij}^{(\ell)}, z_{ij}^{\mathbf{g},(\ell)} \right] \tag{25}$$

where $\lambda_{\mathbf{g}}^{(\ell)}$ is a learnable scalar initialized at zero. Unlike the invariant features used by the fully equivariant model, this projection records the orientation of an edge relative to the preferred direction. It can therefore influence the transport coefficients and channel gates through the same edge networks used elsewhere in the architecture. Figure 3 illustrates the resulting symmetry regimes: full $E(n)$-equivariance when the directional pathway is inactive, and $E_{\mathbf{g}}(n)$-equivariance when it is active.

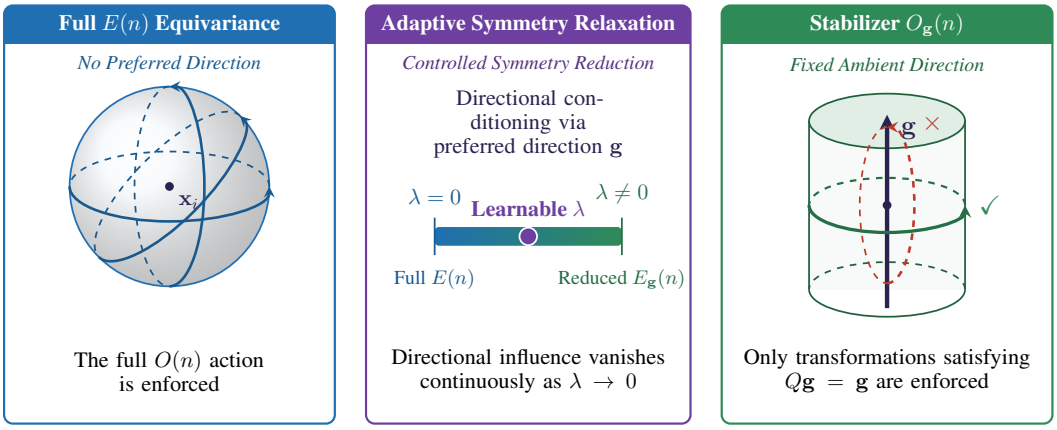

Figure 3: *Controlled symmetry relaxation through a preferred ambient direction.* When all $\lambda_{\mathbf{g}}^{(\ell)} = 0$, the directional pathway is inactive and ESNN retains full $E(n)$-equivariance. Activating the directional conditioning $\lambda_{\mathbf{g}}^{(\ell)} \langle \mathbf{r}_{ij}, \mathbf{g} \rangle$ reduces the guaranteed symmetry to $E_{\mathbf{g}}(n) = O_{\mathbf{g}}(n) \ltimes \mathbb{R}^n$, where $O_{\mathbf{g}}(n) = \{Q \in O(n) : Q\mathbf{g} = \mathbf{g}\}$. The layerwise relaxation coefficients are initialized at zero; $\mathbf{g}$ may be specified or learned, but is held fixed when the input geometry is transformed.

Table 1: *N-body dynamics.* Mean squared error (MSE) for charged and gravity-augmented prediction. Charged ESNN and gravity results are mean ± standard deviation over five runs; baseline uncertainties are reported as in the original sources. For gravity, $\max_\ell |\lambda_g^{(\ell)}| \|\mathbf{g}\|_2$ measures directional-pathway activation and A_g alignment with the gravity axis; for *Fixed*, $A_g = 1$ by construction. Lower MSE and higher A_g are better.

(a) *N-body*

Method	MSE ↓
SE(3)-Transformer	0.0248 ± 0.0010
TFN	0.0154 ± 0.0023
Radial Field	0.0106 ± 0.0001
GMN	0.0082 ± 0.0003
EGNN	0.0076 ± 0.0003
GVP-GNN	0.0069 ± 0.0002
FA-GNN	0.0057 ± 0.0002
ESNN-Id	0.0084 ± 0.0042
ESNN-Diag	0.0053 ± 0.0002
ESNN-Ortho	0.0060 ± 0.0003
ESNN-RadTan	0.0057 ± 0.0006

(b) *N-body + Gravity*

Mode	Transport	MSE ↓	$\max_\ell \lvert\lambda_g^{(\ell)}\rvert \lVert\mathbf{g}\rVert_2$	A_g ↑
None	ESNN-Diag	0.1072 ± 0.0346	–	–
	ESNN-Ortho	0.1252 ± 0.0490	–	–
	ESNN-RadTan	0.1019 ± 0.0324	–	–
Learned	ESNN-Diag	0.0205 ± 0.0016	0.2674 ± 0.0704	1.000
	ESNN-Ortho	0.0197 ± 0.0015	0.4009 ± 0.0722	1.000
	ESNN-RadTan	0.0213 ± 0.0033	2.0552 ± 0.6337	1.000
Fixed	ESNN-Diag	0.0208 ± 0.0026	0.3464 ± 0.0357	1.000
	ESNN-Ortho	0.0198 ± 0.0017	0.6100 ± 0.1771	1.000
	ESNN-RadTan	0.0234 ± 0.0054	1.3792 ± 0.3014	1.000

Theorem 6.1 (Stabilizer Subequivariance)**.** *Let $\mathbf{g} \neq \mathbf{0}$ be held fixed under transformations of the input geometry, and define its stabilizer:*

$$O_{\mathbf{g}}(n) = \{Q \in O(n) \ : \ Q\mathbf{g} = \mathbf{g}\} \tag{26}$$

An ESNN conditioned on Equation 25 is equivariant under translations and every $Q \in O_{\mathbf{g}}(n)$, with guaranteed symmetry group:

$$E_{\mathbf{g}}(n) = O_{\mathbf{g}}(n) \ltimes \mathbb{R}^n \tag{27}$$

When $\lambda_{\mathbf{g}}^{(\ell)} = 0$ for every layer, the directional conditioning vanishes and full $E(n)$-equivariance is recovered.

For any $Q \in O_{\mathbf{g}}(n)$, we have $\langle Q\mathbf{r}_{ij}, \mathbf{g}\rangle = \langle \mathbf{r}_{ij}, Q^\top \mathbf{g}\rangle = \langle \mathbf{r}_{ij}, \mathbf{g}\rangle$, so the directional term is invariant under the stabilizer of $\mathbf{g}$. Together with the covariant transport maps of Section 4, this gives the subgroup equivariance. The full proof is provided in Appendix D.6.

Symmetry-Preserving Initialization. The coefficients $\lambda_{\mathbf{g}}^{(\ell)}$ are initialized at zero, so the directional conditioning is inactive at initialization and the model begins in the fully $E(n)$-equivariant regime. During training, these coefficients can become nonzero when the data support directional dependence. The preferred direction $\mathbf{g}$ can either be specified from known physical structure or learned jointly with the rest of the model.

7 Experiments

We evaluate ESNN through four questions: **Q1** whether richer equivariant transport improves dynamics prediction when full $E(3)$ symmetry is appropriate; **Q2** whether ESNN can exploit or recover a preferred ambient direction; **Q3** how the transport mechanisms perform on irregular mesh-based physical systems; and **Q4** whether the same framework remains robust to unseen rotations in point-cloud classification. We consider charged and gravity-augmented N-body dynamics (Kipf et al., 2018), three MeshGraphNets benchmarks (Pfaff et al., 2021), and ModelNet40 (Wu et al., 2015). Appendix E.5 additionally evaluates invariant QM9 targets, where the best ESNN transport improves over EGNN on ten of twelve properties. Full experimental details are provided in Appendix E.

Q1: Particle Dynamics — Charged N-Body Dynamics. We first consider the standard five-particle charged N-body benchmark, where the model predicts future coordinates from positions, velocities, and charges. Table 1(a) shows that all three learned transports achieve lower mean error than EGNN, whereas ESNN-Id is less accurate and substantially more variable. ESNN-Diag achieves the lowest ESNN error at 0.0053 ± 0.0002, approximately 30% below EGNN (0.0076 ± 0.0003), followed by ESNN-RadTan at 0.0057 ± 0.0006 and ESNN-Ortho at 0.0060 ± 0.0003. The learned transport variants outperform both EGNN and the Identity reference in this setting, supporting the benefit of richer edge-wise geometric interactions under full Euclidean symmetry.

Table 2: *ModelNet40 rotation generalization.* Accuracy (%) under z/z, $z/\mathrm{SO}(3)$, and $\mathrm{SO}(3)/\mathrm{SO}(3)$. The $z/\mathrm{SO}(3)$ result evaluates the same z-trained checkpoint under unseen $\mathrm{SO}(3)$ rotations. Δ_{OOD} is the absolute z/z–$z/\mathrm{SO}(3)$ accuracy change in percentage points. Literature baselines follow Lippmann et al. (2025); EGNN and ESNN use the same evaluation protocol. Higher accuracy and lower Δ_{OOD} are better.

Method	z/z	$z/\mathrm{SO}(3)$	$\mathrm{SO}(3)/\mathrm{SO}(3)$	$\Delta_{\mathrm{OOD}} \downarrow$
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
EGNN	57.6	66.6	62.7	9.0
ESNN-Id	84.7	85.7	85.6	1.1
ESNN-Diag	84.6	84.3	84.4	0.3
ESNN-Ortho	84.9	85.2	84.7	0.2
ESNN-RadTan	85.4	84.6	86.3	0.7

Q2: Particle Dynamics — Gravity N-Body Dynamics. We next introduce a uniform gravitational field, $\mathbf{a}_{\mathbf{g}} = (0, 0, -9.81)^\top$, with preferred axis $\widehat{\mathbf{g}}_{\mathrm{true}} = (0, 0, -1)^\top$. We compare the fully $E(3)$-equivariant *None* model with two symmetry-relaxed variants: *Fixed*, which is given the gravity axis, and *Learned*, which estimates it from data using the mechanism of Section 6. Table 1(b) shows that both variants reduce MSE from approximately

Table 3: *Mesh-based physical dynamics.* RMSE ($\times 10^{-3}$) for one-step, 50-step, and full autoregressive rollouts on CylinderFlow, DeformingPlate, and Airfoil. Results are mean $\pm$ standard deviation over five matched runs. Best values are bold; *Diverged* denotes an unstable autoregressive rollout. Lower is better.

	CylinderFlow			DeformingPlate			Airfoil		
Method	**1-Step**	**50-Step**	**Full**	**1-Step**	**50-Step**	**Full**	**1-Step**	**50-Step**	**Full**
MeshGraphNet	4.05 ± 0.08	12.5 ± 0.6	$\mathbf{46.14 \pm 6.3}$	0.15 ± 0.02	2.2 ± 0.5	13.6 ± 1.4	1141 ± 68	1529 ± 68	$\mathbf{8033 \pm 662}$
EGNN	9.20 ± 1.80	40.5 ± 2.5	158.09 ± 2.5	0.35 ± 0.01	1.6 ± 0.1	11.5 ± 0.5	5001 ± 134	18948 ± 609	diverged
ESNN-Id	5.35 ± 0.32	20.6 ± 1.0	84.09 ± 3.4	0.15 ± 0.01	1.3 ± 0.1	7.7 ± 0.4	4744 ± 154	7579 ± 287	16893 ± 740
ESNN-Diag	2.77 ± 0.17	8.8 ± 0.4	50.84 ± 2.0	0.17 ± 0.01	1.4 ± 0.1	9.2 ± 0.5	2413 ± 49	4966 ± 126	11014 ± 321
ESNN-Ortho	2.56 ± 0.15	8.5 ± 0.4	55.42 ± 2.2	0.13 ± 0.01	1.1 ± 0.1	8.1 ± 0.4	1459 ± 27	2807 ± 66	8821 ± 259
ESNN-RadTan	3.16 ± 0.19	10.2 ± 0.5	49.79 ± 2.0	0.08 ± 0.01	1.0 ± 0.1	$\mathbf{5.8 \pm 0.3}$	1540 ± 29	2386 ± 59	8600 ± 259

0.10–0.13 to about 0.02 across all three transports. The *Learned* models achieve errors comparable to *Fixed* while their inferred directions attain sign-invariant alignment A_g that rounds to 1.000 in every case. ESNN therefore recovers the relevant symmetry-breaking axis from data without requiring it to be specified explicitly.

Q3: Mesh-Based Physical Dynamics. We evaluate ESNN on CylinderFlow, DeformingPlate, and Airfoil from MeshGraphNets (Pfaff et al., 2021), covering fluid flow, structural deformation, and compressible aerodynamics on irregular meshes. We report RMSE for one-step prediction, 50-step autoregressive rollout, and the full trajectory; evaluation details are provided in Appendix E.4. Table 3 shows task- and horizon-dependent performance. On DeformingPlate, ESNN-RadTan achieves the lowest error at every horizon, reducing full-rollout RMSE from 13.6 for MeshGraphNet to 5.8. On CylinderFlow, learned transports improve one- and 50-step prediction, while MeshGraphNet remains strongest over the full trajectory. On Airfoil, MeshGraphNet remains best at all three horizons. Across the three tasks, ESNN consistently improves over EGNN, with the largest gains on DeformingPlate.

Q4: Rotation Generalization — ModelNet40. We evaluate whether the transport framework extends beyond physical simulation to point-cloud classification. ModelNet40 shapes are represented as local k-nearest-neighbor graphs (Wu et al., 2015), and we consider the z/z, $z/\mathrm{SO}(3)$, and $\mathrm{SO}(3)/\mathrm{SO}(3)$ protocols. The $z/\mathrm{SO}(3)$ setting evaluates the same z-trained checkpoint under unseen three-dimensional rotations. Table 2 shows that ESNN remains stable under this shift, with Δ_{OOD} between 0.2 and 1.1 percentage points across transport variants, compared with 9.0 points for EGNN. Specialized rotation-robust architectures achieve higher absolute accuracy, while ESNN retains approximately 84–86% accuracy across all three protocols. These results extend ESNN to point-cloud classification while preserving robustness to unseen global rotations.

8 Conclusions and Limitations

We introduced ESNN, a first-order Cartesian framework that enriches equivariant message passing through structured matrix-valued transport of vector features across graph edges. ESNN increases geometric expressivity in the edge interactions without requiring higher-order representations. We characterize displacement-conditioned linear $O(n)$-equivariant transport and show how learned covariant features enable richer feature-conditioned transformations. We also introduce controlled symmetry relaxation for systems with a preferred ambient direction, preserving exact equivariance to the remaining symmetry while recovering full $E(n)$-equivariance as a special case. Across diverse geometric learning problems, ESNN improves predictive performance over first-order equivariant baselines, learns a preferred direction from data, and remains robust to unseen rotations. The current formulation is limited to scalar and first-order vector features; the completeness result applies only to displacement-conditioned linear transport, and symmetry relaxation assumes a single global preferred direction. Future work could extend ESNN to richer representations, local or gauge-equivariant transport, and more general symmetry-breaking fields.

