A Plug-and-Play Quaternion Message-Passing Module for Molecular Conformation Representation

Angxiao Yue[1], Dixin Luo[2], Hongteng Xu[1,3*]

[1]Gaoling School of Artifical Intelligence, Renmin University of China, Beijing, China
[2]School of Computer Science and Technology, Beijing Institute of Technology, Beijing, China
[3] Beijing Key Laboratory of Big Data Management and Analysis Methods, Beijing, China
angxiaoyue@ruc.edu.cn, dixin.luo@bit.edu.cn, hongtengxu@ruc.edu.cn

Abstract

Graph neural networks have been widely used to represent 3D molecules, which capture molecular attributes and geometric information through various message-passing mechanisms. This study proposes a novel quaternion message-passing (QMP) module that can be plugged into many existing 3D molecular representation models and enhance their power for distinguishing molecular conformations. In particular, our QMP module represents the 3D rotations between one chemical bond and its neighbor bonds as a quaternion sequence. Then, it aggregates the rotations by the chained Hamilton product of the quaternions. The real part of the output quaternion is invariant to the global 3D rotations of molecules but sensitive to the local torsions caused by twisting bonds, providing discriminative information for training molecular conformation representation models. In theory, we prove that considering these features enables invariant GNNs to distinguish the conformations caused by bond torsions. We encapsulate the QMP module with acceleration, so combining existing models with the QMP requires merely one-line code and little computational cost. Experiments on various molecular datasets show that plugging our QMP module into existing invariant GNNs leads to consistent and significant improvements in molecular conformation representation and downstream tasks.

Introduction

With the advent of 3D graph neural networks (GNNs) (Duvenaud et al. 2015; Fout et al. 2017; Xie and Grossman 2018; Gao and Ji 2019; Liu et al. 2021b; Li et al. 2022a), 3D molecular representation learning has shown great promise in molecular dynamics tasks, e.g., predicting the energy (Li et al. 2022b; Hu et al. 2021), atomic force (Zitnick et al. 2022; Gasteiger, Yeshwanth, and Günnemann 2021), and other quantum mechanical properties (Cai et al. 2022; Stärk et al. 2022; Li, Zhao, and Zeng 2022) of molecules. Message-passing is the most widely used mechanism to realize such GNNs, whose input features and model architectures significantly impact the quality and efficiency of 3D molecular representation. When implementing message-passing, existing equivariant GNNs (Thomas et al. 2018; Satorras, Hoogeboom, and Welling 2021; Fuchs et al. 2020; Finzi et al. 2020; Schütt, Unke, and Gastegger 2021; Köhler, Klein, and Noé 2020) often treat molecules as point clouds and embed equivariant information (i.e., atoms' coordinates (Satorras, Hoogeboom, and Welling 2021) or orientations (Thomas et al. 2018)) and relative position information (i.e., pairwise distances among atoms) jointly. Invariant GNNs (Schütt et al. 2017; Gasteiger, Groß, and Günnemann 2020; Gasteiger et al. 2020; Liu et al. 2022; Wang et al. 2022) often model molecules as graphs and consider their SE(3)/E(3)[1]-invariant information, e.g., bond lengths, rotations, and torsions caused by twisting bonds.

Although these GNNs have been applied in many tasks, they suffer from some computational efficiency issues and have limited performance on molecular conformation analysis. On the one hand, the equivariant GNNs have high computational complexity due to their high-order message-passing (Thomas et al. 2018; Fuchs et al. 2020; Finzi et al. 2020; Köhler, Klein, and Noé 2020) within atoms' point clouds, and the lack of bonds' rotation and torsion angles make them unable to distinguish molecular conformations. On the other hand, an ideal SE(3)-invariant GNN represents 3D structures of molecules by recording the dihedrals defined on each bond's 2-hop neighborhood, whose computational complexity can be as high as $\mathcal{O}(ND^3)$ for a 3D molecule with N atoms and D degrees on average. The simplified invariant GNNs (Gasteiger, Groß, and Günnemann 2020; Gasteiger et al. 2020; Liu et al. 2022) can reduce the complexity to $\mathcal{O}(ND^2)$ but sacrifice the power for distinguishing molecular conformations.

This study proposes a plug-and-play quaternion message-passing (QMP) module to improve invariant GNNs in molecular conformation representation and analysis tasks. As illustrated in Figure 1, our QMP achieves a new mechanism to encode the 3D rotation information of chemical bonds. For each bond of a 3D molecule, we consider the 3D rotations from its 1-hop neighboring bonds to itself. The QMP represents each 3D rotation as a quaternion (Dam, Koch, and Lillholm 1998) and achieves a novel message-passing mechanism, i.e., sorting the quaternions in descending order of the corresponding bond rotation angles and

*Corresponding author.

[1]SE(3) is shorthand for "the special Euclidean group of rigid body displacements in 3D". E(3) $=$ SE(3) $\otimes$ $\{\mathrm{E}, \mathrm{I}\}$, is the Euclidean group in 3D space which contains translations, rotations, and reflection.

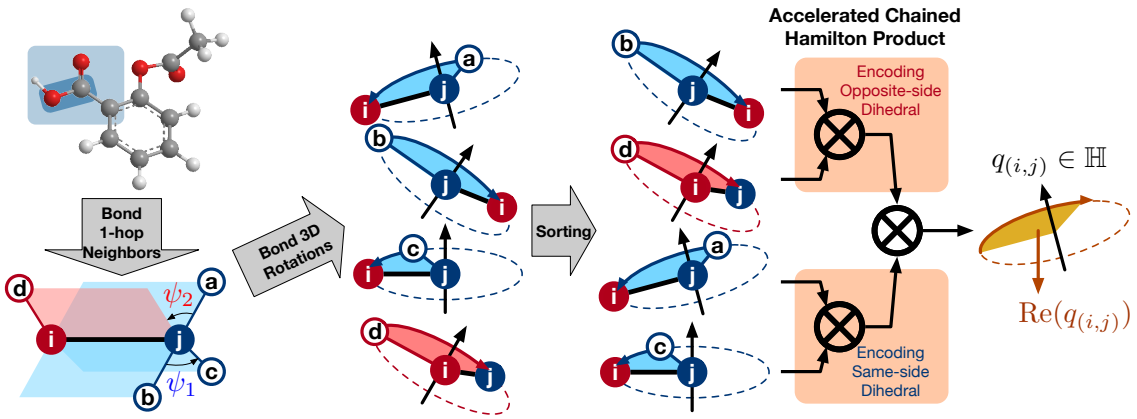


Figure 1: An illustration of our QMP module. Given a molecule, for each bond (i, j), we first obtain the 3D rotations from its 1-hop neighboring bonds to itself and sort them according to the rotation angles. Representing the sorted 3D rotations by a sequence of quaternions, we apply accelerated chained Hamilton product to obtain a quaternion $q_{(i,j)}$, whose real-part $\mathrm{Re}(q_{(i,j)})$ encodes all rotation angles, same-side dihedrals (e.g., the ψ_1 in blue), and opposite-side dihedrals (e.g., the ψ_2 in red) jointly. As shown in the orange blocks, the multiplications of quaternion pairs encode dihedrals.

