Unifying O(3) equivariant neural networks design with tensor-network formalism

Zimu Li, Zihan Pengmei, Han Zheng, Erik Thiede, Junyu Liu, and Risi Kondor

1. Introduction

Graph neural networks (GNNs) have recently gained prominence in the field of chemistry, owing to their ability to learn from the structural properties of molecules and materials. Nevertheless, devising an efficient and accurate GNN architecture for investigating dynamic properties of chemical systems remains a formidable challenge. GNNs are adept at learning the structure of chemical systems and predicting their properties, including potential energy, dipole moment, and atomic forces. Recently, there has been a surge of interest in employing deep learning to forecast chemical properties and expedite first-principles dynamics simulations [1–5]. Specifically, GNNs have been utilized to estimate the potential energy with distinct atomic coordinates, where the negative gradient concerning the input coordinates naturally corresponds to the atomic force. Accurate prediction of potential energy and atomic force [6] necessitates adherence to spatial symmetries, such as translational and rotational covariance, since these properties are continuous functions defined on three-dimensional Euclidean space.

Abstract

Many learning tasks, including learning potential energy surfaces from ab initio calculations, involve global spatial symmetries and permutational symmetry between atoms or general particles. Equivariant graph neural networks are a standard approach to such problems, with one of the most successful methods employing tensor products between various tensors that transform under the spatial group. However, as the number of different tensors and the complexity of relationships between them increase, maintaining parsimony and equivariance becomes increasingly challenging. In this paper, we propose using fusion diagrams, a technique widely employed in simulating SU(2)-symmetric quantum many-body problems, to design new spatial equivariant components for neural networks. This results in a diagrammatic approach to constructing novel neural network architectures. When applied to particles within a given local neighborhood, the resulting components, which we term ‘fusion blocks,’ serve as universal approximators of any continuous equivariant function defined on the neighborhood. We incorporate a fusion block into pre-existing equivariant architectures (Cormorant and MACE), leading to improved performance with fewer parameters on a range of challenging chemical problems. Furthermore, we apply group-equivariant neural networks to study non-adiabatic molecular dynamics of stilbene cis-trans isomerization. Our approach, which combines tensor networks with equivariant neural networks, suggests a potentially fruitful direction for designing more expressive equivariant neural networks.
Machine learning algorithms employed to predict properties such as potential energy and atomic forces must yield consistent results, regardless of the molecule’s rotational pose or ordering. To address this challenge, researchers have developed group-equivariant neural networks that preserve these symmetries [5, 7–10]. In a group-equivariant network, symmetry operations on the data, including rotations of pictures and molecules, and permutations of the labels of each particle, commute with the network’s layers, ensuring that the same physical property is predicted irrespective of the input’s orientation. Many state-of-the-art spatially equivariant neural networks [11, 12] leverage the representation theory of the spatial rotation group in the so-called Fourier space [13]. These Fourier space methods employ the Clebsch–Gordan nonlinearities [13, 14]. In fact, as elucidated in the supplementary material (SM), Clebsch–Gordan nonlinearities are the sole source of nonlinearity in Fourier space determined by invariant theory in mathematics [15–17]. The Clebsch–Gordan nonlinearities further constrain the use of linear weights, which can only act on the multiplicity space corresponding to each irreducible representation [14].

Independently of work in machine learning, physicists have been using network models, called tensor networks, to represent complicated quantum many-body systems. Tensor networks are a family of methods for approximating larger tensors by contracting together a large collection of smaller tensors. Tensor networks have been used to successfully approximate large quantum states with low entanglement accurately by making use of the density matrix renormalization group in one dimension [18, 19] and introducing low-rank tensors to represent the quantum states [20]. Applications of tensor networks include quantum simulation of quantum physics problems [21–23], quantum computing and quantum supremacy experiments [24, 25], machine learning and data science [26–28], and quantum gravity [29, 30]. A special type of tensor networks concerns with global on-site SU(2) symmetry, called the spin networks, where multiple sites are fused by a prescribed fusion diagrams [31, 32] so that the global wavefunctions are SU(2) symmetric. The fusion diagrams, as we will show later and further in SM, are natural and sparse generalization of Clebsch–Gordan products among multiple sites. Fusion diagrams have found great success in simulate SU(2)-symmetric quantum systems [20, 32–34], and we will show their potential for constructing universal equivariant neural networks.

Fusion diagrams facilitate the classification of existing neural network architectures and inspire the development of novel equivariant blocks. We showcase the computational power of these blocks using classical results from invariant theory, which establish that under certain conditions, they can achieve universality. For instance, we employ fusion diagrams to construct a new SO(3)-equivariant block, which we incorporate into two state-of-the-art neural network architectures: Cormorant [35] and MACE [12]. We demonstrate that integrating the new equivariant layer significantly enhances the performance of both architectures, with a comparable or substantially fewer number of parameters.