AI use statement

We used generative AI tools only for language editing, including clarity, grammar, phrasing, and presentation. We did not use these tools to generate research ideas, conduct experiments, analyze

results, or formulate scientific claims. All AI-assisted text was reviewed and verified by the authors, who take full responsibility for the final manuscript.

Ethics statement

This work does not involve human subjects, personally identifiable information, or sensitive data. The datasets used in our experiments are publicly available research benchmarks and are used in accordance with their respective licenses. We are not aware of additional ethical concerns specific to the proposed method or its experimental evaluation.

Reproducibility statement

Appendix E reports dataset splits, preprocessing, model configurations, hyperparameter selection, and evaluation protocols. Machine-readable configurations accompany the experimental material. The source code and configuration files used to produce the reported results will be publicly released upon acceptance.

References

Sarp Aykent and Tian Xia. Gotennet: Rethinking efficient 3D equivariant graph neural networks. In *The Thirteenth International Conference on Learning Representations*, 2025.

Federico Barbero, Cristian Bodnar, Haitz Sáez de Ocáriz Borde, Michael Bronstein, Petar Veličković, and Pietro Liò. Sheaf neural networks with connection laplacians. In *Proceedings of Topological, Algebraic, and Geometric Learning Workshops 2022*, volume 196 of *Proceedings of Machine Learning Research*, pp. 28–36. PMLR, 2022a. URL https://proceedings.mlr.press/v196/barbero22a.html.

Federico Barbero, Cristian Bodnar, Haitz Sáez de Ocáriz Borde, and Pietro Liò. Sheaf attention networks. In *NeurIPS 2022 Workshop on Symmetry and Geometry in Neural Representations*, 2022b.

Ilyes Batatia, David P. Kovács, Gregor Simm, Christoph Ortner, and Gábor Csányi. Mace: Higher order equivariant message passing neural networks for fast and accurate force fields. In *Advances in Neural Information Processing Systems*, volume 35, pp. 11423–11436, 2022.

Ilyes Batatia, Simon Batzner, Dávid Péter Kovács, Albert Musaelian, Gregor N. C. Simm, Ralf Drautz, Christoph Ortner, Boris Kozinsky, and Gábor Csányi. The design space of E(3)-equivariant atom-centred interatomic potentials. *Nature Machine Intelligence*, 7:56–67, 2025. doi: 10.1038/s42256-024-00956-x.

Claudio Battiloro, Zhiyang Wang, Hans Riess, Paolo Di Lorenzo, and Alejandro Ribeiro. Tangent bundle convolutional learning: From manifolds to cellular sheaves and back. *IEEE Transactions on Signal Processing*, 72:1892–1909, 2024. doi: 10.1109/TSP.2024.3379862.

Claudio Battiloro, Ege Karaismailoğlu, Mauricio Tec, George Dasoulas, Michelle Audirac, and Francesca Dominici. E(n) equivariant topological neural networks. In *International Conference on Learning Representations*, 2025.

Cristian Bodnar, Francesco Di Giovanni, Benjamin Paul Chamberlain, Pietro Lió, and Michael M. Bronstein. Neural sheaf diffusion: A topological perspective on heterophily and oversmoothing in GNNs. In *Advances in Neural Information Processing Systems*, volume 35, pp. 18561–18577, 2022. URL https://arxiv.org/abs/2202.04579.

Alessio Borgi, Fabrizio Silvestri, and Pietro Liò. Polynomial neural sheaf diffusion: A spectral filtering approach on cellular sheaves. *arXiv preprint arXiv:2512.00242*, 2025.

Luke Braithwaite, Alessio Borgi, Gabriele Onorato, Kristjan Tarantelli, Francesco Restuccia, Fabrizio Silvestri, and Pietro Liò. Heterogeneous sheaf neural networks. *arXiv preprint arXiv:2409.08036*, 2024.

Johannes Brandstetter, Rob Hesselink, Elise van der Pol, Erik J. Bekkers, and Max Welling. Geometric and physical quantities improve E(3) equivariant message passing. In *International Conference on Learning Representations*, 2022. URL https://arxiv.org/abs/2110.02905.

Jiacheng Cen, Anyi Li, Ning Lin, Yuxiang Ren, Zihe Wang, and Wenbing Huang. Are high-degree representations really unnecessary in equivariant graph neural networks? In *Advances in Neural Information Processing Systems*, volume 37, pp. 26238–26266, 2024. doi: 10.52202/079017-0826.

Taco S. Cohen, Maurice Weiler, Berkay Kicanaoglu, and Max Welling. Gauge equivariant convolutional networks and the icosahedral CNN. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97, pp. 1321–1330. PMLR, 2019. URL https://arxiv.org/abs/1902.04615.

Weitao Du, Yuanqi Du, Limei Wang, Dieqiao Feng, Guifeng Wang, Shuiwang Ji, Carla P. Gomes, and Zhi-Ming Ma. A new perspective on building efficient and expressive 3D equivariant graph neural networks. In *Advances in Neural Information Processing Systems*, volume 36, pp. 66647–66674, 2023.

Iulia Duta, Giulia Cassarà, Fabrizio Silvestri, and Pietro Liò. Sheaf hypergraph networks. *Advances in Neural Information Processing Systems*, 36:12087–12099, 2023.

Alexandre Duval, Victor Schmidt, Alex Hernández-García, Santiago Miret, Fragkiskos D. Malliaros, Yoshua Bengio, and David Rolnick. FAENet: Frame averaging equivariant GNN for materials modeling. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202, pp. 9013–9033. PMLR, 2023. URL https://proceedings.mlr.press/v202/duval23a.html.

Stefano Fiorini, Hakan Emre Aktas, Iulia Duta, Pietro Morerio, Alessio Del Bue, Pietro Liò, and Stefano Coniglio. Sheaves reloaded: A direction awakening. In *International Conference on Learning Representations*, 2026.

Fabian B. Fuchs, Daniel E. Worrall, Volker Fischer, and Max Welling. SE(3)-transformers: 3D roto-translation equivariant attention networks. In *Advances in Neural Information Processing Systems*, volume 33, pp. 1970–1981, 2020. URL https://arxiv.org/abs/2006.10503.

Mustafa Hajij, Lennart Bastian, Sarah Osentoski, Hardik Kabaria, John L. Davenport, Sheik Dawood, Balaji Cherukuri, Joseph G. Kocheemoolayil, Nastaran Shahmansouri, Adrian Lew, Theodore Papamarkou, and Tolga Birdal. Copresheaf topological neural networks: A generalized deep learning framework. In *Advances in Neural Information Processing Systems*, volume 38, 2025.

Jiaqi Han, Wenbing Huang, Hengbo Ma, Jiachen Li, Josh Tenenbaum, and Chuang Gan. Learning physical dynamics with subequivariant graph neural networks. *Advances in Neural Information Processing Systems*, 35:26256–26268, 2022.

Jakob Hansen and Thomas Gebhart. Sheaf neural networks. In *NeurIPS 2020 Workshop on TDA and Beyond*, 2020. URL https://arxiv.org/abs/2012.06333.

Ferran Hernandez Caralt, Mar Gonzàlez i Català, Adrián Bazaga, and Pietro Liò. On the necessity of learnable sheaf laplacians. In *ICLR 2026 Workshop on Geometry-grounded Representation Learning and Generative Modeling*, 2026. URL https://arxiv.org/abs/2603.05395. Tiny Paper Track.

Elyssa Hofgard, Rui Wang, Robin Walters, and Tess Smidt. Relaxed equivariant graph neural networks. *arXiv preprint arXiv:2407.20471*, 2024. URL https://arxiv.org/abs/2407.20471.

Bowen Jing, Stephan Eismann, Pratham N. Soni, and Ron O. Dror. Equivariant graph neural networks for 3D macromolecular structure. *arXiv preprint arXiv:2106.03843*, 2021. URL https://arxiv.org/abs/2106.03843.

Sékou-Oumar Kaba, Arnab Kumar Mondal, Yan Zhang, Yoshua Bengio, and Siamak Ravanbakhsh. Equivariance with learned canonicalization functions. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202, pp. 15546–15566. PMLR, 2023. URL https://proceedings.mlr.press/v202/kaba23a.html.

Thomas Kipf, Ethan Fetaya, Kuan-Chieh Wang, Max Welling, and Richard Zemel. Neural relational inference for interacting systems. In *International conference on machine learning*, pp. 2688–2697. Pmlr, 2018.

Miltiadis Kofinas, Erik J. Bekkers, Naveen Shankar Nagaraja, and Efstratios Gavves. Latent field discovery in interacting dynamical systems with neural fields. *Advances in Neural Information Processing Systems (NeurIPS)*, 2023. URL https://arxiv.org/abs/2310.20679.

Veljko Kovač, Erik J. Bekkers, Pietro Liò, and Floor Eijkelboom. E(n) equivariant message passing cellular networks. In *Proceedings of the Geometry-grounded Representation Learning and Generative Modeling Workshop*, volume 251 of *Proceedings of Machine Learning Research*, pp. 173–186. PMLR, 2024. URL https://proceedings.mlr.press/v251/kovac-24a.html.

Daniel Levy, Sékou-Oumar Kaba, Carmelo Gonzales, Santiago Miret, and Siamak Ravanbakhsh. Using multiple vector channels improves E(n)-equivariant graph neural networks. In *ICML Workshop on Machine Learning for Astrophysics*, 2023.

Daiyuan Li, Shreya Arya, and Robert Ghrist. Learning from frustration: Torsor CNNs on graphs. In *Proceedings of the 4th (2025) and 3rd (2024) NeurIPS Workshops on Symmetry and Geometry in Neural Representations*, volume 282 of *Proceedings of Machine Learning Research*, pp. 298–315. PMLR, 2026. URL https://proceedings.mlr.press/v282/li26a.html.

Yi-Lun Liao, Brandon M. Wood, Abhishek Das, and Tess Smidt. Equiformerv2: Improved equivariant transformer for scaling to higher-degree representations. In *International Conference on Learning Representations*, volume 2024, pp. 39282–39309, 2024. URL https://openreview.net/forum?id=mCOBKZmrzD.

Peter Lippmann, Gerrit Gerhartz, Roman Remme, and Fred A. Hamprecht. Beyond canonicalization: How tensorial messages improve equivariant message passing. In *International Conference on Learning Representations*, 2025.

Cong Liu, David Ruhe, and Patrick Forré. Multivector neurons: Better and faster $O(n)$-equivariant clifford graph neural networks. *arXiv preprint arXiv:2406.04052*, 2024a. URL https://arxiv.org/abs/2406.04052.

Yang Liu, Jiashun Cheng, Haihong Zhao, Tingyang Xu, Peilin Zhao, Fugee Tsung, Jia Li, and Yu Rong. SEGNO: Generalizing equivariant graph neural networks with physical inductive biases. In *The Twelfth International Conference on Learning Representations*, 2024b. URL https://openreview.net/forum?id=3oTPsORaDH.

Yoshihiro Maruyama. Foundations of equivariant deep learning: Unifying graph and sheaf neural networks. *arXiv preprint arXiv:2607.03798*, 2026. URL https://arxiv.org/abs/2607.03798.

Yuhan Peng, Junwen Dong, Yuzhi Zeng, Hao Li, Ce Ju, Huitao Feng, Diaaeldin Taha, Anna Wienhard, and Kelin Xia. Sheaf neural networks on SPD manifolds: Second-order geometric representation learning. *arXiv preprint arXiv:2604.20308*, 2026. URL https://arxiv.org/abs/2604.20308.

Tobias Pfaff, Meire Fortunato, Alvaro Sanchez-Gonzalez, and Peter W. Battaglia. Learning mesh-based simulation with graph networks. *International Conference on Learning Representations (ICLR)*, 2021. URL https://arxiv.org/abs/2010.03409.

Omri Puny, Matan Atzmon, Heli Ben-Hamu, Ishan Misra, Aditya Grover, Edward J. Smith, and Yaron Lipman. Frame averaging for invariant and equivariant network design. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=zIU4k8hK3H.

André Ribeiro, Ana Luiza Tenório, Juan Belieni, Amauri H. Souza, and Diego Mesquita. Cooperative sheaf neural networks. In *International Conference on Learning Representations*, 2026.

Victor Garcia Satorras, Emiel Hoogeboom, and Max Welling. E(n) equivariant graph neural networks. In *International Conference on Machine Learning*, pp. 9323–9332. PMLR, 2021. URL https://arxiv.org/abs/2102.09844.

Kristof Schütt, Oliver Unke, and Michael Gastegger. Equivariant message passing for the prediction of tensorial properties and molecular spectra. In *International Conference on Machine Learning*, pp. 9377–9388. PMLR, 2021.

Guillem Simeon and Gianni De Fabritiis. TensorNet: Cartesian tensor representations for efficient learning of molecular potentials. In *Advances in Neural Information Processing Systems*, volume 36, pp. 37334–37353, 2023.

Nathaniel Thomas, Tess Smidt, Steven Kearnes, Lusann Yang, Li Li, Kai Kohlhoff, and Patrick Riley. Tensor field networks: Rotation- and translation-equivariant neural networks for 3D point clouds. *arXiv preprint arXiv:1802.08219*, 2018. URL https://arxiv.org/abs/1802.08219.

Tycho F. A. van der Ouderaa, David W. Romero, and Mark van der Wilk. Relaxing equivariance constraints with non-stationary continuous filters. In *Advances in Neural Information Processing Systems*, volume 35, pp. 25776–25790, 2022.

Rui Wang, Robin Walters, and Rose Yu. Approximately equivariant networks for imperfectly symmetric dynamics. In *Proceedings of the 39th International Conference on Machine Learning*, pp. 23078–23091. PMLR, 2022. URL https://arxiv.org/abs/2201.11969.

Rui Wang, Elyssa Hofgard, Han Gao, Robin Walters, and Tess Smidt. Discovering symmetry breaking in physical systems with relaxed group convolution. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 50599–50621. PMLR, 2024a. URL https://proceedings.mlr.press/v235/wang24y.html.

Yusong Wang, Tong Wang, Shaoning Li, Xinheng He, Mingyu Li, Zun Wang, Nanning Zheng, Bin Shao, and Tie-Yan Liu. Enhancing geometric representations for molecules with equivariant vector-scalar interactive message passing. *Nature Communications*, 15(1):313, 2024b. doi: 10.1038/s41467-023-43720-2.

Maurice Weiler, Patrick Forré, Erik Verlinde, and Max Welling. Coordinate independent convolutional networks–isometry and gauge equivariant convolutions on riemannian manifolds. *arXiv preprint arXiv:2106.06020*, 2021.

