# SE3Set: Harnessing Equivariant Hypergraph Neural Networks for Molecular Representation Learning

**Hongfei Wu, Lijun Wu, Guoqing Liu, Zhirong Liu, Bin Shao, Zun Wang**

Peking University · Microsoft Research AI4Science

TMLR 2025

J2C Certification

ICML International Conference On Machine Learning

Paper Link

Code Link

## 1 Why Go Beyond Pairwise Graphs?

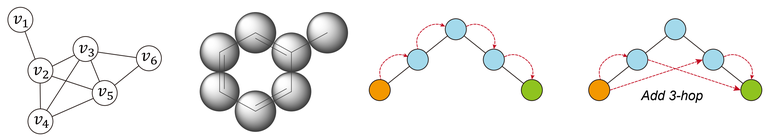
- **Many-body interactions** in large molecules.
- Pairwise GNNs message passing relies **only on atom–atom edges**.
- Hypergraphs let one hyperedge represent a **multi-atom substructure**.

HGNNs *vs* GNNs

**Explicit Many-body Interaction Modeling**

**Higher Message Passing Efficiency**

**Graph** Multi-hop message passing




**Hypergraph** Fragment-level message passing

## 2 Main Contributions

### First SE(3)-equivariant HGNN

We extend equivariant molecular representation learning from pairwise graphs to hypergraphs through V2E and E2V updates.

### Chemistry-guided Hypergraph Construction

We build molecular hypergraphs from fragments that combine 2D topology, functional groups, rings, and 3D spatial context.

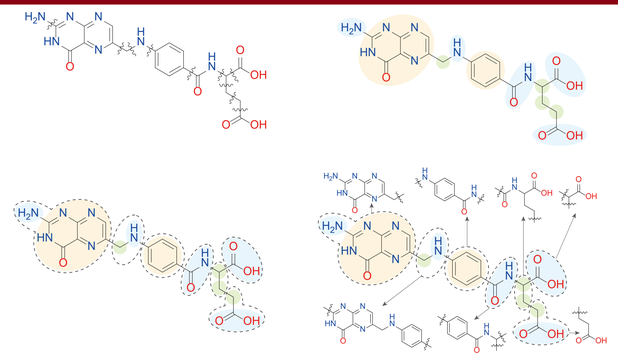
### Many-body Molecular Representation

Hyperedges naturally encode multi-atom substructures, enabling fragment-level modeling beyond atom–atom message passing.

### Large-molecule Advantage

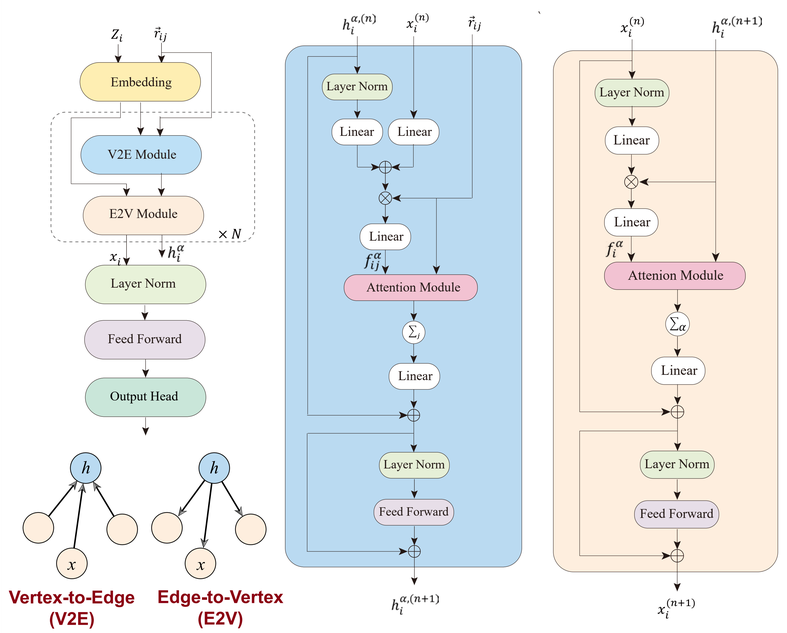
Competitive results on QM9/MD17 and stronger gains on MD22/OE62, where many-body interactions are more important.