merging the quaternion sequence to a single quaternion by a chained Hamilton product. The real part of the output quaternion is invariant to the global 3D rotation of the molecule and, moreover, sensitive to the torsions caused by twisting bonds because it encodes opposite-side dihedrals implicitly. Therefore, **taking the output of QMP as the input of invariant GNNs helps to enhance their discriminative power on representing molecular conformations while keeping their global SE(3)-invariance.**

We accelerate the chained Hamilton product using a multi-thread reduction strategy and encapsulate the QMP as one-line code. Therefore, we can plug the QMP into existing invariant GNNs with little computational cost. In theory, we demonstrate that the computational complexity of our QMP is $\mathcal{O}(ND^2)$, and with the help of the accelerated chained Hamilton product, its runtime can be further reduced to $\mathcal{O}(ND \log D)$. We plug the QMP into three representative E(3)-invariant GNNs (i.e., SchNet (Schütt et al. 2017), DimeNet (Gasteiger, Groß, and Günnemann 2020), and DimeNet++ (Gasteiger et al. 2020)) and one typical SE(3)-invariant GNN (i.e., SphereNet (Liu et al. 2022)), and analyze its impacts on model performance quantitatively. Experimental results show that our QMP leads to consistent and significant improvements of these invariant GNNs in molecular dynamics prediction tasks.

Proposed Quaternion Message-Passing

Motivations and Preliminaries

In this study, we denote a 3D molecule with N atoms and M chemical bonds as a graph $G(\mathcal{V}, \mathcal{E}, \boldsymbol{H}^{(0)}, \boldsymbol{Z}^{(0)}, \boldsymbol{X})$. Here, $\mathcal{V}$ and $\mathcal{E}$ denote the set of atoms and that of bonds, respectively. $\boldsymbol{H}^{(0)} \in \mathbb{R}^{N \times D_v}$ represents the atom feature matrix, and the i-th row of $\boldsymbol{H}^{(0)}$ denotes the initial feature of the i-th atom, denoted as $\boldsymbol{h}_i \in \mathbb{R}^{D_v}$. $\boldsymbol{Z}^{(0)} \in \mathbb{R}^{M \times D_e}$ represents the bond feature matrix. We denote the row corresponding to the bond (i, j) as $\boldsymbol{z}_{ij} \in \mathbb{R}^{D_e}$. The matrix $\boldsymbol{X} \in \mathbb{R}^{N \times 3}$ contains the 3D coordinates of atoms, whose rows are denoted as $\boldsymbol{x}_i \in \mathbb{R}^3$ for $i = 1, ..., N$. Based on the coordinate matrix $\boldsymbol{X}$, we can extract the following to SE(3)-invariant information:

1. **Bond Length Matrix:** We denote the bond length matrix of G as $\boldsymbol{D} = [d_{ij}] \in \mathbb{R}_+^{N \times N}$, where $d_{ij} = \|\boldsymbol{x}_i - \boldsymbol{x}_j\|_2$ if $(i, j) \in \mathcal{E}$, otherwise, $d_{ij} = 0$.
2. **Bond Rotation Set:** Given two edges connected at atom j, i.e., $(j, i), (j, k) \in \mathcal{E}$, the bond rotation from (j, k) to (j, i) is described as follows:

$$\boldsymbol{u}_{kji} = \frac{\boldsymbol{p}_{jk} \times \boldsymbol{p}_{ji}}{\|\boldsymbol{p}_{jk} \times \boldsymbol{p}_{ji}\|}, \ \theta_{kji} = \arccos \frac{\langle \boldsymbol{p}_{jk}, \boldsymbol{p}_{ji} \rangle}{\|\boldsymbol{p}_{jk}\| \|\boldsymbol{p}_{ji}\|}, \quad (1)$$

   where $\boldsymbol{p}_{ji} = \boldsymbol{x}_i - \boldsymbol{x}_j$, $\boldsymbol{u}_{kji}$ is the rotation axis, and θ_{kji} is the bond rotation angle. $\langle \cdot, \cdot \rangle$ and $\times$ represent the inner and cross product, respectively. The rotation angle is SE(3)-invariant, and the collection of all rotation angles is denoted as $\boldsymbol{\Theta}$.
3. **Bond Dihedral Set:** As shown in Figure 1, given one bond and its two arbitrary 1-hop neighboring bonds, we can construct a *same-side dihedral* or an *opposite-side dihedral*. The same-side dihedral is obtained when the neighboring bonds connect to the same atom of the central bond. The opposite-side dihedral is obtained when the neighboring bonds connect to the two different atoms of the central bond. We denote the sets of these two kinds of dihedrals as $\boldsymbol{\Psi}_s$ and $\boldsymbol{\Psi}_o$, respectively.

Invariant GNNs (Gasteiger, Groß, and Günnemann 2020; Gasteiger et al. 2020; Liu et al. 2022; Wang et al. 2022) take the above information as input and represent molecules accordingly. The representations are invariant to the global 3D rotations of the molecules. However, as shown in Figure 2, a molecular conformation can be generated by twisting chemical bonds, and the changes of opposite-side dihedrals capture the geometrical discrepancy between different conformations. Existing invariant models often ignore the informa-

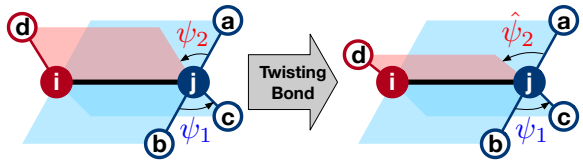


Figure 2: An illustration of the torsion caused by twisting bond. When twisting (i, j) with a torsion δ, the same-side dihedral ψ_1 is unchanged, while the opposite-side dihedral ψ_2 changes to $\hat{\psi}_2 = \psi_2 + \delta$.

tion of the opposite-side dihedrals, leading to sub-optimal performance when representing molecular conformations.

Motivated by this observation, we require the representations of conformations to be invariant to global 3D rotations and sensitive to the local twisting of chemical bonds. To achieve this aim, we would like to develop a new message-passing module to encode sufficient information (e.g., dihedrals) for representing molecular conformations and ensure that the module has high efficiency, high compatibility, and plug-and-play property. In the following content, we will design such a message-passing module with the help of the Quaternion Algebra (Dam, Koch, and Lillholm 1998).