Zhirong Wu, Shuran Song, Aditya Khosla, Fisher Yu, Linguang Zhang, Xiaoou Tang, and Jianxiong Xiao. 3d shapenets: A deep representation for volumetric shapes. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1912–1920, 2015.

Equivariant Sheaf Neural Networks: Learning Geometric Transport on Graphs

Supplementary Material

CONTENTS

A Extended Related Work **15**

A.1 Cartesian and Steerable Equivariant Models 15

A.2 Frames, Gauges, and Topological Domains 15

A.3 Sheaf Learning and Directionality 16

A.4 Symmetry Relaxation and Subequivariant Inductive Biases 16

B Algebraic Foundations of Equivariant Spatial Transport **17**

B.1 Tensor-Product Structure of Vector Transport 17

B.2 Completeness Under Displacement-Only Conditioning 17

B.3 Feature-Conditioned Spatial Operators 19

C Cellular Sheaves, Directed Transport, and Connections **20**

C.1 Cellular Sheaves and Induced Node-to-Node Coupling 21

C.2 Directed Transport as a Quiver Representation 21

C.3 Adjoint-Consistent and Connection-Sheaf Specializations 22

C.4 Ambient Equivariance and Local Gauge Transformations 24

D Proofs of Equivariance and Symmetry Results **24**

D.1 Proof of Proposition 3.2: $O(n)$-Equivariance of the Transport Map 25

D.2 $O(n)$-Equivariance of Orthogonal Transport 25

D.3 $O(n)$-Equivariance of Radial–Tangential Transport 26

D.4 Completeness of Displacement-Conditioned Linear Transport 26

D.5 Proof of Theorem 5.1: $E(n)$-Equivariance of the ESNN Layer 27

D.6 Proof of Theorem 6.1: Stabilizer Subequivariance 28

E Experimental Details **28**

E.1 Charged N-Body Dynamics 29

E.2 Gravity-Augmented N-Body Dynamics 29

E.3 ModelNet40 Point-Cloud Classification 30

E.4 Mesh-Based Physical Dynamics 31

E.5 QM9 Molecular-Property Prediction 32

E.6 Implementation, Model Selection, and Evaluation 33

A Extended Related Work

This appendix extends the discussion in Section 2 by placing ESNN within three closely related areas: equivariant geometric learning, sheaf-based representation learning, and methods that relax or reduce symmetry in the presence of external structure. We focus on approaches that act on geometric vector features, learn edge-dependent transformations between local representations, or adapt the symmetry imposed on message passing.

A.1 Cartesian and Steerable Equivariant Models

A central distinction among $E(n)$-equivariant graph networks lies in the geometric representations they propagate and how these representations are coupled across neighboring nodes. One broad family works directly with invariant scalars and low-order Cartesian features. EGNN (Satorras et al., 2021) constructs coordinate updates from invariant edge functions acting on relative displacement vectors, while PaiNN (Schütt et al., 2021) and GVP-based models (Jing et al., 2021) maintain scalar and vector channels coupled through equivariant operations. Subsequent architectures have expanded this Cartesian design space through richer scalar–vector interactions (Wang et al., 2024b), multiple vector channels (Levy et al., 2023), and Cartesian tensor features (Simeon & De Fabritiis, 2023). More recent models such as GotenNet (Aykent & Xia, 2025) similarly seek greater geometric expressivity while avoiding explicit Clebsch–Gordan contractions. A complementary family uses steerable features transforming under higher-order group representations. Tensor Field Networks (Thomas et al., 2018), SE(3)-Transformers (Fuchs et al., 2020), SEGNN (Brandstetter et al., 2022), MACE (Batatia et al., 2022), and EquiformerV2 (Liao et al., 2024) represent angular structure through spherical harmonics and tensor-product or related steerable operations. Higher representation degrees can capture increasingly rich angular information, but also increase the cost of the representations and their couplings. Recent work has therefore investigated how much high-degree information is needed and how it can be processed more efficiently (Cen et al., 2024). ESNN instead retains a first-order scalar–vector representation and increases geometric expressivity through matrix-valued transport of vector features across graph edges.

A.2 Frames, Gauges, and Topological Domains

Another route to equivariance is to resolve orientation dependence through the choice or averaging of reference frames. Frame Averaging (Puny et al., 2022) constructs exactly equivariant models by averaging a backbone over an equivariant frame, with FAENet (Duval et al., 2023) adapting this idea to atomistic systems. Learned canonicalization (Kaba et al., 2023) instead predicts a canonical representative before applying a generic backbone, while subsequent work has examined the benefit of retaining tensorial messages beyond canonicalized scalar representations (Lippmann et al., 2025). Local gauge equivariance addresses a related but distinct setting, in which different sites may use independently chosen reference frames. Gauge-equivariant CNNs (Cohen et al., 2019) and coordinate-independent convolutions (Weiler et al., 2021) require the network to transform consistently under changes of these local frames. Tangent-bundle convolutional networks connect this viewpoint to connection Laplacians and cellular sheaves through vector-field diffusion on Riemannian manifolds (Battiloro et al., 2024), while Torsor CNNs extend local group-valued transport to arbitrary graphs through edge potentials relating neighboring frames (Li et al., 2026). These models concern local gauge equivariance, whereas ESNN enforces equivariance under a common ambient $O(n)$ action. Its Orthogonal Transport therefore has a connection-like geometric interpretation, but is not a general gauge-equivariant construction. A separate line of work extends Euclidean equivariance to richer combinatorial domains. $E(n)$-Equivariant Topological Neural Networks (Battiloro et al., 2025) generalize equivariant message passing from graphs to combinatorial complexes, while $E(n)$-equivariant message-passing cellular networks (Kovač et al., 2024) develop related constructions on cellular structures. These approaches enrich the underlying domain through higher-order cells. ESNN is complementary: it remains graph-based and instead enriches the transport between geometric feature spaces. A related theoretical perspective is provided by (Maruyama, 2026), who formulates order-equivariant neural networks on equivariant vector bundles over face posets and recovers graph and sheaf layers within a common framework. Their construction organizes equivariance over the combinatorial structure of the domain, whereas ESNN focuses on the continuous Euclidean action on

geometric feature fibers and constrains edge transport accordingly. The two perspectives therefore address distinct but compatible forms of symmetry.

A.3 Sheaf Learning and Directionality

Cellular sheaves associate local vector spaces with graph cells and relate them through incidence restriction maps. Early SNNs (Hansen & Gebhart, 2020) used the resulting sheaf Laplacian to generalize graph convolution, while Neural Sheaf Diffusion (Bodnar et al., 2022) introduced learnable restriction maps and studied their effect on heterophily, class separation, and oversmoothing. These works established neural sheaf diffusion as a framework for propagation through learned compatibility maps between local feature spaces. Subsequent architectures have explored richer sheaf structures and propagation operators. Connection-Laplacian SNNs (Barbero et al., 2022a) use orthogonal restriction maps motivated by local tangent-space alignment, providing a particularly relevant precedent for geometric transport. Sheaf Attention Networks (Barbero et al., 2022b) incorporate attention into sheaf propagation, Sheaf Hypergraph Networks (Duta et al., 2023) extend the construction to higher-order relations, and Heterogeneous Sheaf Neural Networks (Braithwaite et al., 2024) introduce heterogeneous stalk and restriction structures for typed graphs. Polynomial Neural Sheaf Diffusion (Borgi et al., 2025) develops higher-order spectral filters of the sheaf operator, while recent work has also considered non-Euclidean stalk geometries such as SPD-valued representations (Peng et al., 2026). Directionality has become an important extension of this framework. Classical sheaf Laplacians are self-adjoint and therefore do not directly represent independent propagation rules for the two orientations of an edge. Cooperative Sheaf Neural Networks (Ribeiro et al., 2026) introduce directed sheaf constructions with in- and out-degree Laplacians, while Directed Sheaf Neural Networks (Fiorini et al., 2026) develop a directed cellular-sheaf formulation and a corresponding directional Laplacian. These models show how asymmetric information flow can be incorporated without forcing opposite edge orientations to encode the same interaction. A related categorical perspective is provided by Copresheaf Topological Neural Networks (Hajij et al., 2025), which organize directed maps between local feature spaces and their compositions. ESNN builds on this matrix-valued view of information transport while imposing the additional Euclidean transformation laws required by geometric vector features. As detailed in Appendix C, classical restriction maps induce node-to-node couplings such as $\rho_{i\to e}^{*}\rho_{j\to e}$, whereas ESNN parameterizes the directed node-to-node transport directly. Adjoint-consistent transports together with symmetric aggregation weights yield a self-adjoint specialization, while additionally imposing full-fiber orthogonality gives the exact connection-sheaf case. The general ESNN operator instead remains a directed matrix-valued transport between neighboring feature spaces. Recent empirical work also suggests that unrestricted learned sheaf maps are not universally advantageous: identity-sheaf baselines can be competitive with learned operators on standard heterophilic benchmarks (Hernandez Caralt et al., 2026). ESNN therefore targets a more structured setting: geometric vector features with a known Euclidean transformation law, for which edge-wise transport can be designed to remain compatible with that symmetry.

A.4 Symmetry Relaxation and Subequivariant Inductive Biases

External fields and environmental structure can reduce the symmetry of a physical system. One approach is to enforce the subgroup that preserves the relevant external direction. Subequivariant Graph Neural Networks (Han et al., 2022), for example, incorporate fields such as gravity and construct message-passing operations equivariant to the corresponding stabilizer subgroup. ESNN follows the same group-theoretic principle, but introduces the preferred direction within the transport framework and scales its contribution through a learnable coefficient. A different line of work relaxes equivariance rather than replacing it with an exact subgroup symmetry. Non-stationary continuous filters (van der Ouderaa et al., 2022), approximately equivariant networks (Wang et al., 2022), relaxed group convolutions (Wang et al., 2024a), and Relaxed Equivariant Graph Neural Networks (Hofgard et al., 2024) learn controlled departures from a prescribed symmetry. External effects can also be modeled separately, as in Latent Field Discovery (Kofinas et al., 2023), which augments equivariant local interactions with an additional global neural field. ESNN is closest to the subequivariant setting: when directional conditioning is active, it preserves exact equivariance to the stabilizer of the preferred direction, while full $E(n)$-equivariance is recovered when the directional coefficients vanish across all layers.

B Algebraic Foundations of Equivariant Spatial Transport

This appendix develops the algebraic structure underlying the transport maps introduced in Section 3. We first derive their spatial–channel factorization, then characterize the complete class of $O(n)$-equivariant transports available when relative displacement is the only covariant geometric input, and finally show how additional learned covariant features give rise to richer spatial operators. Throughout, the transport is linear with respect to the vector message for a fixed set of context-dependent coefficients. Since these coefficients may themselves be predicted from the current feature state, the resulting message function need not be linear in the evolving features.

B.1 Tensor-Product Structure of Vector Transport

The vector component of an ESNN stalk is:

$$\mathcal{F}_{\mathrm{vec}}(i) = \mathbb{R}^n \otimes \mathbb{R}^{c_v} \tag{28}$$

where $\mathbb{R}^n$ is the spatial factor carrying the $O(n)$ action and $\mathbb{R}^{c_v}$ indexes the vector channels. After identifying the tensor product with $\mathbb{R}^{n\times c_v}$, the group therefore acts as:

$$\mathbf{V} \mapsto Q\mathbf{V}, \qquad Q \in O(n) \tag{29}$$

This separation between spatial and channel factors also carries over to linear operators. In finite dimensions:

$$\mathrm{End}\left(\mathbb{R}^n \otimes \mathbb{R}^{c_v}\right) \cong \mathrm{End}(\mathbb{R}^n) \otimes \mathrm{End}(\mathbb{R}^{c_v}) \tag{30}$$

so every linear operator on the vector stalk can be expressed as a finite sum of separable spatial and channel transformations.

Proposition B.1 (Separable Expansion of Linear Vector Transport)**.** *Let $\mathcal{L} : \mathbb{R}^{n\times c_v} \longrightarrow \mathbb{R}^{n\times c_v}$ be linear. Then there exist spatial matrices $\mathbf{S}^{(k)} \in \mathbb{R}^{n\times n}$ and channel matrices $\mathbf{M}^{(k)} \in \mathbb{R}^{c_v\times c_v}$ such that:*

$$\mathcal{L}(\mathbf{V}) = \sum_{k=1}^{K} \mathbf{S}^{(k)}\mathbf{V}\mathbf{M}^{(k)} \tag{31}$$

for some finite K.

Proof. Equation 30 implies that any endomorphism of $\mathbb{R}^n \otimes \mathbb{R}^{c_v}$ can be written as a finite sum of elementary tensor-product operators. Under the matrix identification $\mathbb{R}^n \otimes \mathbb{R}^{c_v} \simeq \mathbb{R}^{n\times c_v}$, the standard vectorization identity:

$$\mathrm{vec}\left(\mathbf{S}\mathbf{V}\mathbf{M}\right) = \left(\mathbf{M}^\top \otimes \mathbf{S}\right)\mathrm{vec}(\mathbf{V}) \tag{32}$$

maps each elementary tensor-product operator to a left spatial action and a right channel action, yielding Equation 31. □

Proposition B.1 is purely algebraic and does not by itself enforce equivariance. In ESNN, the spatial and channel matrices may depend on the local edge context, so this dependence must itself respect the group action. For a transformed context $Q\cdot\mathcal{C}_{ij}$, the sufficient component-wise conditions introduced in Section 3 are:

$$\mathbf{S}_{ij}^{(k)}(Q\cdot\mathcal{C}_{ij}) = Q\mathbf{S}_{ij}^{(k)}(\mathcal{C}_{ij})Q^\top \qquad \mathbf{M}_{ij}^{(k)}(Q\cdot\mathcal{C}_{ij}) = \mathbf{M}_{ij}^{(k)}(\mathcal{C}_{ij}) \tag{33}$$

These are precisely the conditions used in Proposition 3.2. The decomposition separates the spatial and channel actions, while the corresponding covariance and invariance constraints ensure compatibility with the ambient $O(n)$ action. Figure 4 illustrates this spatial–channel factorization.