## 3 Fragmentation to Hypergraph



(1) preserve rings and functional groups
(2) selectively break single bonds
(3) merge fragments with 3D context
(4) allow overlap to retain local chemical effects

## 4 SE3Set Architecture




***V2E***

$$\Delta h_i^\alpha = \phi_{\text{aggr}}^{\text{V2E}}\left(\sum a_{ij}^\alpha \phi_{\text{msg}}^{\text{V2E}}\left(x_i, x_j\right)\right)$$

***E2V***

$$\Delta x_i = \phi_{\text{aggr}}^{\text{E2V}}\left(\sum a_{ij}^\alpha \phi_{\text{msg}}^{\text{E2V}}\left(h_i^\alpha, h_j^\alpha\right)\right)$$

## 5 Results

**Competitive on Small Molecules** QM9/MD17

**~ 20% Error Reduction** SOTA on MD22

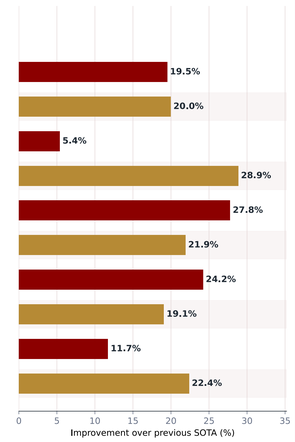
**Large Molecule Advantage** Across OE62 Mol Size

### MD22: SE3Set vs. Previous SOTA

Mean absolute error on five MD22 molecules; lower MAE is better.

| Molecule / Target | Prev. SOTA | SE3Set | Improve |
|---|---|---|---|
| **Ac-Ala3-NHMe** Energy \| 5500/500 \| kcal/mol | 0.0620 MACE | **0.0499** | +19.5% |
| **Ac-Ala3-NHMe** Force \| 5500/500 \| kcal/(mol·Å) | 0.0681 QuinNet | **0.0545** | +20.0% |
| **DHA** Energy \| 7500/500 \| kcal/mol | 0.0873 ViSNet-LSRM | **0.0826** | +5.4% |
| **DHA** Force \| 7500/500 \| kcal/(mol·Å) | 0.0506 Equiformer | **0.0360** | +28.9% |
| **Stachyose** Energy \| 7500/500 \| kcal/mol | 0.1055 ViSNet-LSRM | **0.0762** | +27.8% |
| **Stachyose** Force \| 7500/500 \| kcal/(mol·Å) | 0.0543 QuinNet | **0.0424** | +21.9% |
| **AT-AT** Energy \| 2500/500 \| kcal/mol | 0.0772 ViSNet-LSRM | **0.0585** | +24.2% |
| **AT-AT** Force \| 2500/500 \| kcal/(mol·Å) | 0.0687 QuinNet | **0.0556** | +19.1% |
| **AT-AT-CG-CG** Energy \| 1500/500 \| kcal/mol | 0.1135 ViSNet-LSRM | **0.1002** | +11.7% |
| **AT-AT-CG-CG** Force \| 1500/500 \| kcal/(mol·Å) | 0.1063 ViSNet-LSRM | **0.0825** | +22.4% |

**Average improvement across 10 MD22 tasks: 20.1%**

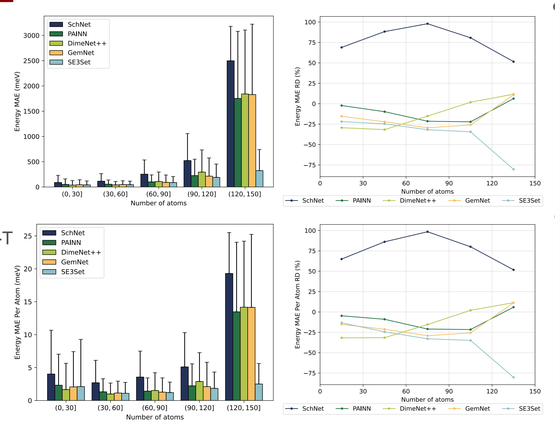



### OE62: Large-Molecule Advantage

SE3Set shows pronounced gains on **large molecules.**

Best overall MAE: 51.7 meV
SE3Set outperforms SchNet, PaiNN, DimeNet++, GemNet-T

**SE3Set reduces the additional error growth caused by complex many-body interactions.**




## Takeaway Message

- SE3Set introduces an **SE(3)-equivariant molecular hypergraph neural network** that brings chemical fragments into geometric representation learning.
- By modeling **fragment-level many-body interactions**, SE3Set achieves stronger performance on larger molecular systems where pairwise graph models become less sufficient.

*Chemically meaningful hypergraphs for SE(3)-equivariant many-body molecular representation learning.*