Implementation of QMP

Quaternion (Dam, Koch, and Lillholm 1998; Zhu et al. 2018) is a kind of hypercomplex number having one real part and three imaginary parts, denoted as $q = s + x\mathtt{i} + y\mathtt{j} + z\mathtt{k} \in \mathbb{H}$. Here, $\mathbb{H}$ is the quaternion domain, $s, x, y, z \in \mathbb{R}$, and the imaginary units $\{\mathtt{i}, \mathtt{j}, \mathtt{k}\}$ obey the following rule: $\mathtt{i}^2 = \mathtt{j}^2 = \mathtt{k}^2 = \mathtt{ijk} = -1$. For convenience, we can represent a quaternion as a vector, i.e., $q = [s, \boldsymbol{u}^\top]^\top$ where $\boldsymbol{u}^\top = [x, y, z]^\top$. The multiplication of two quaternions is achieved by **Hamilton product** (Zhao et al. 2020): for $q_1 = [s_1, \boldsymbol{u}_1^\top]^\top, q_2 = [s_2, \boldsymbol{u}_2^\top]^\top \in \mathbb{H}$, we have

$$q_1 \otimes q_2 = [s_1 s_2 - \langle \boldsymbol{u}_1, \boldsymbol{u}_2 \rangle, \boldsymbol{u}_1 \times \boldsymbol{u}_2 + s_1 \boldsymbol{u}_2 + s_2 \boldsymbol{u}_1]. \quad (2)$$

Note that the Hamilton product is not commutative, i.e., $q_1 \otimes q_2 \neq q_2 \otimes q_1$ in general.

Quaternion is a powerful mathematical tool to describe 3D rotations. In particular, suppose that we rotate as a point $\boldsymbol{v}_1 \in \mathbb{R}^3$ with an angle θ around a unit rotation axis $\boldsymbol{u}$ and obtain a point $\boldsymbol{v}_2 \in \mathbb{R}^3$. When representing $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ as two pure quaternions (whose real parts are zeros), we can achieve the 3D rotation operation by

$$[0, \boldsymbol{v}_2^\top]^\top = q \otimes [0, \boldsymbol{v}_1^\top]^\top \otimes q^*, \quad (3)$$

where $q = [\cos \frac{\theta}{2}, \sin \frac{\theta}{2} \boldsymbol{u}^\top]^\top$ and $q^* = [\cos \frac{\theta}{2}, -\sin \frac{\theta}{2} \boldsymbol{u}^\top]^\top$ represents the conjugation of q.

Our QMP module leverages the above algebraic properties of quaternion to encode the 3D rotations associated with each bond. As illustrated in Figure 1, our QMP consists of the following key steps:

1. **Quaternion-based Representation of 3D Rotations** Given a molecule G with an atom set $\mathcal{V}$ and an edge set $\mathcal{E}$. For each atom i, we denote its 1-hop neighbors as a set $\mathcal{N}_i := \{j \mid (i, j) \in \mathcal{E}\}$. Accordingly, for each bond $(i, j) \in \mathcal{E}$, we consider its 1-hop neighbors connecting to i and j, denoted as $\mathcal{N}_{j \to i} := \{(i, k) \in \mathcal{E} \mid k \in \mathcal{N}_i \setminus \{j\}\}$ and $\mathcal{N}_{i \to j} := \{(k, j) \in \mathcal{E} \mid k \in \mathcal{N}_j \setminus \{i\}\}$, respectively. We represent the 3D rotation from each bond in the 1-hop neighboring sets to the bond (i, j) as a unit quaternion. Take $(k, j) \in \mathcal{N}_{i \to j}$ as an example. The rotation from (k, j) to (i, j) is represented as
$$q_{kji} = \left[\cos \frac{\theta_{kji}}{2}, \sin \frac{\theta_{kji}}{2} \boldsymbol{u}_{kji}^\top\right]^\top, \quad (4)$$
where θ_{kji} and $\boldsymbol{u}_{kji}$ are derived based on (1). As a result, we obtain two sets, i.e., $\mathcal{Q}_{i \to j} = \{q_{kji} | (k, j) \in \mathcal{N}_{i \to j}\}$ and $\mathcal{Q}_{j \to i} = \{q_{kij} | (i, k) \in \mathcal{N}_{j \to i}\}$, respectively.
2. **Angle-based Quaternion Sorting and Merging** Given the two sets of quaternions, our QMP encodes the corresponding rotation information by a chained Hamilton product. Because of the non-commutativity of the Hamilton product, we first need to convert the two quaternion sets to a quaternion sequence. In this study, we implement this conversion by an angle-based sorting and selection method. In particular, we select the quaternions with top-K rotation angles for each set. Then, we sort the selected $2K$ quaternions in descending order of their rotation angles, leading to a quaternion sequence, denoted as $\mathcal{Q}_{(i,j)} = \{q_k^{(i,j)}\} \in \mathbb{H}^{2K}$. Therefore, we can merge the quaternion sequence by a chained Hamilton product:
$$q_{(i,j)} = q_1^{(i,j)} \otimes q_2^{(i,j)} \otimes \ldots \otimes q_{2K}^{(i,j)} = \bigotimes\nolimits_{q \in \mathcal{Q}_{(i,j)}} q. \quad (5)$$
In practice, we set $K \leq D_{\text{max}}$ in general, where D_{max} is the maximum degree of atoms. For the bond whose $\mathcal{Q}_{i \to j}$ or $\mathcal{Q}_{j \to i}$ has fewer than K quaternions, we pad the set with the dumy quaternion $[1, \mathbf{0}_3^\top]^\top$.
3. **Mixed Encoding of Rotations and Dihedrals** The real part of $q_{(i,j)}$ contains important geometric information. Given two unit quaternions $q_k = [\cos \frac{\theta_k}{2}, \sin \frac{\theta_k}{2} \boldsymbol{u}_k]$ and $q_{k'} = [\cos \frac{\theta_{k'}}{2}, \sin \frac{\theta_{k'}}{2} \boldsymbol{u}_{k'}]$, the real part of $q_k \otimes q_{k'}$ is $\cos \frac{\theta_k}{2} \cos \frac{\theta_{k'}}{2} - \sin \frac{\theta_k}{2} \sin \frac{\theta_{k'}}{2} \langle \boldsymbol{u}_k, \boldsymbol{u}_{k'} \rangle$ according to (2). The first term encodes the information of the bonds' rotation angles, and the second term records the dihedral angle defined by the two axes, i.e., $\cos \psi = \langle \boldsymbol{u}_k, \boldsymbol{u}_{k'} \rangle$. Therefore, our QMP outputs the real part $\text{Re}(q_{(i,j)})$ as the feature of the bond (i, j).

In summary, given a molecule G with M bonds, our QMP achieves a map from the bond set $\mathcal{E}$ to a M-dimensional vector, i.e., QMP : $\mathcal{E} \mapsto [-1, 1]^M$.