B.2 Completeness Under Displacement-Only Conditioning

This subsection establishes the completeness result stated in Theorem 4.1. The key assumption is that relative displacement is the only covariant geometric quantity available to the transport, while any additional conditioning is $O(n)$-invariant. Under this restriction, equivariance with respect to the stabilizer of a non-zero displacement forces the spatial action to separate into radial and tangential

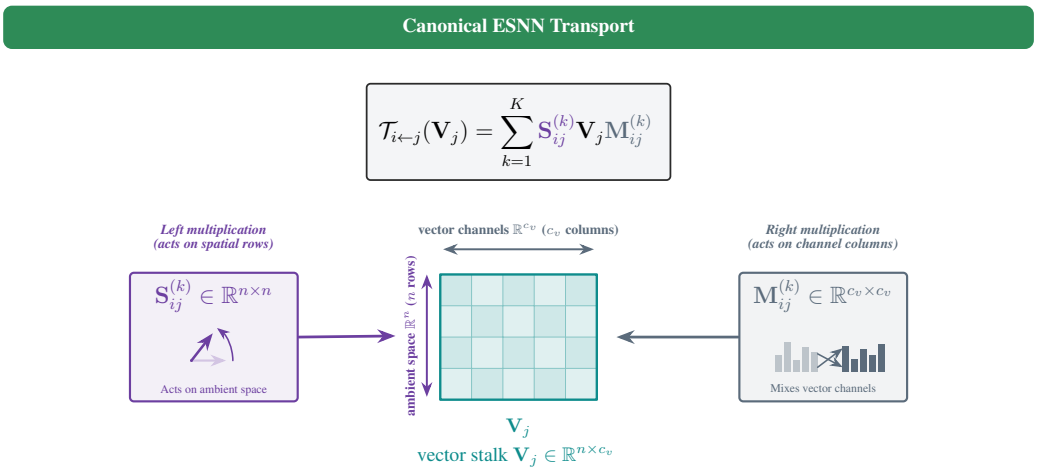

Figure 4: *Spatial–channel factorization of ESNN transport.* Each term acts on $\mathbf{V}_j \in \mathbb{R}^{n\times c_v}$ along two independent axes: $\mathbf{S}_{ij}^{(k)}$ acts on the ambient spatial dimension by left multiplication, while $\mathbf{M}_{ij}^{(k)}$ acts on the vector channels by right multiplication. Summing the separable terms gives the canonical transport $\mathcal{T}_{i\leftarrow j}$. ESNN enforces equivariance by constraining the spatial operators to be $O(n)$-covariant and the channel maps to be invariant.

components. The proof below makes this constraint explicit and characterizes the resulting transport class.

Theorem 4.1. *Let $n \geq 2$, and consider a linear transport on $\mathbb{R}^n \otimes \mathbb{R}^{c_v}$ whose only covariant geometric conditioning variable is a non-zero displacement $\mathbf{r} \in \mathbb{R}^n$, with arbitrary additional $O(n)$-invariant scalar conditioning. Every such $O(n)$-equivariant transport can be written as:*

$$\mathcal{T}_{\mathbf{r},\boldsymbol{\xi}}(\mathbf{V}) = \mathbf{P}^{\parallel}(\mathbf{r})\mathbf{V}\mathbf{A}_{\mathbf{r},\boldsymbol{\xi}} + \mathbf{P}^{\perp}(\mathbf{r})\mathbf{V}\mathbf{B}_{\mathbf{r},\boldsymbol{\xi}} \tag{34}$$

where $\mathbf{A}_{\mathbf{r},\boldsymbol{\xi}}$ and $\mathbf{B}_{\mathbf{r},\boldsymbol{\xi}}$ are channel endomorphisms whose dependence on the geometric conditioning is $O(n)$-invariant.

Proof. Let

$$\widehat{\mathbf{r}} = \frac{\mathbf{r}}{\|\mathbf{r}\|} \qquad \mathbf{P}^{\parallel} = \widehat{\mathbf{r}}\widehat{\mathbf{r}}^{\top} \qquad \mathbf{P}^{\perp} = \mathbf{I}_n - \mathbf{P}^{\parallel} \tag{35}$$

For a fixed non-zero displacement $\mathbf{r}$, consider its stabilizer:

$$H_{\mathbf{r}} = \{Q \in O(n) : Q\mathbf{r} = \mathbf{r}\} \cong O(n-1) \tag{36}$$

Under this subgroup, the spatial representation decomposes as:

$$\mathbb{R}^n = \mathrm{span}\{\widehat{\mathbf{r}}\} \oplus \widehat{\mathbf{r}}^{\perp} \tag{37}$$

The stabilizer acts trivially on the radial component $\mathrm{span}\{\widehat{\mathbf{r}}\}$ and through the standard $O(n-1)$ representation on the tangential component $\widehat{\mathbf{r}}^{\perp}$.

Let $\boldsymbol{\xi}$ denote any additional $O(n)$-invariant scalar conditioning available to the transport, and consider:

$$\mathcal{T}_{\mathbf{r},\boldsymbol{\xi}} : \mathbb{R}^n \otimes \mathbb{R}^{c_v} \longrightarrow \mathbb{R}^n \otimes \mathbb{R}^{c_v}$$

Since $\boldsymbol{\xi}$ is invariant, $O(n)$-equivariance requires:

$$\mathcal{T}_{Q\mathbf{r},\boldsymbol{\xi}}(Q\mathbf{V}) = Q\,\mathcal{T}_{\mathbf{r},\boldsymbol{\xi}}(\mathbf{V}), \qquad \forall Q \in O(n) \tag{38}$$

For every $Q \in H_{\mathbf{r}}$, the displacement is unchanged, so Equation 38 reduces to:

$$\mathcal{T}_{\mathbf{r},\boldsymbol{\xi}}(Q\mathbf{V}) = Q\,\mathcal{T}_{\mathbf{r},\boldsymbol{\xi}}(\mathbf{V}) \tag{39}$$

Thus, for fixed $(\mathbf{r}, \boldsymbol{\xi})$, the transport must commute with the action of the stabilizer $H_{\mathbf{r}}$.

The two spatial subspaces in Equation 37 carry inequivalent representations of $H_{\mathbf{r}}$. The radial line carries the trivial representation, whereas $\widehat{\mathbf{r}}^{\perp}$ carries the standard representation of $O(n-1)$. An intertwining operator therefore cannot mix the radial and tangential components. On the tangential subspace, the spatial part of the commutant is proportional to the identity. Since $\mathbb{R}^{c_v}$ carries no non-trivial $O(n)$ action, arbitrary linear maps remain available on the channel dimension. Consequently:

$$\mathrm{End}_{H_{\mathbf{r}}}\left(\mathbb{R}^n \otimes \mathbb{R}^{c_v}\right) = \Big(\mathbf{P}^{\parallel} \otimes \mathrm{End}(\mathbb{R}^{c_v})\Big) \oplus \Big(\mathbf{P}^{\perp} \otimes \mathrm{End}(\mathbb{R}^{c_v})\Big) \tag{40}$$

Hence the transport must take the form:

$$\boxed{\mathcal{T}_{\mathbf{r},\boldsymbol{\xi}}(\mathbf{V}) = \mathbf{P}^{\parallel}(\mathbf{r})\mathbf{V}\mathbf{A}_{\mathbf{r},\boldsymbol{\xi}} + \mathbf{P}^{\perp}(\mathbf{r})\mathbf{V}\mathbf{B}_{\mathbf{r},\boldsymbol{\xi}}} \tag{41}$$

for channel endomorphisms $\mathbf{A}_{\mathbf{r},\boldsymbol{\xi}}$ and $\mathbf{B}_{\mathbf{r},\boldsymbol{\xi}}$ in $\mathbb{R}^{c_v \times c_v}$.

It remains to determine how these channel maps may depend on the orientation of $\mathbf{r}$. From Equation 38 and

$$\mathbf{P}^{\parallel}(Q\mathbf{r}) = Q\mathbf{P}^{\parallel}(\mathbf{r})Q^{\top}, \qquad \mathbf{P}^{\perp}(Q\mathbf{r}) = Q\mathbf{P}^{\perp}(\mathbf{r})Q^{\top}$$

the uniqueness of the radial–tangential block decomposition gives:

$$\mathbf{A}_{Q\mathbf{r},\boldsymbol{\xi}} = \mathbf{A}_{\mathbf{r},\boldsymbol{\xi}}, \qquad \mathbf{B}_{Q\mathbf{r},\boldsymbol{\xi}} = \mathbf{B}_{\mathbf{r},\boldsymbol{\xi}} \tag{42}$$

The channel endomorphisms can therefore depend on the displacement only through $O(n)$-invariant quantities such as $\|\mathbf{r}\|$, together with the additional invariant conditioning $\boldsymbol{\xi}$. This establishes Equation 41 as the complete class of linear $O(n)$-equivariant transports under the stated conditioning assumptions. □

Additional $O(n)$-invariant scalar information does not alter the radial–tangential structure. It may change the channel maps $\mathbf{A}_{\mathbf{r},\boldsymbol{\xi}}$ and $\mathbf{B}_{\mathbf{r},\boldsymbol{\xi}}$, but cannot introduce additional spatial operators. ESNN instantiates these channel maps through the structured factorization $\mathbf{WD}(\boldsymbol{\eta}_{ij})$.

The completeness statement relies on the full orthogonal group $O(n)$. Under $SO(n)$ alone, additional orientation-sensitive spatial operators may be admissible. In three dimensions, for example, consider $[\widehat{\mathbf{r}}]_{\times}$ defined by:

$$[\widehat{\mathbf{r}}]_{\times}\mathbf{v} = \widehat{\mathbf{r}} \times \mathbf{v} \tag{43}$$

For $Q \in O(3)$, this operator transforms as:

$$[Q\widehat{\mathbf{r}}]_{\times} = \det(Q)\, Q[\widehat{\mathbf{r}}]_{\times}Q^{\top} \tag{44}$$

It therefore satisfies the required conjugation law for $Q \in SO(3)$ but acquires an additional sign under reflections. Theorem 4.1 should consequently be understood as a completeness result for displacement-conditioned linear $O(n)$-equivariant transport, not for the broader class of $SO(n)$-equivariant kernels.

B.3 Feature-Conditioned Spatial Operators

The completeness result of Appendix B.2 applies when relative displacement is the only covariant geometric input available to the transport. ESNN can also construct spatial operators from the learned vector features at the endpoints of an edge. These additional covariant features provide a richer local geometric description and support transport mechanisms beyond the radial–tangential class. The resulting spatial operators may vary with the evolving feature state while retaining the spatial–channel structure introduced in Section 3.

A simple example is the cross-feature operator:

$$\mathbf{C}_{ij} = \mathbf{V}_j\mathbf{V}_i^{\top} \tag{45}$$

which transforms by conjugation under a common orthogonal transformation:

$$\mathbf{C}_{ij} \mapsto Q\mathbf{C}_{ij}Q^{\top} \tag{46}$$

Its skew-symmetric component:

$$\boldsymbol{\Omega}^{V}_{ij} = \mathbf{V}_j\mathbf{V}_i^{\top} - \mathbf{V}_i\mathbf{V}_j^{\top} \tag{47}$$

inherits the same transformation law:

$$\mathbf{\Omega}^V_{ij} \mapsto Q\mathbf{\Omega}^V_{ij}Q^\top \tag{48}$$

Since the Frobenius norm is invariant under orthogonal conjugation, the normalized operator:

$$\widehat{\mathbf{\Omega}}^V_{ij} = \frac{\mathbf{\Omega}^V_{ij}}{\|\mathbf{\Omega}^V_{ij}\|_F + \varepsilon} \tag{49}$$

is also a valid $O(n)$-covariant spatial operator. This gives a feature-conditioned skew component that can be combined with the displacement-based operators introduced in Section 4.

Orthogonal Transport follows the same principle, with normalization applied before extracting the skew component. Starting from $\mathbf{C}_{ij} = \mathbf{V}_j\mathbf{V}_i^\top$, define:

$$\widetilde{\mathbf{C}}_{ij} = \frac{\mathbf{C}_{ij}}{\|\mathbf{C}_{ij}\|_F + \varepsilon}, \qquad \mathbf{\Omega}_{ij} = \widetilde{\mathbf{C}}_{ij} - \widetilde{\mathbf{C}}_{ij}^\top \tag{50}$$

The invariance of the Frobenius norm gives:

$$\mathbf{\Omega}_{ij} \mapsto Q\mathbf{\Omega}_{ij}Q^\top \tag{51}$$

and for an invariant scalar β_{ij}, the matrix exponential preserves the same conjugation law:

$$\mathbf{R}_{ij} = \exp(\beta_{ij}\mathbf{\Omega}_{ij}) \mapsto Q\mathbf{R}_{ij}Q^\top \tag{52}$$

Since $\mathbf{\Omega}_{ij}$ is skew-symmetric, $\mathbf{R}_{ij} \in SO(n)$, recovering the spatial factor used by Orthogonal Transport in Section 4.

These constructions explain why the displacement-only completeness result does not extend directly to feature-conditioned transport. Additional covariant vector features provide geometric information beyond the single direction $\mathbf{r}_{ij}$, so the stabilizer argument restricting the spatial action to $\mathbf{P}^\parallel$ and $\mathbf{P}^\perp$ no longer applies in general.

Unified Transport. Displacement- and feature-conditioned operators can be combined within the general transport form of Equation 4. Let

$$\mathcal{K} \subseteq \{\mathrm{id}, \parallel, \perp, \mathrm{skew}\}$$

denote the active spatial components, with

$$\mathbf{B}^{(\mathrm{id})}_{ij} = \mathbf{I}_n \qquad \mathbf{B}^{(\parallel)}_{ij} = \mathbf{P}^\parallel_{ij} \qquad \mathbf{B}^{(\perp)}_{ij} = \mathbf{P}^\perp_{ij} \qquad \mathbf{B}^{(\mathrm{skew})}_{ij} = \widehat{\mathbf{\Omega}}^V_{ij} \tag{53}$$

The corresponding transport is

$$\mathcal{T}_{i\leftarrow j}(\mathbf{V}_j) = \sum_{b\in\mathcal{K}} \mathbf{B}^{(b)}_{ij}\mathbf{V}_j\mathbf{M}^{(b)}_{ij} \tag{54}$$

Each spatial component satisfies the covariance condition of Equation 5, so the resulting map is $O(n)$-equivariant whenever the corresponding channel maps are invariant.

This unified construction shows how learned covariant features can extend the spatial operator family beyond the displacement-only class characterized by Theorem 4.1. It provides a convenient way to combine several admissible spatial actions within a single transport map, without implying that this particular set of operators spans the complete class of feature-conditioned $O(n)$-equivariant transports.

C Cellular Sheaves, Directed Transport, and Connections

ESNN directly parameterizes transport between neighboring vector features. This appendix places these maps within the sheaf-theoretic perspective introduced in Section 2.1. We first relate classical cellular-sheaf restriction maps to the effective couplings they induce between node stalks, then show how parameterizing these couplings directly leads to a directed quiver formulation. We subsequently identify the additional conditions required for self-adjoint transport and for an exact connection-sheaf realization. We also distinguish the common ambient $O(n)$ action enforced by ESNN from independent changes of local reference frame.

C.1 Cellular Sheaves and Induced Node-to-Node Coupling

The connection between cellular sheaves and ESNN is most easily seen through the node-to-node interaction induced by the sheaf Laplacian. Let $G = (\mathcal{V}, \mathcal{E})$ be a graph. A cellular sheaf $\mathcal{F}$ assigns a vector space $\mathcal{F}(i)$ to each vertex $i \in \mathcal{V}$ and a vector space $\mathcal{F}(e)$ to each edge $e \in \mathcal{E}$. For every vertex–edge incidence $i \trianglelefteq e$, these local spaces are related by a linear restriction map:

$$\rho_{i \to e} : \mathcal{F}(i) \longrightarrow \mathcal{F}(e) \tag{55}$$

The restriction maps place the features associated with the endpoints of an edge in a common edge space where their compatibility can be compared. After choosing an orientation for each edge $e = \{i, j\}$, the degree-0 coboundary measures this disagreement as:

$$(\delta_{\mathcal{F}} \mathbf{h})_e = \rho_{i \to e} \mathbf{h}_i - \rho_{j \to e} \mathbf{h}_j \tag{56}$$

up to the chosen orientation convention. With standard Euclidean inner products, the corresponding cellular-sheaf Laplacian is $\mathbf{L}_{\mathcal{F}} = \delta_{\mathcal{F}}^* \delta_{\mathcal{F}}$ where * denotes the adjoint. For a single edge $e = \{i, j\}$, its contribution to the Laplacian on the two endpoint stalks is:

$$\mathbf{L}_e = \begin{pmatrix} \rho_{i \to e}^* \rho_{i \to e} & -\rho_{i \to e}^* \rho_{j \to e} \\ -\rho_{j \to e}^* \rho_{i \to e} & \rho_{j \to e}^* \rho_{j \to e} \end{pmatrix} \tag{57}$$

The off-diagonal blocks reveal the effective interaction between neighboring node stalks. In particular, information from node j contributes to node i through the composition:

$$\rho_{i \to e}^* \rho_{j \to e} : \mathcal{F}(j) \longrightarrow \mathcal{F}(i) \tag{58}$$

up to the conventional minus sign in the Laplacian. The restriction at j maps the feature into the common edge space while the adjoint of the restriction at i maps the resulting representation back to the stalk at i. Their composition therefore acts as an induced linear transport between the neighboring node spaces. This effective node-to-node coupling is the point of departure for ESNN. Rather than parameterizing separate restriction maps through an explicit edge stalk and obtaining their composition indirectly, ESNN parameterizes the neighboring vector transport directly:

$$\mathcal{T}_{i \leftarrow j} : \mathcal{F}_{\mathrm{vec}}(j) \longrightarrow \mathcal{F}_{\mathrm{vec}}(i) \tag{59}$$

The transport is then constrained to satisfy the transformation law of the ambient Euclidean representation as developed in Sections 3 and 4. This retains the sheaf perspective of local feature spaces connected by linear maps while allowing the two orientations of an interaction to be parameterized directly. In general, however, the resulting ESNN operator need not arise from a classical cellular-sheaf Laplacian. The conditions under which more structured sheaf realizations are recovered are developed below.

C.2 Directed Transport as a Quiver Representation

Directly parameterizing node-to-node transport naturally accommodates asymmetric interactions. A convenient language for describing this structure is that of quiver representations. A quiver is a directed graph whose vertices are assigned vector spaces and whose arrows are assigned linear maps between those spaces. In ESNN, the vertices correspond to node vector stalks and the arrows to the directed transport maps used during message passing.

For a given layer state, once the context-dependent transport operators have been determined, let

$$\mathcal{Q} = (\mathcal{V}, \mathcal{A}) \tag{60}$$

denote the directed interaction graph with an arrow $j \to i$ for every directed message-passing interaction. Assign the vector space:

$$\mathcal{F}_{\mathrm{vec}}(i) = \mathbb{R}^n \otimes \mathbb{R}^{c_v} \tag{61}$$

to each vertex i and the realized linear transport:

$$\mathcal{T}_{i \leftarrow j} : \mathcal{F}_{\mathrm{vec}}(j) \longrightarrow \mathcal{F}_{\mathrm{vec}}(i) \tag{62}$$

to each arrow $j \to i$. These assignments define a representation of the quiver $\mathcal{Q}$ for the realized transport operators at that layer.

Equivalently, the construction extends to the free path category of $\mathcal{Q}$. Each directed path

$$i_0 \longrightarrow i_1 \longrightarrow \cdots \longrightarrow i_m \tag{63}$$

is associated with the composition

$$\mathcal{T}_{i_m \leftarrow i_{m-1}} \circ \cdots \circ \mathcal{T}_{i_1 \leftarrow i_0} \tag{64}$$

so the vertex spaces and directed transports extend naturally to a covariant functor from the free path category of $\mathcal{Q}$ to finite-dimensional vector spaces. This categorical viewpoint is not required to define an ESNN layer but makes explicit that directed transports compose consistently along paths.

The ambient equivariance of the individual edge maps is preserved under this composition. If every arrow satisfies:

$$\mathcal{T}'_{i \leftarrow j} \circ Q = Q \circ \mathcal{T}_{i \leftarrow j} \tag{65}$$

where Q denotes the natural $O(n)$ action on $\mathcal{F}_{\mathrm{vec}} = \mathbb{R}^n \otimes \mathbb{R}^{c_v}$, then for any path $i_0 \to i_1 \to \cdots \to i_m$

$$\mathcal{T}'_{i_m \leftarrow i_{m-1}} \circ \cdots \circ \mathcal{T}'_{i_1 \leftarrow i_0} \circ Q = Q \circ \mathcal{T}_{i_m \leftarrow i_{m-1}} \circ \cdots \circ \mathcal{T}_{i_1 \leftarrow i_0} \tag{66}$$

Thus the transport associated with a path obeys the same global $O(n)$ intertwining relation as the individual edge maps. This describes the algebraic composition of the transports while a single ESNN layer still performs the one-hop aggregation defined in Section 5.

The quiver formulation does not require any compatibility condition between opposite edge orientations. If both $i \to j$ and $j \to i$ are present, the general ESNN construction does not impose:

$$\mathcal{T}_{i \leftarrow j} = \mathcal{T}^*_{j \leftarrow i}, \qquad \mathcal{T}_{i \leftarrow j} = \mathcal{T}^{-1}_{j \leftarrow i} \tag{67}$$

The two directions may therefore represent genuinely different local interactions. This distinguishes the general directional ESNN operator from diffusion generated by a classical self-adjoint sheaf Laplacian and is consistent with recent directed extensions of sheaf-based learning (Ribeiro et al., 2026; Fiorini et al., 2026).

This perspective is also closely related to copresheaf neural constructions which organize information through directed maps between local feature spaces and their compositions (Hajij et al., 2025). ESNN shares this directed and compositional view of information flow while imposing an additional geometric requirement: each transport must intertwine the common ambient $O(n)$ action carried by the vector features. The resulting structure is therefore directional at the level of the interaction graph and equivariant with respect to the ambient geometry.

The transport operators are state dependent, with coefficients determined from the current node, edge, and geometric features as described in Section 3. Once these coefficients are fixed, each realized edge operator retains the linear spatial–channel form introduced above, even though the overall message function may depend nonlinearly on the evolving feature state.

C.3 Adjoint-Consistent and Connection-Sheaf Specializations

The directed formulation of Appendix C.2 allows the two orientations of an edge to carry independent transport maps. Classical sheaf diffusion is more structured: the off-diagonal blocks of a cellular-sheaf Laplacian occur in adjoint pairs which makes the resulting operator self-adjoint. We first identify the corresponding specialization of ESNN and then give the stronger conditions under which the edge transport admits an exact connection-sheaf realization.

Recall the normalized transport operator:

$$(\mathcal{A}_{\mathcal{T}}\mathbf{V})_i = \nu_{ii}\mathbf{V}_i + \sum_{j \in \mathcal{N}(i)} \nu_{ij}\mathcal{T}_{i \leftarrow j}(\mathbf{V}_j) \tag{68}$$

For the left–right transport form introduced in Equation 4

$$\mathcal{T}(\mathbf{V}) = \sum_k \mathbf{S}^{(k)}\mathbf{V}\mathbf{M}^{(k)} \tag{69}$$

the adjoint with respect to the Frobenius inner product is:

$$\mathcal{T}^*(\mathbf{U}) = \sum_k \left(\mathbf{S}^{(k)}\right)^\top \mathbf{U} \left(\mathbf{M}^{(k)}\right)^\top \tag{70}$$

Indeed

$$\langle \mathbf{USVM} \rangle_F = \left\langle \mathbf{S}^\top \mathbf{U} \mathbf{M}^\top \mathbf{V} \right\rangle_F \tag{71}$$

For Radial–Tangential Transport, the spatial projectors are symmetric so the adjoint acts by transposing the corresponding channel maps:

$$\mathcal{T}^*(\mathbf{U}) = \mathbf{P}^{\parallel}\mathbf{U}\left(\mathbf{M}^{\parallel}\right)^\top + \mathbf{P}^{\perp}\mathbf{U}\left(\mathbf{M}^{\perp}\right)^\top \tag{72}$$

This makes the condition for self-adjoint transport explicit. On a bidirected graph, the scalar normalization must agree across the two orientations and the transport in one direction must be the adjoint of the transport in the other.

Proposition C.1 (Self-Adjoint Transport Operator)**.** *Fix the layer state so that the transport maps and aggregation weights are determined. Suppose the graph is bidirected and that for every adjacent pair i, j*

$$\nu_{ij} = \nu_{ji} \qquad \mathcal{T}_{j\leftarrow i} = \mathcal{T}^*_{i\leftarrow j} \tag{73}$$

Then $\mathcal{A}_\mathcal{T}$ is self-adjoint with respect to the direct-sum Frobenius inner product on the node vector features.

Proof. For two node signals $\mathbf{U}$ and $\mathbf{V}$, an off-diagonal contribution to $\langle \mathbf{U}\mathcal{A}_\mathcal{T}\mathbf{V}\rangle$ is

$$\nu_{ij}\left\langle \mathbf{U}_i \mathcal{T}_{i\leftarrow j}\mathbf{V}_j \right\rangle_F \tag{74}$$

Using the adjoint relation together with $\nu_{ij} = \nu_{ji}$

$$\nu_{ij}\left\langle \mathbf{U}_i \mathcal{T}_{i\leftarrow j}\mathbf{V}_j \right\rangle_F = \nu_{ji}\left\langle \mathcal{T}_{j\leftarrow i}\mathbf{U}_i \mathbf{V}_j \right\rangle_F \tag{75}$$

Summing over the bidirected edge set gives

$$\langle \mathbf{U}\mathcal{A}_\mathcal{T}\mathbf{V}\rangle = \langle \mathcal{A}_\mathcal{T}\mathbf{U}\mathbf{V}\rangle \tag{76}$$

The diagonal self-loop terms are real scalar multiples of the identity and are therefore self-adjoint. □

Proposition C.1 identifies the adjoint-consistent specialization of the general directional operator. This recovers the self-adjoint structure characteristic of classical sheaf diffusion, but does not by itself make $\mathcal{A}_\mathcal{T}$ a cellular-sheaf Laplacian. An exact connection-sheaf realization additionally requires the transport between neighboring fibers to be orthogonal, with the reverse direction given by its inverse and hence by its adjoint. Under these stronger conditions, the ESNN edge transport can be represented exactly through classical sheaf restriction maps.

Proposition C.2 (Connection-Sheaf Realization)**.** *Let $\mathcal{H} = \mathbb{R}^n \otimes \mathbb{R}^{c_v}$ be the vector fiber and consider an undirected edge $e = \{i, j\}$. Suppose the full transport*

$$\mathcal{U}_{i\leftarrow j} : \mathcal{H} \longrightarrow \mathcal{H} \tag{77}$$

is orthogonal and that the reverse transport satisfies

$$\mathcal{U}_{j\leftarrow i} = \mathcal{U}^*_{i\leftarrow j} = \mathcal{U}^{-1}_{i\leftarrow j} \tag{78}$$

Choose vertex and edge stalks equal to $\mathcal{H}$ and restrictions

$$\rho_{i\to e} = \mathbf{I}_\mathcal{H} \qquad \rho_{j\to e} = \mathcal{U}_{i\leftarrow j} \tag{79}$$

Then the contribution of e to the cellular-sheaf Laplacian is

$$\mathbf{L}_e = \begin{pmatrix} \mathbf{I}_\mathcal{H} & -\mathcal{U}_{i\leftarrow j} \\ -\mathcal{U}^*_{i\leftarrow j} & \mathbf{I}_\mathcal{H} \end{pmatrix} \tag{80}$$

Thus the off-diagonal node couplings are exactly the orthogonal transport and its adjoint up to the conventional Laplacian sign.

Proof. Substituting Equation 79 into Equation 57 gives

$$\rho^*_{i\to e}\rho_{i\to e} = \mathbf{I}_\mathcal{H} \qquad \rho^*_{i\to e}\rho_{j\to e} = \mathcal{U}_{i\leftarrow j} \tag{81}$$

Since $\mathcal{U}_{i\leftarrow j}$ is orthogonal

$$\rho^*_{j\to e}\rho_{j\to e} = \mathcal{U}^*_{i\leftarrow j}\mathcal{U}_{i\leftarrow j} = \mathbf{I}_\mathcal{H} \tag{82}$$

while

$$\rho^*_{j\to e}\rho_{i\to e} = \mathcal{U}^*_{i\leftarrow j} \tag{83}$$

Equation 80 follows. □

Proposition C.2 gives the discrete orthogonal-connection counterpart of ESNN transport and is closely related to connection-Laplacian SNNs (Barbero et al., 2022a). In this sense, Orthogonal Transport matches the required geometry through its spatial factor $\mathbf{R}_{ij} \in SO(n)$. The full ESNN edge map, however, also includes channel mixing and edge-dependent gating, and is therefore not generally orthogonal on $\mathbb{R}^n \otimes \mathbb{R}^{c_v}$. The exact connection-sheaf realization is recovered when the complete fiber transport is orthogonal and the reverse direction is given by its adjoint.

This yields a simple hierarchy. General ESNN transport allows independent directed edge maps. Enforcing Equation 73 gives a self-adjoint transport operator, while additionally requiring orthogonality of the full fiber maps recovers the classical connection-sheaf case of Proposition C.2.