Theoretical Properties and Rationality Analysis

Global SE(3)-Invariance Our QMP is SE(3)-invariant. As shown in Proposition 1 in (Zhang et al. 2020), for arbitrary two quaternions $q_1 = [s_1, \boldsymbol{u}_1^\top]^\top$ and $q_2 = [s_2, \boldsymbol{u}_2^\top]^\top$, when applying a rotation R on the rotation axes, we have

$$\begin{aligned} &[s_1, R(\boldsymbol{u}_1)^\top]^\top \otimes [s_2, R(\boldsymbol{u}_2)^\top]^\top \\ =&[s_1 s_2 - \langle \boldsymbol{u}_1, \boldsymbol{u}_2 \rangle, R(\boldsymbol{u}_1 \times \boldsymbol{u}_2 + s_1 \boldsymbol{u}_2 + s_2 \boldsymbol{u}_1)]^\top, \end{aligned} \quad (6)$$

where the real part is invariant to the rotation operation, while the imaginary part is equivariant to the rotation operation. Therefore, for the $\text{Re}(q_{(i,j)})$ derived by our QMP,

the global 3D rotation of the molecule does not change its value. Additionally, because the quaternions in (5) are derived based on the relative positions of chemical bonds, the $\mathrm{Re}(q_{(i,j)})$ is invariant to the global translation of molecule.

Sensitivity to Local Torsions Besides preserving the global SE(3)-invariance, our QMP is sensitive to the torsions of chemical bonds and thus benefits molecular conformation representation. In particular, the torsion caused by twisting a bond leads to the change of opposite-side dihedrals, which is ignored by most existing SE(3)-invariant GNNs (e.g., the SphereNet (Liu et al. 2022) only encodes the same-side dihedrals). As a result, these models fail to distinguish the conformations caused by twisting bonds.

On the contrary, the output of our QMP is sensitive to bond torsions. Take Figure 2 as an example. The twisting of the bond (i, j) can be represented as rotating the atom "d" with a torsion angle δ along the axis defined by (i, j), which equivalently corresponds to imposing a rotation R of the axis $\boldsymbol{u}_{dij}$ associated with the quaternion q_{dij}, i.e., $R(\boldsymbol{u}_{dij})$. When computing the Hamilton product between q_{dij} and any other $q = [s, \boldsymbol{u}] \in \{q_{aji}, q_{bji}, q_{cji}\}$, the real part of the output is $[\cos\frac{\theta_{dij}}{2}s - \sin\frac{\theta_{dij}}{2}\langle \boldsymbol{u}_{dij}, \boldsymbol{u}\rangle]$. Here, each $\langle \boldsymbol{u}_{dij}, \boldsymbol{u}\rangle$ corresponds to an opposite-side dihedral. It is changed by the bond torsion because $\langle R(\boldsymbol{u}_{dij}), \boldsymbol{u}\rangle \neq \langle \boldsymbol{u}_{dij}, \boldsymbol{u}\rangle$ in general. In theory, we have the following proposition.

Proposition 1 *Given K quaternions $\{q_k = [s_k, \boldsymbol{u}_k^\top]^\top\}_{k=1}^K$, their chained Hamilton product involves all $K(K-1)/2$ inner products of paired imaginary parts, i.e., $\langle \boldsymbol{u}_k, \boldsymbol{u}_{k'}\rangle$ for all $k \neq k'$.*

Proposition 1 means that given a bond, our QMP encodes all associated (same-side and opposite-side) dihedrals, whose output changes according to the bond torsion.

Rationality of Angle-based Sorting As a key step of our QMP module, the angle-base sorting operation determines the order of quaternions in the chained Hamilton product. We apply the angles of bond rotations to determine the order because they are relatively stable for molecular conformations — the molecular conformation is mainly generated by twisting bonds, and the twisting of bonds changes opposite-side dihedrals while keeping the rotation angles unchanged.

Plugging QMP into Invariant GNNs

Based on the above analysis, combining our QMP with existing invariant GNNs can preserve the global SE(3)-invariance of the original models and, at the same time, enhance their sensitivity to local bond torsions. As shown in Figure 3, we propose a simple and generic framework to plug our QMP module into representative invariant GNNs, which has high efficiency and compatibility. In particular, given a graph $G(\mathcal{V}, \mathcal{E}, \boldsymbol{H}^{(0)}, \boldsymbol{Z}^{(0)}, \boldsymbol{X})$, we keep the architectures of the GNNs and only modify the input of their bond embedding module by the output of our QMP. Specifically, in the l-th message-passing layer, we update bond and atom embeddings as

$$\boldsymbol{z}_{ij}^{(l+1)} = f_e(\boldsymbol{e}_{ij}),\ \boldsymbol{h}_i^{(l+1)} = f_v\big(\boldsymbol{h}_i^{(l)}, \sum\nolimits_{j\in\mathcal{N}_i} \boldsymbol{z}_{ij}^{(l+1)}\big), \quad (7)$$

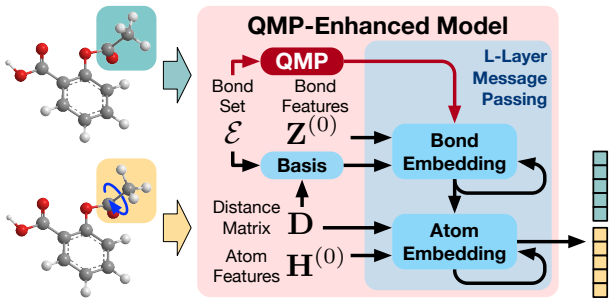


Figure 3: An illustration of how to plug our QMP into SE(3)-invariant GNNs. Given two molecular conformations, the QMP-enhanced model can capture bond torsions (i.e., the blue rotated arrow) and output different representations (i.e., the vectors with different colors).

where f_e and f_v are embedding functions for bonds and atoms, respectively, which are implemented by message-passing in general (Gilmer et al. 2017). e_{ij} is the intermediate embedding of the bond (i, j). For DimeNet, DimeNet++ and SphereNet,

$$\boldsymbol{e}_{ij} = \boldsymbol{z}_{ij}^{(l)} \oplus \sum\nolimits_{k\in\mathcal{N}_i\setminus j} f_a(\boldsymbol{z}_{ik}^{(l)}, b_{\boldsymbol{D},\boldsymbol{\Theta},\boldsymbol{\Psi}_s}) \| \mathrm{Re}(q_{(i,j)}), \quad (8)$$

where f_a is an interaction function that aggregates information from neighboring bonds. The operator "$\|$" means the concatenation of features, and "$\oplus$" denotes element-wise addition. b is the base function that transforms the geometries like the bond length matrix $\boldsymbol{D}$, the bond rotation set $\boldsymbol{\Theta}$, and optionally, the same-side dihedral set $\boldsymbol{\Psi}_s$ (Liu et al. 2022) into physically-meaningful representations. Typically, we can set b based on spherical Bessel and harmonics basis (Hu et al. 2021). For SchNet,

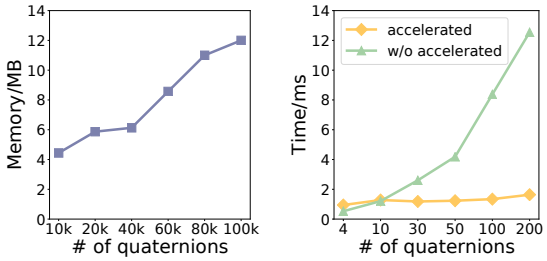
$$\boldsymbol{e}_{ij} = f_s(\boldsymbol{h}_j^{(l)}) \odot f_{cf}(b_{\boldsymbol{D}}) \parallel \mathrm{Re}(q_{(i,j)}), \quad (9)$$

where f_s is an embedding function, f_{cf} is a filter-generating function, and "$\odot$" represents element-wise multiplication. The main difference between (8) and (9) is whether the bond embeddings in $(l+1)$-th layer depend on l-th layer. According to (7)-(9), the only modification we made is concatenating the hidden vector with the corresponding output of our QMP, which merely requires one-line code.