C.4 Ambient Equivariance and Local Gauge Transformations

The connection interpretation above should not be confused with local gauge equivariance. In a classical cellular sheaf, the bases of different stalks may be changed independently. If $\mathbf{Q}_i$ and $\mathbf{Q}_e$ are orthogonal changes of basis on a vertex stalk and an incident edge stalk, the coordinate representation of the corresponding restriction map transforms as:

$$\rho'_{i\to e} = \mathbf{Q}_e \rho_{i\to e} \mathbf{Q}_i^\top \tag{84}$$

This expresses covariance under independent changes of local reference frame. ESNN enforces a different symmetry. A single orthogonal transformation $Q \in O(n)$ acts simultaneously on the physical coordinates and all vector features:

$$\mathbf{x}_i \mapsto Q\mathbf{x}_i + \mathbf{t}, \qquad \mathbf{V}_i \mapsto Q\mathbf{V}_i \tag{85}$$

while the spatial transport transforms as:

$$\mathbf{S}_{ij} \mapsto Q\mathbf{S}_{ij}Q^\top \tag{86}$$

This is the ambient $E(n)$-equivariance established in Sections 3–5. It corresponds to transforming the entire geometric system in a common Cartesian frame rather than independently changing the basis attached to each node.

In particular, the feature-derived Orthogonal Transport is not designed to be covariant under arbitrary node-wise frame changes. For $\mathbf{C}_{ij} = \mathbf{V}_j\mathbf{V}_i^\top$ independent transformations:

$$\mathbf{V}_i \mapsto \mathbf{Q}_i\mathbf{V}_i, \qquad \mathbf{V}_j \mapsto \mathbf{Q}_j\mathbf{V}_j \tag{87}$$

give:

$$\mathbf{C}'_{ij} = \mathbf{Q}_j\mathbf{C}_{ij}\mathbf{Q}_i^\top \tag{88}$$

and the skew-symmetric construction used by Orthogonal Transport does not in general transform as a gauge-covariant map between the two independently chosen frames. When $\mathbf{Q}_i = \mathbf{Q}_j = Q$, however, the same construction reduces to conjugation by the common ambient transformation which is precisely the covariance required by ESNN. Thus, Orthogonal Transport has a connection-like geometric interpretation under the common ambient frame but is not a general gauge-equivariant transport between independently chosen local frames.

A gauge-equivariant ESNN would instead associate each node with its own local frame and require directed transports to transform as:

$$\mathcal{T}'_{i\leftarrow j} = \mathbf{Q}_i\mathcal{T}_{i\leftarrow j}\mathbf{Q}_j^\top \tag{89}$$

under independent local transformations $\mathbf{Q}_i$ and $\mathbf{Q}_j$. Such a construction would require transport operators built from gauge-covariant quantities rather than from the common-frame feature representations used here. Extending ESNN to this locally gauge-equivariant setting is left for future work.

D Proofs of Equivariance and Symmetry Results

This appendix provides the proofs of the equivariance and symmetry results stated in the main text. The algebraic proof of the completeness of displacement-conditioned linear transport is given separately in Appendix B.2, while the self-adjoint and connection-sheaf specializations are established in Appendix C.

D.1 Proof of Proposition 3.2: $O(n)$-Equivariance of the Transport Map

Proposition 3.2 ($O(n)$-Equivariance of the Transport Map). *Suppose that, for every component k, the spatial operator satisfies*

$$\mathbf{S}_{ij}^{(k)}(Q\cdot\mathcal{C}_{ij}) = Q\mathbf{S}_{ij}^{(k)}(\mathcal{C}_{ij})Q^\top \tag{90}$$

and the channel operator $\mathbf{M}_{ij}^{(k)}$ is unchanged under the $O(n)$ action. Then the transport map in Equation 4 is $O(n)$-equivariant

$$\mathcal{T}'_{i\leftarrow j}(Q\mathbf{V}_j) = Q\,\mathcal{T}_{i\leftarrow j}(\mathbf{V}_j) \tag{91}$$

where $\mathcal{T}'_{i\leftarrow j}$ denotes the transport evaluated from the transformed edge context.

Proof. Let $Q \in O(n)$ and consider the transport evaluated from the transformed edge context $Q\cdot\mathcal{C}_{ij}$. By assumption

$$\mathbf{S}_{ij}^{(k)}(Q\cdot\mathcal{C}_{ij}) = Q\mathbf{S}_{ij}^{(k)}(\mathcal{C}_{ij})Q^\top \tag{92}$$

while the channel operators satisfy

$$\mathbf{M}_{ij}^{(k)}(Q\cdot\mathcal{C}_{ij}) = \mathbf{M}_{ij}^{(k)}(\mathcal{C}_{ij}) \tag{93}$$

Evaluating the transport on the transformed vector feature gives

$$\mathcal{T}'_{i\leftarrow j}(Q\mathbf{V}_j) = \sum_{k=1}^{K}\left(Q\mathbf{S}_{ij}^{(k)}Q^\top\right)(Q\mathbf{V}_j)\mathbf{M}_{ij}^{(k)} \tag{94}$$

$$= Q\sum_{k=1}^{K}\mathbf{S}_{ij}^{(k)}\mathbf{V}_j\mathbf{M}_{ij}^{(k)} \tag{95}$$

$$= Q\,\mathcal{T}_{i\leftarrow j}(\mathbf{V}_j) \tag{96}$$

which is the required $O(n)$-equivariance relation. □

D.2 $O(n)$-Equivariance of Orthogonal Transport

Proposition D.1 (Equivariance of Orthogonal Transport)**.** *Let $Q \in O(n)$ act on the vector features as $\mathbf{V}_i \mapsto Q\mathbf{V}_i$. Then the Orthogonal Transport of Equation 13 satisfies*

$$\mathcal{T}'_{i\leftarrow j}(Q\mathbf{V}_j) = Q\,\mathcal{T}_{i\leftarrow j}(\mathbf{V}_j) \tag{97}$$

where the primed transport is evaluated from the transformed edge context.

Proof. Recall that Orthogonal Transport can be written as

$$\mathcal{T}_{i\leftarrow j}(\mathbf{V}_j) = \mathbf{R}_{ij}\mathbf{V}_j\mathbf{M}_{ij} \qquad \mathbf{M}_{ij} = \mathbf{W}\mathbf{D}(\boldsymbol{\eta}_{ij}) \tag{98}$$

with

$$\mathbf{R}_{ij} = \exp(\beta_{ij}\boldsymbol{\Omega}_{ij}) \tag{99}$$

and

$$\boldsymbol{\Omega}_{ij} = \widetilde{\mathbf{C}}_{ij} - \widetilde{\mathbf{C}}_{ij}^\top \qquad \widetilde{\mathbf{C}}_{ij} = \frac{\mathbf{C}_{ij}}{\|\mathbf{C}_{ij}\|_F + \varepsilon} \qquad \mathbf{C}_{ij} = \mathbf{V}_j\mathbf{V}_i^\top \tag{100}$$

Since $\boldsymbol{\Omega}_{ij}^\top = -\boldsymbol{\Omega}_{ij}$, we have $\beta_{ij}\boldsymbol{\Omega}_{ij} \in \mathfrak{so}(n)$ and therefore $\mathbf{R}_{ij} \in SO(n)$.

Under the transformations $\mathbf{V}'_i = Q\mathbf{V}_i$ and $\mathbf{V}'_j = Q\mathbf{V}_j$ the cross-feature matrix becomes

$$\mathbf{C}'_{ij} = (Q\mathbf{V}_j)(Q\mathbf{V}_i)^\top = Q\mathbf{C}_{ij}Q^\top \tag{101}$$

The Frobenius norm is invariant under orthogonal conjugation, hence

$$\widetilde{\mathbf{C}}'_{ij} = Q\widetilde{\mathbf{C}}_{ij}Q^\top \tag{102}$$

and consequently

$$\mathbf{\Omega}'_{ij} = Q\mathbf{\Omega}_{ij}Q^\top \tag{103}$$

The coefficient β_{ij} is predicted from invariant edge features and therefore $\beta'_{ij} = \beta_{ij}$. The matrix exponential is compatible with similarity transformations

$$\exp(Q\mathbf{A}Q^{-1}) = Q\exp(\mathbf{A})Q^{-1} \tag{104}$$

and since $Q^{-1} = Q^\top$

$$\mathbf{R}'_{ij} = Q\mathbf{R}_{ij}Q^\top \tag{105}$$

The gating vector $\boldsymbol{\eta}_{ij}$ is likewise constructed from invariant quantities, so $\mathbf{M}'_{ij} = \mathbf{M}_{ij}$. Evaluating the transport from the transformed context gives

$$\mathcal{T}'_{i\leftarrow j}(Q\mathbf{V}_j) = (Q\mathbf{R}_{ij}Q^\top)(Q\mathbf{V}_j)\mathbf{M}_{ij} \tag{106}$$
$$= Q\mathbf{R}_{ij}\mathbf{V}_j\mathbf{M}_{ij} \tag{107}$$
$$= Q\,\mathcal{T}_{i\leftarrow j}(\mathbf{V}_j) \tag{108}$$

which proves the claim. □

D.3 $O(n)$-Equivariance of Radial–Tangential Transport

Proposition D.2 (Equivariance of Radial–Tangential Transport)**.** *For $\mathbf{r}_{ij} \neq \mathbf{0}$, the Radial–Tangential Transport*

$$\mathcal{T}_{i\leftarrow j}(\mathbf{V}_j) = \mathbf{P}^{\parallel}_{ij}\mathbf{V}_j\mathbf{M}^{\parallel}_{ij} + \mathbf{P}^{\perp}_{ij}\mathbf{V}_j\mathbf{M}^{\perp}_{ij} \tag{109}$$

is $O(n)$-equivariant.

Proof. Let $Q \in O(n)$. Since $\mathbf{r}'_{ij} = Q\mathbf{r}_{ij}$ and Q is orthogonal

$$\|\mathbf{r}'_{ij}\| = \|\mathbf{r}_{ij}\| \tag{110}$$

and therefore

$$\widehat{\mathbf{r}}'_{ij} = Q\widehat{\mathbf{r}}_{ij} \tag{111}$$

The radial projector transforms by conjugation

$$(\mathbf{P}^{\parallel}_{ij})' = \widehat{\mathbf{r}}'_{ij}(\widehat{\mathbf{r}}'_{ij})^\top = Q\mathbf{P}^{\parallel}_{ij}Q^\top \tag{112}$$

Since $\mathbf{P}^{\perp}_{ij} = \mathbf{I}_n - \mathbf{P}^{\parallel}_{ij}$ we also obtain

$$(\mathbf{P}^{\perp}_{ij})' = Q\mathbf{P}^{\perp}_{ij}Q^\top \tag{113}$$

The channel maps $\mathbf{M}^{\parallel}_{ij}$ and $\mathbf{M}^{\perp}_{ij}$ are constructed from invariant edge features and are therefore unchanged by the $O(n)$ action. Hence

$$\mathcal{T}'_{i\leftarrow j}(Q\mathbf{V}_j) = (Q\mathbf{P}^{\parallel}_{ij}Q^\top)(Q\mathbf{V}_j)\mathbf{M}^{\parallel}_{ij} + (Q\mathbf{P}^{\perp}_{ij}Q^\top)(Q\mathbf{V}_j)\mathbf{M}^{\perp}_{ij} \tag{114}$$
$$= Q\left(\mathbf{P}^{\parallel}_{ij}\mathbf{V}_j\mathbf{M}^{\parallel}_{ij} + \mathbf{P}^{\perp}_{ij}\mathbf{V}_j\mathbf{M}^{\perp}_{ij}\right) \tag{115}$$
$$= Q\,\mathcal{T}_{i\leftarrow j}(\mathbf{V}_j) \tag{116}$$

which proves the claim. □

D.4 Completeness of Displacement-Conditioned Linear Transport

Theorem 4.1 is proved in Appendix B.2. The argument characterizes the commutant of the stabilizer

$$H_{\mathbf{r}} = \{Q \in O(n) : Q\mathbf{r} = \mathbf{r}\} \cong O(n-1)$$

for a non-zero displacement. This yields the two spatial projectors $\mathbf{P}^{\parallel}$ and $\mathbf{P}^{\perp}$ together with arbitrary channel endomorphisms whose dependence on the geometric conditioning is $O(n)$-invariant.

D.5 Proof of Theorem 5.1: $E(n)$-Equivariance of the ESNN Layer

Theorem 5.1 ($E(n)$-Equivariance). *Assume that the graph topology is fixed or constructed from $E(n)$-invariant quantities, that edge attributes are $O(n)$-invariant, and that the transport maps satisfy Proposition 3.2. Assume also that $\mathbf{r}_{ij} \neq \mathbf{0}$ whenever a normalized edge direction is used. In the absence of the symmetry-relaxing inputs introduced in Section 6, the ESNN layer of Section 5 is $E(n)$-equivariant.*

Proof. Let $(Q, \mathbf{t}) \in E(n)$ with $Q \in O(n)$ and transform

$$\overline{\mathbf{x}}_i = Q\mathbf{x}_i + \mathbf{t} \qquad \overline{\mathbf{V}}_i = Q\mathbf{V}_i \qquad \overline{\mathbf{s}}_i = \mathbf{s}_i \tag{117}$$

For the first layer, the lift condition in Equation 18 ensures that any task-specific vector initialization obeys the same transformation law. By assumption, the graph topology and edge attributes are unchanged.

For each interaction

$$\overline{\mathbf{r}}_{ij} = \overline{\mathbf{x}}_i - \overline{\mathbf{x}}_j = Q\mathbf{r}_{ij} \tag{118}$$

Orthogonality preserves distances, vector norms, and inner products. All quantities entering the invariant construction of Equation 19 are therefore unchanged

$$\overline{\mathbf{z}}_{ij} = \mathbf{z}_{ij} \tag{119}$$

By Proposition 3.2, the vector messages satisfy

$$\overline{\mathbf{m}}^V_{i\leftarrow j} = Q\mathbf{m}^V_{i\leftarrow j} \tag{120}$$

The scalar messages depend only on invariant quantities and therefore satisfy

$$\overline{\mathbf{m}}^s_{i\leftarrow j} = \mathbf{m}^s_{i\leftarrow j} \tag{121}$$

The normalized coefficients ν_{ij} and ν_{ii} are invariant scalars. Equation 21 therefore gives

$$\overline{\mathbf{V}}^{\mathrm{agg}}_i = Q\mathbf{V}^{\mathrm{agg}}_i \qquad \overline{\mathbf{s}}^{\mathrm{agg}}_i = \mathbf{s}^{\mathrm{agg}}_i \tag{122}$$

The matrix $\mathbf{W}_V$ acts only on vector channels and therefore commutes with the spatial action

$$(Q\mathbf{V}^{\mathrm{agg}}_i)\mathbf{W}_V = Q(\mathbf{V}^{\mathrm{agg}}_i\mathbf{W}_V) \tag{123}$$

The radial non-linearity is equivariant because

$$\sigma_V(Q\mathbf{v}) = a(\|Q\mathbf{v}\|)Q\mathbf{v} = Q\sigma_V(\mathbf{v}) \tag{124}$$

The scalar update acts only on invariant quantities through channel maps and scalar nonlinearities and therefore remains invariant. Equation 22 consequently yields

$$\overline{\mathbf{V}}'_i = Q\mathbf{V}'_i \qquad \overline{\mathbf{s}}'_i = \mathbf{s}'_i \tag{125}$$

Finally, γ_{ij} in Equation 23 is predicted from invariant node and edge features and is therefore unchanged. Using Equation 118

$$\overline{\Delta\mathbf{x}}_i = \frac{1}{|\mathcal{N}(i)|}\sum_{j\in\mathcal{N}(i)} \gamma_{ij}Q\mathbf{r}_{ij} = Q\Delta\mathbf{x}_i \tag{126}$$

and hence

$$\overline{\mathbf{x}}'_i = Q\mathbf{x}'_i + \mathbf{t} \tag{127}$$

The scalar, vector, and coordinate outputs therefore satisfy the required $E(n)$ transformation laws. □

Remark (Invariant attention). The same argument applies when the coefficients ν_{ij} are obtained from receiver-normalized attention over invariant edge features. The attention weights remain invariant under $E(n)$, so replacing degree-normalized coefficients by these invariant scalars does not alter the covariance of the aggregated vector messages.

D.6 Proof of Theorem 6.1: Stabilizer Subequivariance

Theorem 6.1 (Stabilizer Subequivariance). *Let $\mathbf{g} \neq \mathbf{0}$ be held fixed under transformations of the input geometry and define*

$$O_{\mathbf{g}}(n) = \{Q \in O(n) : Q\mathbf{g} = \mathbf{g}\} \tag{128}$$

An ESNN conditioned on Equation 25 is equivariant under translations and every $Q \in O_{\mathbf{g}}(n)$ with guaranteed symmetry group

$$E_{\mathbf{g}}(n) = O_{\mathbf{g}}(n) \ltimes \mathbb{R}^n \tag{129}$$

When $\lambda_{\mathbf{g}}^{(\ell)} = 0$ for every layer, the directional conditioning vanishes and full $E(n)$-equivariance is recovered.

Proof. At layer ℓ, the directional feature is

$$z_{ij}^{\mathbf{g},(\ell)} = \lambda_{\mathbf{g}}^{(\ell)} \langle \mathbf{r}_{ij}, \mathbf{g} \rangle \tag{130}$$

Translations leave $\mathbf{r}_{ij}$ unchanged, so this feature is translation invariant.

Now let $Q \in O_{\mathbf{g}}(n)$. Since $Q\mathbf{g} = \mathbf{g}$ and Q is orthogonal, $Q^\top \mathbf{g} = \mathbf{g}$. Therefore

$$\langle Q\mathbf{r}_{ij}, \mathbf{g} \rangle = \langle \mathbf{r}_{ij}, Q^\top \mathbf{g} \rangle = \langle \mathbf{r}_{ij}, \mathbf{g} \rangle \tag{131}$$

Because $\lambda_{\mathbf{g}}^{(\ell)}$ is a scalar model parameter, $z_{ij}^{\mathbf{g},(\ell)}$ is invariant under $O_{\mathbf{g}}(n)$ at every layer.

The remaining edge features are invariant by Equation 119. Hence all scalar coefficients generated from the augmented edge context remain unchanged under $O_{\mathbf{g}}(n)$, while the spatial transport operators retain the covariance laws of Section 4. The proof of Theorem 5.1 therefore applies with the orthogonal action restricted from $O(n)$ to $O_{\mathbf{g}}(n)$. Together with translation equivariance, this gives

$$E_{\mathbf{g}}(n) = O_{\mathbf{g}}(n) \ltimes \mathbb{R}^n \tag{132}$$

For a transformation $Q \notin O_{\mathbf{g}}(n)$, the directional projection is not invariant in general because

$$Q^\top \mathbf{g} \neq \mathbf{g} \tag{133}$$

and hence there exist displacements for which

$$\langle Q\mathbf{r}, \mathbf{g} \rangle \neq \langle \mathbf{r}, \mathbf{g} \rangle \tag{134}$$

If $\lambda_{\mathbf{g}}^{(\ell)} \neq 0$ for at least one layer, the architecture therefore does not guarantee equivariance outside $O_{\mathbf{g}}(n)$. Particular learned parameters may ignore the directional feature and realize a larger symmetry, but that larger symmetry is not enforced by the construction.

Finally, if $\lambda_{\mathbf{g}}^{(\ell)} = 0$ for every layer, all directional features vanish and the network reduces to the fully $E(n)$-equivariant architecture of Section 5. □

Remark (Learned preferred direction). The same symmetry statement applies when $\mathbf{g}$ is learned jointly with the network provided its learned value is non-zero. For fixed model parameters, $\mathbf{g}$ is a single global ambient parameter and is not transformed together with the input. The guaranteed symmetry is therefore the stabilizer of its learned value. If $\mathbf{g} = \mathbf{0}$, the directional feature vanishes and no symmetry reduction is induced.

E Experimental Details

This appendix describes the datasets, geometric inputs, task-specific lifting procedures, training objectives, model selection, and evaluation protocols used in our experiments. Selected configurations and model-selection details are summarized in Appendix E.6.

E.1 Charged N-Body Dynamics

Dataset and prediction task. We follow the standard charged N-body benchmark protocol used for EGNN comparisons in prior work (Satorras et al., 2021). Each example contains five charged particles in three dimensions. The input consists of the initial position $\mathbf{x}_i$, velocity $\mathbf{v}_i$, and charge $q_i \in \{-1, +1\}$ of every particle, and the target is the particle position after $1{,}000$ simulation steps:

$$\left\{\mathbf{x}_i^{(0)}, \mathbf{v}_i^{(0)}, q_i\right\}_{i=1}^{5} \longmapsto \left\{\widehat{\mathbf{x}}_i^{(1000)}\right\}_{i=1}^{5} \tag{135}$$

The fixed train/validation/test split contains $3{,}000/2{,}000/2{,}000$ simulated systems. Initial positions are Gaussian, initial velocity directions are Gaussian and rescaled to norm 0.5, and charges are sampled independently and uniformly from $\{-1, +1\}$. No observation noise is added.

The charged-particle literature baselines in Table 1(a) use evaluations of the same five-particle benchmark. SE(3)-Transformer, TFN, Radial Field, and GMN follow the five-run comparison of Liu et al. (2024b); EGNN and GVP-GNN follow Liu et al. (2024a); and FA-GNN follows Levy et al. (2023). Their uncertainties are reported as given in the corresponding sources.

Graph and lifting. Each system is represented as a complete directed graph without self-loops. ESNN uses the charge q_i as the invariant scalar feature $\mathbf{s}_i$, the particle position as the coordinate state, and the velocity $\mathbf{v}_i$ as a covariant vector input. The lift initializes the first vector group of $\mathbf{V}_i$ from $\mathbf{v}_i$, replicated across its vector channels. Although the Identity model has no latent vector group, its coordinate update remains conditioned on velocity.

Training and evaluation. Models minimize mean squared coordinate error,

$$\mathcal{L}_{\mathrm{NBody}} = \frac{1}{BN} \sum_{b=1}^{B} \sum_{i=1}^{N} \left\| \widehat{\mathbf{x}}_{b,i}^{(1000)} - \mathbf{x}_{b,i}^{(1000)} \right\|_2^2 \tag{136}$$

Validation MSE selects the checkpoint, and the corresponding held-out test MSE is reported. All selected ESNN models use batch size 100, AdamW, cosine decay, twenty warm-up epochs, degree-normalized aggregation, and gradient clipping at 5.

E.2 Gravity-Augmented N-Body Dynamics

This task follows the same graph construction, inputs, data split, and coordinate-MSE objective as the charged setting, with an additional uniform gravitational acceleration

$$\mathbf{a}_{\mathbf{g}} = (0, 0, -9.81)^{\top}, \qquad \widehat{\mathbf{g}}_{\mathrm{true}} = (0, 0, -1)^{\top} \tag{137}$$

The model predicts particle positions $1{,}000$ integration steps ahead, with $\Delta t = 10^{-3}$ and therefore a physical horizon of 1.0:

$$\left\{\mathbf{x}_i^{(0)}, \mathbf{v}_i^{(0)}, q_i\right\}_{i=1}^{5} \longmapsto \left\{\widehat{\mathbf{x}}_i^{(1000)}\right\}_{i=1}^{5} \tag{138}$$

The simulation contains no floor or boundary conditions. For numerical stability, the summed Coulomb force is clipped componentwise before adding gravity. Translation symmetry remains exact, while the $O_{\mathbf{g}}(3)$ symmetry of the ideal dynamics is approximate on steps where the clipping is active. Figure 5 illustrates the charged and gravity-augmented settings.

Preferred-direction variants. For each non-identity transport we compare three settings. *None* contains no preferred-direction pathway and remains fully $E(3)$-equivariant. *Fixed* supplies the normalized physical direction $\mathbf{g} = \widehat{\mathbf{g}}_{\mathrm{true}}$ as a non-trainable vector, while *Learned* uses one global trainable vector $\mathbf{g}_\theta$, shared across layers and initialized as $\mathbf{g}_\theta = 10^{-2}\boldsymbol{\epsilon}$ with $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, I_3)$. No auxiliary alignment loss or direct supervision of $\mathbf{g}_\theta$ is used. Each active layer receives

$$z_{ij}^{\mathbf{g},(\ell)} = \lambda_{\mathbf{g}}^{(\ell)} \langle \mathbf{r}_{ij}, \mathbf{g} \rangle \tag{139}$$

where the layer-specific $\lambda_{\mathbf{g}}^{(\ell)}$ is initialized exactly at zero. Within each transport, the same selected backbone and training protocol are used for all three direction modes. Identity Transport is omitted because it has no directional transport pathway.

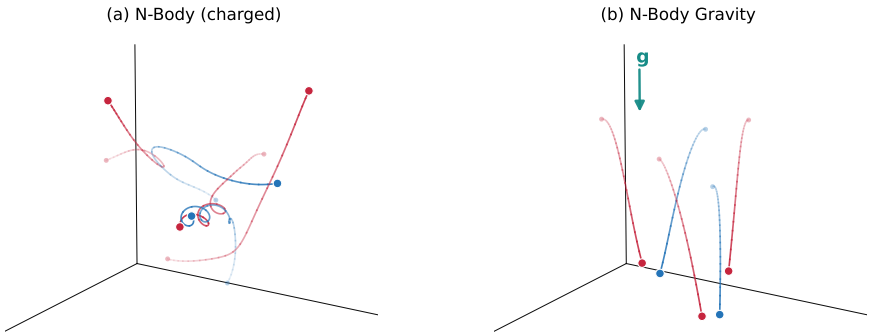

Figure 5: Charged-particle dynamics without and with a uniform gravitational field. Translation symmetry is preserved, while the ideal unclipped gravity–Coulomb dynamics retain only orthogonal transformations that preserve $\widehat{\mathbf{g}}_{\mathrm{true}}$.

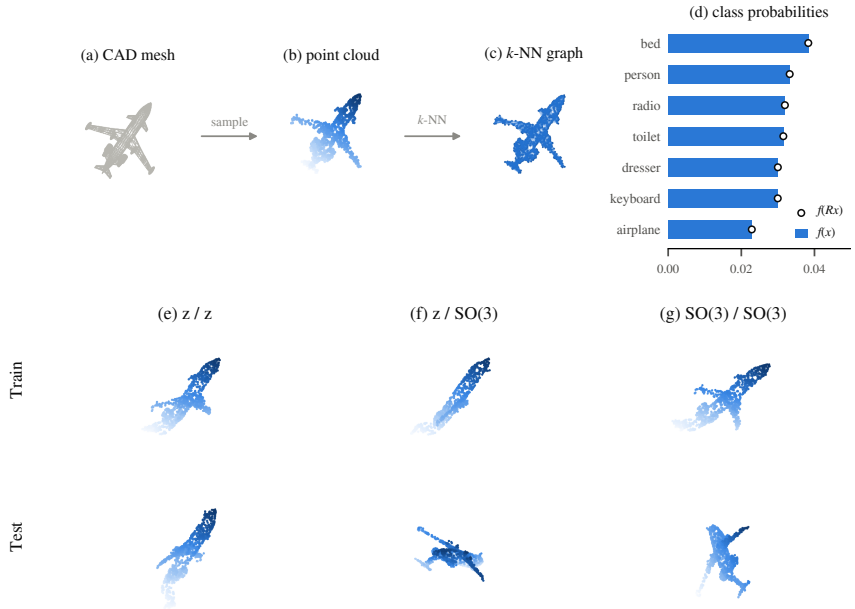

Figure 6: ModelNet40 graph construction and rotation protocols. Each centered point cloud is converted into a local k-nearest-neighbor graph and evaluated under the z/z, $z/\mathrm{SO}(3)$, and $\mathrm{SO}(3)/\mathrm{SO}(3)$ train/test regimes.

For the *Learned* setting we measure recovery of the gravity axis up to sign:

$$A_{\mathbf{g}} = \left| \left\langle \frac{\mathbf{g}_\theta}{\|\mathbf{g}_\theta\|_2}, \widehat{\mathbf{g}}_{\mathrm{true}} \right\rangle \right| \tag{140}$$

For both symmetry-relaxed settings we additionally report $\max_\ell |\lambda_{\mathbf{g}}^{(\ell)}| \, \|\mathbf{g}\|_2$, where $\mathbf{g} = \widehat{\mathbf{g}}_{\mathrm{true}}$ for *Fixed* and $\mathbf{g} = \mathbf{g}_\theta$ for *Learned*. This measures the effective scale of the directional pathway after training.

E.3 ModelNet40 Point-Cloud Classification

Figure 6 summarizes the point-cloud representation and rotation protocols used in this task.

Data and graph construction. We use the resampled ModelNet40 dataset (Wu et al., 2015), with $9{,}843$ training and $2{,}468$ test shapes. Each shape is represented by $1{,}024$ points, centered and rescaled to the unit ball, with surface normals discarded. A class-stratified 10% validation split is drawn from the training set using seed 42. We construct a symmetrized 20-nearest-neighbor graph without self-loops. The in-repository EGNN baseline uses the same data, split, graph construction, rotation protocols, and evaluation procedure as ESNN. Additional baselines in Table 2 are taken from Lippmann et al. (2025).

Lifting and readout. Each point receives the constant invariant feature $\mathbf{s}_i = 1$. The covariant lift is

$$\mathbf{V}_i^{\text{lift}} = \left[\mathbf{x}_i, \frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \mathbf{r}_{ij} \right], \tag{141}$$

combining the centered position with the mean incident displacement, followed by learned channel-wise scaling. Both components transform as polar vectors under $O(3)$. For classification, invariant vector-channel energies are pooled together with the scalar features and passed to a two-layer classifier.