Computational Efficiency

Accelerated Chained Hamilton Product As shown in Figure 3, our QMP is a plug-and-play module — given a molecule with M bonds, we only need to apply the QMP once to obtain a feature vector $[\mathrm{Re}(q_{(i,j)})] \in \mathbb{R}^M$. Without updating the output, introducing our QMP to existing models requires little computational cost. Figure 4(a) illustrates the memory cost of our QMP achieved under different numbers of quaternions. The memory cost is small (e.g., the chained Hamilton product of 100k quaternions only occupies 12MB memory), and its increasing rate is linear.

Moreover, we leverage the associativity of the Hamilton product, applying the multi-thread reduction strategy



(a) Memory Cost (b) Runtime Comparison

Figure 4: (a) The memory cost of chained Hamilton product. (b) The runtime comparison for the accelerated chained Hamilton product and the non-accelerated one.

in (Santos 2002; Qin et al. 2022) to accelerate the computation of the chained Hamilton product. As illustrated in Figure 1, this approach involves constructing a complete binary tree for the quaternion sequence to compute the chained Hamilton product recursively. Leveraging the tree structure, we can compute multiple Hamilton products in parallel and reduce the runtime significantly. Specifically, for the sequence with $2K$ quaternions, the runtime of our accelerated chained Hamilton product is $\mathcal{O}(\log K)$. As shown in Figure 4(b), the horizontal axis represents the number of quaternions, and the vertical axis represents the time consumed by performing chained product on these quaternions. As the number of quaternions increases, our non-accelerated method exhibited a rapidly increasing time consumption trend, whereas our accelerated version remained stable, consistently below 2ms.

Complexity Analysis In this study, we follow the work in (Liu et al. 2022), constructing radius graphs for molecular conformations and obtaining dense adjacency matrices. For a molecule having N atoms, and D node degrees on average, the complexity of typical invariant GNNs (e.g., DimeNet (Gasteiger, Groß, and Günnemann 2020), DimeNet++ (Gasteiger et al. 2020), and SphereNet (Liu et al. 2022)) is $\mathcal{O}(ND^2)$ when operating within each bond's 1-hop neighborhood and computing bond rotations. As aforementioned, these methods fail to distinguish molecular conformations because of lacking the information on opposite-side dihedrals. These models require each atom's 2-hop neighborhood information to capture the opposite-side dihedrals, resulting in a complexity of $\mathcal{O}(ND^3)$. **Our QMP suppresses this issue by encoding dihedrals implicitly, achieving a trade-off between performance and efficiency.** In particular, for each bond, our QMP records and merges $2K$ bond rotations from its 1-hop neighbors to itself. According to Proposition 1, although our QMP encodes all K^2 opposite-side and $K(K-1)$ same-side dihedrals, it only requires $2K-1$ quaternion multiplications. Therefore, the computational complexity of our QMP is $\mathcal{O}(MK)$ or equivalently, $\mathcal{O}(ND^2)$ (because the number of bonds $M = \mathcal{O}(ND)$ and $K = \mathcal{O}(D)$), which is comparable to that of invariant GNNs. The runtime in practice can be $\mathcal{O}(M \log K)$ (or $\mathcal{O}(ND \log D)$) with the help of the accelerated chained Hamilton product.

Connections to Related Work

As aforementioned, invariant GNNs favorably achieve the invariance merit since the relative information is contained in massage-passing. SchNet (Schütt et al. 2017) encodes bond length by continuous-filter convolutional layers. DimeNet (Gasteiger, Groß, and Günnemann 2020) and Dimenet++ (Gasteiger et al. 2020) utilize embeddings associated with the directions to neighboring atoms, encoding both bond length and angle by spherical harmonics. Nevertheless, these methods fail to provide a comprehensive representation as they overlook dihedral information. SphereNet (Liu et al. 2022) is developed based on DimeNet and moves forward to approximate completeness by considering dihedral information. However, due to the bottleneck of computation cost, the calculation only considers the same-side dihedrals. Recently, ComeENet (Wang et al. 2022) builds a complete 3D graph to represent a molecule and selects the nearest neighbor for each atom to calculate bond length, angle, and dihedral. However, ignoring the remaining neighboring atoms hurts the model's expressiveness. Our QMP works as a plug-and-play module, providing a simple, efficient, and generic solution to enhance various backbone models' discriminative power when representing molecular conformations.

Besides our QMP module, some quaternion-based neural networks have been built for modeling graphs (Zhu et al. 2018; Zhang et al. 2020) and point clouds (Shen et al. 2020), e.g., the QuaterNet in (Pavllo, Grangier, and Auli 2018), the quaternion convolution neural network in (Zhu et al. 2018), and the quaternion product unit (QPU) (Zhang et al. 2020; Qin et al. 2022). Among these models, the QPU model applies a similar technical route, aggregating 3D rotations by chained Hamilton product. However, QPU is designed for modeling the rotation information of 3D action skeletons, which requires different skeletons to be modeled in the same tree structure and merges the 3D rotations of joints in the same order, which is unsuitable for 3D molecular modeling. Moreover, different from QPU and other quaternion models, which have learnable parameters and thus require time-consuming training, our QMP is a parameter-free module for feature extraction. It is highly compatible with existing models, and plugging it into them does not significantly increase the computational cost.