Rotation protocols. We consider two augmentation regimes. For z-training, training and validation shapes are independently rotated about the vertical axis, and validation accuracy selects an EMA checkpoint. The same checkpoint is evaluated under vertical-axis test rotations (z/z) and independent Haar $\mathrm{SO}(3)$ rotations ($z/\mathrm{SO}(3)$), so the latter directly measures generalization to unseen three-dimensional rotations. For $\mathrm{SO}(3)/\mathrm{SO}(3)$, a separate model is trained and selected using Haar $\mathrm{SO}(3)$ augmentation before evaluation under independent $\mathrm{SO}(3)$ test rotations. Models are trained with cross-entropy and label smoothing, using an EMA decay of 0.999 for model selection and final evaluation. Reported EGNN and ESNN results use one run per selected configuration.

E.4 Mesh-Based Physical Dynamics

We evaluate CylinderFlow, DeformingPlate, and Airfoil from the MeshGraphNets benchmark (Pfaff et al., 2021). Figure 7 summarizes the three tasks and their rollout setting.

Common protocol. We use the official train/validation/test trajectory splits: $1{,}000/100/100$ for CylinderFlow and Airfoil, and $1{,}200/100/100$ for DeformingPlate. Mesh edges are represented in both directions, with additional world-space edges within radius 0.03 for DeformingPlate. Scalar features are standardized componentwise and vector features by an isotropic RMS using training-set statistics. Models are trained on randomly sampled one-step transitions with normalized MSE and task-specific input noise. Checkpoints are selected by one-step validation performance, while hyperparameter configurations are chosen by validation rollout RMSE. Final models are evaluated autoregressively on the held-out test trajectories, with prescribed boundary or kinematic nodes clamped to the data. We report physical-space RMSE after one step, 50 steps, and over the full trajectory, averaging trajectories equally. MeshGraphNet, EGNN, and ESNN use the same processed datasets, splits, normalization, prediction targets, evaluation supports, and rollout procedure, with results reported over five runs.

CylinderFlow. The scalar input is a five-way node-type encoding, with current velocity as the covariant vector input. The model predicts the velocity increment and next pressure. The ESNN lift combines velocity with the mean incident edge displacement. The primary rollout metric is velocity RMSE over normal and outflow nodes.

DeformingPlate. The scalar input is a three-way node-type encoding, with current world position as the coordinate state. Edge features encode edge type, rest length, and strain. The covariant lift combines the prescribed next displacement on kinematic nodes with the mean incident edge displacement. The model predicts displacement and $\log(1 + \text{von-Mises stress})$, with world-position RMSE over nonkinematic nodes as the primary metric.

Airfoil. The scalar input combines a nine-way node-type encoding with current density, while current velocity provides the covariant vector input. The model predicts velocity and density increments together with next pressure. The ESNN lift combines velocity with the mean incident edge

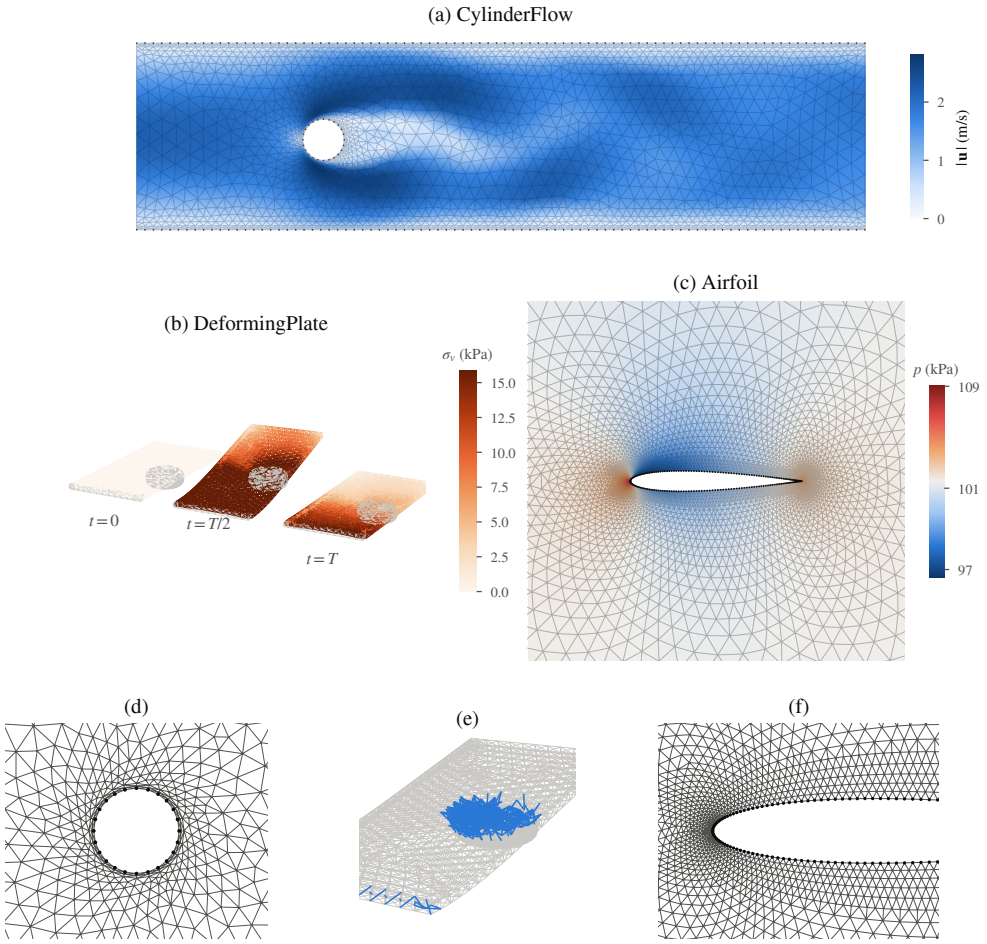

Figure 7: Mesh-based dynamics benchmarks: CylinderFlow, DeformingPlate, and Airfoil. Models are trained on one-step targets and evaluated autoregressively.

displacement, and the primary rollout metric is velocity RMSE over normal nodes. For Airfoil, this evaluation support differs from the near-wing region used for the headline result in the original MeshGraphNets study. We therefore compare ESNN directly with the in-repository MeshGraphNet and EGNN models evaluated on the same support.

E.5 QM9 Molecular-Property Prediction

Dataset and reporting convention. We use the QM9 data partition and reporting convention of GotenNet (Aykent & Xia, 2025). The PyTorch Geometric release contains 130,831 molecules after standard preprocessing. A fixed random 110,000/10,000/10,831 train/validation/test partition is generated with split seed 1. Molecular graphs are radius graphs with cutoff 5 Å, at most 32 neighbors per atom, and no self-loops.

The twelve independently trained targets are

$$\{\mu, \alpha, \epsilon_{\mathrm{HOMO}}, \epsilon_{\mathrm{LUMO}}, \Delta\epsilon, \langle R^2 \rangle, \mathrm{ZPVE}, U_0, U, H, G, C_v\}\,. \tag{142}$$

Validation and test performance are reported as MAE in the same units: mD for μ, $\mathrm{m}a_0^3$ for α, meV for orbital and energy targets, $\mathrm{m}a_0^2$ for $\langle R^2 \rangle$, $\mathrm{mcal\,mol^{-1}\,K^{-1}}$ for C_v, and meV for ZPVE. Literature results obtained with a different data partition are marked explicitly in Table 4.

Table 4: *QM9 molecular-property prediction.* Mean absolute error (MAE) on twelve invariant molecular properties. Baseline values and partition annotations follow Aykent & Xia (2025); † marks a different data partition. ESNN results are reported for the validation-selected model for each property, with the best ESNN result shown in **bold**. Units are given in the second header row. Lower is better.

Method	α	$\Delta\epsilon$	ϵ_{HOMO}	ϵ_{LUMO}	μ	C_v	G	H	$\langle R^2 \rangle$	U	U_0	**ZPVE**
Units	$\mathrm{m}a_0^3$	meV	meV	meV	mD	$\mathrm{mcal\,mol^{-1}\,K^{-1}}$	meV	meV	$\mathrm{m}a_0^2$	meV	meV	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	10.67	10.54	1.54
ESNN-Diag	62.16	43.69	**20.29**	21.43	17.15	**30.57**	**11.48**	**12.09**	110.88	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	**55.90**	**39.01**	25.16	**19.17**	**13.99**	31.57	14.77	13.86	**108.16**	**10.63**	**10.34**	**1.45**

Inputs, lifting, and readout. PyTorch Geometric atom descriptors provide the invariant scalar features $\mathbf{s}_i$, while atomic numbers are retained for atom-reference corrections. Geometry is encoded through the radius graph and relative displacements, without explicit bond attributes. The covariant lift initializes vector features from centered atomic positions with learned scales. The readout combines invariant vector norms with scalar features, followed by permutation-invariant pooling and a scalar prediction head.

Training and model selection. A separate model is trained for each property using MSE on standardized targets. For U_0, U, H, and G, atom-reference contributions are removed before normalization and restored at evaluation. Validation MAE selects the checkpoint for each property and transport, which is then evaluated once on the held-out test set. Table 4 reports the resulting test MAE.

Across the ESNN variants, Radial–Tangential Transport gives the lowest MAE on eight of the twelve properties, while Diagonal Transport is best on the remaining four. Relative to EGNN, the best ESNN result is lower on ten of the twelve targets, with H and $\langle R^2 \rangle$ as the exceptions.

E.6 Implementation, Model Selection, and Evaluation

Training and model selection. Unless stated otherwise, ESNN and EGNN are trained with AdamW, cosine learning-rate decay, degree-normalized aggregation, and gradient clipping; MeshGraphNet uses Adam. Dataset splits and preprocessing statistics are fixed before training.

Hyperparameters are selected using validation data only. Charged and gravity N-body models are selected by validation MSE, ModelNet40 by validation accuracy, mesh models by validation rollout RMSE, and QM9 models by validation MAE. For the mesh tasks, checkpoints are selected using one-step validation performance before configurations are compared by rollout RMSE.

For ModelNet40, $z/\mathrm{SO}(3)$ is not a separate training or selection regime. The same validation-selected z-trained checkpoint is evaluated under both z and $\mathrm{SO}(3)$ test rotations.

Most ModelNet40 ESNN experiments use $C = 128$, $L = 4$, $V = 2$, and learning rate 10^{-3}. The selected z-trained Diagonal model uses $C = 64$, $L = 3$, $V = 2$, and learning rate 1.06×10^{-3}.

Table 5: Selected N-body, gravity, and ModelNet40 configurations. C, L, and V denote scalar width, message-passing layers, and vector groups. For ModelNet40, z/z and $z/\mathrm{SO}(3)$ use the same z-trained checkpoint.

Task	Model	Transport / training regime	C	L	V	Learning rate	Epochs
Charged N-body	ESNN	Identity	64	6	0	3.38×10^{-3}	10000
	ESNN	Diagonal	32	5	1	3.25×10^{-3}	3000
	ESNN	Orthogonal	32	7	1	1.92×10^{-3}	3000
	ESNN	Radial–Tangential	32	5	1	3.41×10^{-3}	3000
Gravity N-body	ESNN	Diagonal	64	7	1	1.01×10^{-3}	3000
	ESNN	Orthogonal	32	7	1	1.24×10^{-3}	3000
	ESNN	Radial–Tangential	32	5	1	3.53×10^{-3}	3000
ModelNet40	EGNN	z-trained	304	4	–	5×10^{-4}	800
	EGNN	$\mathrm{SO}(3)$-trained	224	4	–	10^{-3}	800
	ESNN	Identity, z-trained	128	4	2	10^{-3}	800
	ESNN	Identity, $\mathrm{SO}(3)$-trained	128	4	2	10^{-3}	800
	ESNN	Diagonal, z-trained	64	3	2	1.06×10^{-3}	300
	ESNN	Diagonal, $\mathrm{SO}(3)$-trained	128	4	2	10^{-3}	800
	ESNN	Orthogonal, z-trained	128	4	2	10^{-3}	800
	ESNN	Orthogonal, $\mathrm{SO}(3)$-trained	128	4	2	10^{-3}	800
	ESNN	Radial–Tangential, z-trained	128	4	2	10^{-3}	800
	ESNN	Radial–Tangential, $\mathrm{SO}(3)$-trained	128	4	2	10^{-3}	800

Table 6: Selected mesh configurations. All models use the same processed datasets and rollout evaluation. For MeshGraphNet, width denotes latent width and depth the number of message-passing steps; for ESNN, V denotes vector groups.

Task	Model	Transport	Width	Depth	V	Batch	Learning rate	Epochs
Airfoil	MeshGraphNet	–	192	10	–	2	2×10^{-4}	45
	EGNN	–	246	6	–	2	2.5×10^{-4}	45
	ESNN	Identity	128	7	2	2	3.01×10^{-4}	45
	ESNN	Diagonal	128	5	3	2	2.48×10^{-4}	45
	ESNN	Orthogonal	128	7	2	2	5.86×10^{-4}	45
	ESNN	Radial–Tangential	64	8	3	2	3.93×10^{-4}	45
CylinderFlow	MeshGraphNet	–	192	15	–	8	5×10^{-5}	52
	EGNN	–	244	8	–	8	2.5×10^{-4}	52
	ESNN	Identity	128	7	2	8	3.91×10^{-4}	52
	ESNN	Diagonal	128	9	1	8	1.63×10^{-4}	52
	ESNN	Orthogonal	128	8	2	8	6.29×10^{-4}	52
	ESNN	Radial–Tangential	64	9	2	8	3.33×10^{-4}	52
DeformingPlate	MeshGraphNet	–	192	10	–	4	5×10^{-5}	16
	EGNN	–	318	8	–	4	3×10^{-4}	16
	ESNN	Identity	128	8	2	4	3×10^{-4}	16
	ESNN	Diagonal	128	10	1	4	3.69×10^{-4}	16
	ESNN	Orthogonal	128	8	3	2	2.98×10^{-4}	16
	ESNN	Radial–Tangential	128	8	2	4	3×10^{-4}	16

Repeated-run evaluation. Charged N-body, gravity N-body, and the mesh benchmarks are evaluated over five runs using fixed train/validation/test partitions. Reported uncertainties are sample standard deviations across these runs.

ModelNet40 is reported from one run per selected configuration. For the z-trained models, the z/z and $z/\mathrm{SO}(3)$ evaluations use identical trained weights. QM9 uses a separately validation-selected model for each property and transport, evaluated once on the held-out test set. Literature baselines follow the reporting conventions of their original sources.