Experiments

Effectiveness Tests

To demonstrate the effectiveness of our QMP, we conduct experiments on three datasets, including MD17 (Chmiela et al. 2017), MD17@CCSD (Chmiela et al. 2018), and OC20 (Chanussot et al. 2021b). We perform all experiments on MD17 and MD17@CCSD using the DIG[2] framework (Liu et al. 2021a). As for OC20, we use the Open Catalyst Project (OCP) framework[3] (Chanussot et al. 2021a)

[2]https://github.com/divelab/DIG

[3]https://github.com/Open-Catalyst-Project/ocp

Dataset	Molecule	sGDML	DimeNet	SchNet	DimeNet++	SphereNet	Q-SchNet	Q-DimeNet++	Q-SphereNet
MD17	Aspirin	0.68	0.499	1.339	<u>0.325</u>	0.400	1.289↑	**0.316**↑	0.356↑
	Benzene	0.20	0.187	0.346	<u>0.168</u>	0.193	0.316↑	**0.151**↑	0.177↑
	Ethanol	0.33	0.230	0.738	<u>0.150</u>	0.181	0.468↑	**0.148**↑	0.169↑
	Malonaldehyde	0.41	0.383	1.559	<u>0.263</u>	0.379	1.540↑	**0.241**↑	0.321↑
	Naphthalene	0.11	0.215	0.723	**0.100**	0.159	0.521↑	<u>0.106</u>	0.138↑
	Salicylic acid	0.28	0.374	1.001	**0.231**	0.261	0.971↑	<u>0.238</u>	0.321
	Toluene	0.14	0.216	0.747	<u>0.117</u>	0.142	0.550↑	**0.107**↑	0.136↑
	Uracil	0.24	0.301	1.351	<u>0.189</u>	0.228	1.010↑	**0.178**↑	0.241
MD17@CCSD	Aspirin	-	-	1.471	<u>0.387</u>	0.458	1.298↑	**0.365**↑	0.475
	Benzene	-	-	0.352	<u>0.049</u>	0.061	0.287↑	**0.048**↑	0.063
	Ethanol	-	-	1.321	<u>0.129</u>	0.158	0.786↑	**0.108**↑	0.148↑
	Malonaldehyde	-	-	1.531	**0.211**	0.289	1.086↑	<u>0.231</u>	0.264↑
	Toluene	-	-	0.855	0.133	0.160	0.637↑	<u>0.128</u>↑	**0.052**↑

Table 1: Results in terms of force MAE ($\frac{\text{kcal}}{\text{mol}\mathring{A}}$). The best results are in bold. The second best results are with underlines. ↑ indicates the performance is improved. The results of sGDML and DimeNet are quoted from (Liu et al. 2022).

to produce results. We apply state-of-the-art invariant models as our baselines, including SchNet (Schütt et al. 2017), CGCNN (Xie and Grossman 2018), sGDML (Chmiela et al. 2018), DimeNet (Gasteiger, Groß, and Günnemann 2020), DimeNet++ (Gasteiger et al. 2020), and SphereNet (Liu et al. 2022). For performance comparison, we plug QMP into SchNet, DimeNet, DimeNet++, and SphereNet, respectively, leading to four modified models: Q-SchNet, Q-DimeNet, Q-DimeNet++, and Q-SphereNet. We quantitatively analyze the impacts of QMP on the models in two tasks, i.e., molecular energy prediction (Li et al. 2022b; Hu et al. 2021) and atomic force prediction (Zitnick et al. 2022; Gasteiger, Yeshwanth, and Günnemann 2021), which are influenced by molecular conformations significantly. To ensure experimental fairness, we conduct the experiments using the same hyperparameter settings and hardware environment (NVIDIA GeForce RTX 3090). The code is available at https://github.com/AngxiaoYue/QMP.

MD17 The MD17 dataset consists of eight subsets, each containing a different molecule, requiring a separate model to be trained. We train each model using 1,000 conformations and 1,000 for validation, with the remaining used for testing. Our objective was to predict the atomic forces of the molecules using MAE as the evaluation metric. The joint loss function includes atomic forces and molecular energy, and we set the hyperparameter weight of force over energy (WoFE) to 100 in our experiment. The results in Table 1 demonstrate that QMP-plugged versions gain substantial performance over their backbone models. In particular, Q-SchNet surpasses SchNet by a large margin in performance on all molecules. Unlike SchNet, which only considers the bond length, Q-SchNet incorporates angle and dihedral information through QMP, resulting in significant improvements. Furthermore, Q-DimeNet++ and Q-SphereNet also exhibit superior performance on six molecules even if the original DimeNet++ and SphereNet have considered same-side dihedral information, which demonstrates the necessity of the opposite-side dihedral information.

MD17@CCSD Compared with MD17, MD17@CCSD provides more accurate target atomic forces with an expensive CCSD or CCSD(T) method. We adopt the train/validation/test split from (Gasteiger, Becker, and Günnemann 2021) where 950 samples are used for training, 50 for validation, and 500 for testing. The experimental settings are same with those of MD17. The results in Table 1 demonstrate Q-SchNet's superiority to SchNet on all molecules. Q-DimeNet++ and Q-SphereNet exhibit improved results on four and three molecules, respectively. The experimental results show that our QMP can enhance the model performance given more accurate target atomic forces.

OC20 The OC20 dataset, which has millions of density functional theory (DFT) relaxations, was released to facilitate catalyst modeling and discovery. On average, each molecule in this dataset is larger than those in MD17, containing 77.75 atoms. The dataset comprises three tasks: S2EF, IS2RS, and IS2RE. We focused on the IS2RE task, which involves predicting the adsorption energy of a relaxed structure given its initial structure as input. We employed a training set comprising 10,000 samples and four validation splits: In Domain (ID), Out of Domain Adsorbates (OOD Ads), Out of Domain Catalysts (OOD Cat), and Out of Domain Adsorbates and Catalysts (OOD Both), with the number of structures in each split being 24,943, 24,961, 24,963, and 24,987, respectively. In our evaluation, we assessed performance using the mean absolute error (MAE) of energy predictions and the percentage of energies within a threshold (EwT) of the ground truth energy. Following the baseline model paper, we compared the performance of different methods on the validation set. Results are shown in Table 2. Q-SchNet, Q-DimeNet, and Q-DimeNet++ consistently improve performance compared to their counterparts.

Ablation Study

By fusing angle and dihedral information, QMP enables backbone models to be sensitive to local torsion. We investigate the contributions of different QMP components to demonstrate the rationality of our design. For a bond (i, j),

Metrics		CGCNN	SchNet	DimeNet	DimeNet++	Q-SchNet	Q-DimeNet	Q-DimeNet++
Energy MAE	ID	0.9773	1.0480	0.9314	0.9261	1.0363↑	**0.9132**↑	0.9177↑
	OOD Ads	0.9818	1.0450	1.0720	0.9400	1.0640	0.9865↑	**0.9295** ↑
	OOD Cat	0.9269	1.0630	0.8945	0.8828	1.0370 ↑	**0.8753**↑	0.8816↑
	OOD Both	0.8828	1.0076	0.9643	0.8572	0.9646 ↑	0.9317↑	**0.8513**↑
	Average	0.9422	1.0409	0.9656	0.9015	1.0255 ↑	0.9267↑	**0.8950**↑
EwT (%)	ID	1.84	1.58	1.80	1.79	1.59↑	1.84↑	**1.91**↑
	OOD Ads	1.72	1.53	1.62	1.67	1.52	**1.85**↑	1.79↑
	OOD Cat	1.93	1.56	1.93	1.81	1.46	1.96↑	**2.13**↑
	OOD Both	1.69	1.48	1.57	**1.80**	1.61↑	1.70↑	1.75↑
	Average	1.80	1.54	1.73	1.77	1.55↑	1.84↑	**1.90**↑

Table 2: Results on OC20 IS2RE task in terms of energy MAE (eV) and the percentage of EwT of the ground truth energy. Performance is reported for models trained on the 10k training dataset.

QMP Components			Force MAE	
j-side	i-side	sorting	Aspirin	Benzene
✓	X	✓	0.323	0.152
X	✓	✓	0.360	0.166
✓	✓	X	0.390	0.173
✓	✓	✓	**0.316**	**0.151**

Table 3: Effects of ablating components of QMP.

Method	SE(3) Trans.	TFN	EGNN	Q-EGNN
MSE	2.440	1.550	0.724	**0.668**

Table 4: Mean Squared Error ($\times 10^{-2}$) for the future position estimation in the N-body system experiment.

our QMP aggregates its rotations from both sides. Therefore, we consider two ablating cases, i.e., aggregating the rotations from either the j-side or the i-side. We selected DimeNet++ as the backbone model and performed the ablation analysis on the Aspirin and Benzene in the MD17 dataset. The results are reported in Table 3. When we exclusively embedded rotations from the j-side of the bond, we find that the performance is similar to that of the backbone model. DimeNet++ has encoded bond length and angle from the j-side, and thus the j-side aggregated rotation cannot provide sufficient information to DimeNet++. If we only use rotations from the i-side, We observed that the model's performance is similar or slightly decreases compared to the original model. An intuitive explanation is, the rotations from the i-side while bond length and angle from the j-side, this inconsistency has disrupted the learning of the original model. Additionally, to demonstrate the necessity of the angle-based sorting, we consider applying the chained Hamilton product directly based on the default order of the quaternions. As shown in Table 3, without sorting, the model performance is significantly compromised.

Potentials to Equivariant GNNs

Although our QMP is originally designed for invariant GNNs, it has the potential for improving equivariant GNNs as well. In particular, besides predicting conformation properties, we consider the prediction problem of dynamical N-body systems, in which equivariant models are commonly used and have achieved encouraging performance. We apply 3,000 trajectories of five particles for training, 2,000 for validation, and 2,000 for testing. In the testing phase, we predict the positions of the five particles after 1,000 timesteps. The evaluation measure is the averaged mean-squared-error (MSE) between the estimated positions and the ground truth.

We test three representative equivariant models, i.e., SE(3) Transformer (Fuchs et al. 2020), TFN (Thomas et al. 2018), and EGNN (Satorras, Hoogeboom, and Welling 2021). Furthermore, we combine our QMP with EGNN, taking the real part of QMP as the input of EGNN. The results are shown in Table 4. We can find that EGNN achieves the best performance among the baselines in the N-body system prediction task. Moreover, by applying our QMP, we can further boost the performance of EGNN, leading to smaller errors. This experiment somewhat demonstrates the usefulness of QMP for equivariant models. A combination of QMP and more advanced equivariant models like Equiformer (Liao and Smidt 2022), LEFTNet (Du et al. 2023) can be our feature work. One potential way is using the imaginary part of QMP which is SE(3)-equivariant.

Conclusion

In this study, we have proposed an efficient and effective quaternion message-passing module for molecular conformation representation and analysis. With little computational cost, we can plug this module into most existing invariant GNNs by one-line code, achieving global SE(3)-invariance and enhancing their sensitivity to local bond torsions simultaneously. Experiments show that with the help of QMP module, the enhanced models perform better on distinguishing molecular conformations, leading to consistent improvements in downstream tasks. In the future, we would like to utilize the imaginary part of QMP to enhance equivariant GNNs. Additionally, we plan to design more hypercomplex neural networks for molecular modeling based on Clifford Algebra (Brandstetter et al. 2022; Ruhe et al. 2023).

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 62106271, No. 92270110), CAAI-Huawei MindSpore Open Fund, the Fundamental Research Funds for the Central Universities, and the Research Funds of Renmin University of China. Hongteng Xu also would like to thank the supports from the Beijing Key Laboratory of Big Data Management and Analysis Methods, the Intelligent Social Governance Platform, Major Innovation & Planning Interdisciplinary Platform for the "Double-First Class" Initiative.

References

Brandstetter, J.; Berg, R. v. d.; Welling, M.; and Gupta, J. K. 2022. Clifford neural layers for PDE modeling. *arXiv preprint arXiv:2209.04934*.

Cai, H.; Zhang, H.; Zhao, D.; Wu, J.; and Wang, L. 2022. FP-GNN: a versatile deep learning architecture for enhanced molecular property prediction. *Briefings in Bioinformatics*, 23(6): bbac408.

Chanussot, L.; Das, A.; Goyal, S.; Lavril, T.; Shuaibi, M.; Riviere, M.; Tran, K.; Heras-Domingo, J.; Ho, C.; Hu, W.; Palizhati, A.; Sriram, A.; Wood, B.; Yoon, J.; Parikh, D.; Zitnick, C. L.; and Ulissi, Z. 2021a. Open Catalyst 2020 (OC20) Dataset and Community Challenges. *ACS Catalysis*.

Chanussot, L.; Das, A.; Goyal, S.; Lavril, T.; Shuaibi, M.; Riviere, M.; Tran, K.; Heras-Domingo, J.; Ho, C.; Hu, W.; et al. 2021b. Open catalyst 2020 (OC20) dataset and community challenges. *Acs Catalysis*, 11(10): 6059–6072.

Chmiela, S.; Sauceda, H. E.; Müller, K.-R.; and Tkatchenko, A. 2018. Towards exact molecular dynamics simulations with machine-learned force fields. *Nature communications*, 9(1): 3887.

Chmiela, S.; Tkatchenko, A.; Sauceda, H. E.; Poltavsky, I.; Schütt, K. T.; and Müller, K.-R. 2017. Machine learning of accurate energy-conserving molecular force fields. *Science advances*, 3(5): e1603015.

Dam, E. B.; Koch, M.; and Lillholm, M. 1998. *Quaternions, interpolation and animation*, volume 2. Citeseer.

Du, W.; Du, Y.; Wang, L.; Feng, D.; Wang, G.; Ji, S.; Gomes, C.; and Ma, Z.-M. 2023. A new perspective on building efficient and expressive 3D equivariant graph neural networks. *arXiv preprint arXiv:2304.04757*.

Duvenaud, D. K.; Maclaurin, D.; Iparraguirre, J.; Bombarell, R.; Hirzel, T.; Aspuru-Guzik, A.; and Adams, R. P. 2015. Convolutional networks on graphs for learning molecular fingerprints. *Advances in neural information processing systems*, 28.

Finzi, M.; Stanton, S.; Izmailov, P.; and Wilson, A. G. 2020. Generalizing convolutional neural networks for equivariance to lie groups on arbitrary continuous data. In *International Conference on Machine Learning*, 3165–3176. PMLR.

Fout, A.; Byrd, J.; Shariat, B.; and Ben-Hur, A. 2017. Protein interface prediction using graph convolutional networks. *Advances in neural information processing systems*, 30.

Fuchs, F.; Worrall, D.; Fischer, V.; and Welling, M. 2020. Se (3)-transformers: 3d roto-translation equivariant attention networks. *Advances in Neural Information Processing Systems*, 33: 1970–1981.

Gao, H.; and Ji, S. 2019. Graph u-nets. In *international conference on machine learning*, 2083–2092. PMLR.

Gasteiger, J.; Becker, F.; and Günnemann, S. 2021. Gemnet: Universal directional graph neural networks for molecules. *Advances in Neural Information Processing Systems*, 34: 6790–6802.

Gasteiger, J.; Giri, S.; Margraf, J. T.; and Günnemann, S. 2020. Fast and uncertainty-aware directional message passing for non-equilibrium molecules. *arXiv preprint arXiv:2011.14115*.

Gasteiger, J.; Groß, J.; and Günnemann, S. 2020. Directional message passing for molecular graphs. *arXiv preprint arXiv:2003.03123*.

Gasteiger, J.; Yeshwanth, C.; and Günnemann, S. 2021. Directional message passing on molecular graphs via synthetic coordinates. *Advances in Neural Information Processing Systems*, 34: 15421–15433.

Gilmer, J.; Schoenholz, S. S.; Riley, P. F.; Vinyals, O.; and Dahl, G. E. 2017. Neural message passing for quantum chemistry. In *International conference on machine learning*, 1263–1272. PMLR.

Hu, W.; Shuaibi, M.; Das, A.; Goyal, S.; Sriram, A.; Leskovec, J.; Parikh, D.; and Zitnick, C. L. 2021. Forcenet: A graph neural network for large-scale quantum calculations. *arXiv preprint arXiv:2103.01436*.

Köhler, J.; Klein, L.; and Noé, F. 2020. Equivariant flows: exact likelihood generative learning for symmetric densities. In *International conference on machine learning*, 5361–5370. PMLR.

Li, H.; Zhao, D.; and Zeng, J. 2022. KPGT: Knowledge-Guided Pre-training of Graph Transformer for Molecular Property Prediction. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 857–867.

Li, S.; Zhou, J.; Xu, T.; Dou, D.; and Xiong, H. 2022a. Geomgcl: Geometric graph contrastive learning for molecular property prediction. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36(4), 4541–4549.

Li, Z.; Meidani, K.; Yadav, P.; and Barati Farimani, A. 2022b. Graph neural networks accelerated molecular dynamics. *The Journal of Chemical Physics*, 156(14): 144103.

Liao, Y.-L.; and Smidt, T. 2022. Equiformer: Equivariant graph attention transformer for 3d atomistic graphs. *arXiv preprint arXiv:2206.11990*.

Liu, M.; Luo, Y.; Wang, L.; Xie, Y.; Yuan, H.; Gui, S.; Yu, H.; Xu, Z.; Zhang, J.; Liu, Y.; Yan, K.; Liu, H.; Fu, C.; Oztekin, B. M.; Zhang, X.; and Ji, S. 2021a. DIG: A Turnkey Library for Diving into Graph Deep Learning Research. *Journal of Machine Learning Research*, 22(240): 1–9.

Liu, S.; Wang, H.; Liu, W.; Lasenby, J.; Guo, H.; and Tang, J. 2021b. Pre-training molecular graph representation with 3d geometry. *arXiv preprint arXiv:2110.07728*.

Liu, Y.; Wang, L.; Liu, M.; Lin, Y.; Zhang, X.; Oztekin, B.; and Ji, S. 2022. Spherical message passing for 3d molecular graphs. In *International Conference on Learning Representations (ICLR)*.

Pavllo, D.; Grangier, D.; and Auli, M. 2018. Quaternet: A quaternion-based recurrent model for human motion. *arXiv preprint arXiv:1805.06485*.

Qin, S.; Zhang, X.; Xu, H.; and Xu, Y. 2022. Fast Quaternion Product Units for Learning Disentangled Representations in $\mathbb{SO}(3)$. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Ruhe, D.; Gupta, J. K.; de Keninck, S.; Welling, M.; and Brandstetter, J. 2023. Geometric Clifford Algebra Networks. *arXiv preprint arXiv:2302.06594*.

Santos, E. E. 2002. Optimal and efficient algorithms for summing and prefix summing on parallel machines. *Journal of Parallel and Distributed Computing*, 62(4): 517–543.

Satorras, V. G.; Hoogeboom, E.; and Welling, M. 2021. E (n) equivariant graph neural networks. In *International conference on machine learning*, 9323–9332. PMLR.

Schütt, K.; Kindermans, P.-J.; Sauceda Felix, H. E.; Chmiela, S.; Tkatchenko, A.; and Müller, K.-R. 2017. Schnet: A continuous-filter convolutional neural network for modeling quantum interactions. *Advances in neural information processing systems*, 30.

Schütt, K.; Unke, O.; and Gastegger, M. 2021. Equivariant message passing for the prediction of tensorial properties and molecular spectra. In *International Conference on Machine Learning*, 9377–9388. PMLR.

Shen, W.; Zhang, B.; Huang, S.; Wei, Z.; and Zhang, Q. 2020. 3d-rotation-equivariant quaternion neural networks. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XX 16*, 531–547. Springer.

Stärk, H.; Beaini, D.; Corso, G.; Tossou, P.; Dallago, C.; Günnemann, S.; and Liò, P. 2022. 3d infomax improves gnns for molecular property prediction. In *International Conference on Machine Learning*, 20479–20502. PMLR.

Thomas, N.; Smidt, T.; Kearnes, S.; Yang, L.; Li, L.; Kohlhoff, K.; and Riley, P. 2018. Tensor field networks: Rotation-and translation-equivariant neural networks for 3d point clouds. *arXiv preprint arXiv:1802.08219*.

Wang, L.; Liu, Y.; Lin, Y.; Liu, H.; and Ji, S. 2022. ComENet: Towards complete and efficient message passing for 3D molecular graphs. *Advances in Neural Information Processing Systems*, 35: 650–664.

Xie, T.; and Grossman, J. C. 2018. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Physical review letters*, 120(14): 145301.

Zhang, X.; Qin, S.; Xu, Y.; and Xu, H. 2020. Quaternion product units for deep learning on 3d rotation groups. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 7304–7313.

Zhao, Y.; Birdal, T.; Lenssen, J. E.; Menegatti, E.; Guibas, L.; and Tombari, F. 2020. Quaternion equivariant capsule networks for 3d point clouds. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part I 16*, 1–19. Springer.

Zhu, X.; Xu, Y.; Xu, H.; and Chen, C. 2018. Quaternion convolutional neural networks. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 631–647.

Zitnick, L.; Das, A.; Kolluru, A.; Lan, J.; Shuaibi, M.; Sriram, A.; Ulissi, Z.; and Wood, B. 2022. Spherical channels for modeling atomic interactions. *Advances in Neural Information Processing Systems*, 35: 8054–8067.