To assess the validity of the fusion block, we carried out extensive experiments on various chemical systems, including standard benchmark datasets such as QM-9 [36] and MD-17 [6], which aims to predict the quantum properties of molecules and potential energy surfaces (PES), as well as more challenging systems like the non-adiabatic cis-trans isomerization of stilbene. Non-adiabatic isomerization of stilbene poses a considerable challenge learning the multiple boarder and reactive PESs, necessitating accurate interpolation and extrapolation.

In summary, this paper presents a novel method for constructing group-equivariant neural network blocks using fusion diagrams, a concept borrowed from theoretical physics. Our approach alleviates the combinatorial complexity associated with preserving symmetry constraints in neural networks, enabling the construction of expressive and universal equivariant layers. We demonstrate the effectiveness of the fusion block by incorporating our new SO(3)-equivariant layer into two state-of-the-art neural network architectures, Cormorant and MACE, and evaluating them on a variety of common benchmarks in companion with more complicated molecular isomerization and adsorption process. Our results indicate that the fusion block leads to improved performance with comparable or fewer parameters. Overall, our approach contributes to the developing a new routine that can be used to construct more expressive group equivariant neural networks.

2. Background

Before delving into the specifics of our approach, it is crucial to lay the groundwork with some foundational concepts. In this section, we offer an overview of relevant ideas from both machine learning and physics. We begin with a concise review of molecular dynamics and the significance of symmetry and equivariance in machine learning. Subsequently, we introduce the concept of tensor products and their role in theoretical physics, including a description of the fusion diagram notation.
2.1. Molecular dynamics

Molecular dynamics simulations are essential tools for studying molecular properties at the atomic level within specific timescales. To simulate atomic motion, we need to calculate the potential energy and atomic forces acting on molecules with particular geometric configurations in $\mathbb{R}^3$ space. Generally, potential energy and its gradients can be accurately determined by electronic structure calculations from first principles or approximated classically as simple analytical potential functions within specific chemical environments, such as atomic type, bond length, and bond angle. However, the electronic structure calculations under the ab initio molecular dynamics (AIMD) calculations are expensive.

One popular approach to overcoming this limitation is to use neural networks as interatomic potentials \cite{1,2}, which are trained with reference AIMD data. Training neural networks as interatomic potentials involves regressing on potential energy and atomic forces simultaneously, where predictive forces can be naturally achieved as the negative gradient of energy via back-propagation. The potential energy is invariant to 3D rigid rotations, while atomic forces are covariant to rotations, as they are vector values. The equivariant neural networks that we introduce subsequently are a powerful data-driven approach for an accurate representation of the chemical environment.

2.2. Representation theory of SU(2) and SO(3)

Rotationally equivariant nets are arguably one of the most successful types of equivariant neural networks. Let $X$ and $Y$ be the input and output spaces of a layer $L$, and let $T$ and $T'$ be linear actions of a group $G$ encoding the symmetry on $X$ resp. $Y$. The layer is said to be equivariant to $G$ if

$$T_g' \circ L = L \circ T_g$$

for all $g \in G$. \hspace{1cm} (1)

If the group action on the output space is the identity transformation, i.e. $T'_g y = y$ for all elements of $G$, the above reduces to

$$L = L \circ T_g \hspace{1cm} \forall g \in G$$

and we have an invariant layer. Constructing an equivariant neural network requires that both the learned affine function and the fixed nonlinear function obey equivariance. Kondor and Trivedi \cite{14} showed how to construct learned affine functions that are equivariant to compact groups (such as the group of rotations or the group of permutations) using the theory of linear representations. A linear representation \footnote{Linear representations should not be confused with the different use of the word ‘representation’ in representation learning.} of a compact group $G$ is a pair $(V, \rho)$ such that for each $g \in G$, $\rho(g)$ is a linear transformation of $V$ for which $\forall g_1, g_2 \in G, \rho(g_1)\rho(g_2) = \rho(g_1g_2)$. If the representation is finite-dimensional, the range of $\rho$ is a subset of the space of complex $k \times k$ matrices for some $k$. An irreducible linear representation (irrep) is a representation where $V$ has no proper subspaces preserved under $\rho$. Using well-known results in representation theory, we can apply a linear transformation that decomposes the inputs, outputs, and activations of a neural network into components that transform according to the group’s irreps. Then, one can show that the most general possible equivariant linear transformation can be written as matrix multiplication against each component \cite{14, 37–39}. The construction of linear equivariant layers where inputs and outputs transform according to linear group representations has been widely studied and used in today’s neural networks \cite{9, 10, 14, 37, 38, 40–42}.

For the rest of this work, we will focus on equivariance in the presence of SU(2) and SO(3) symmetries. These groups have fundamental importance in modern quantum physics and machine learning applications on geometric data. The irreps of SU(2) can be indexed by a single non-negative integer or half-integer, called the spin label. For any $g \in SU(2)$ and spin label $j$, we denote the corresponding matrix that arises from evaluating $\rho(g)$ as $W(j)$. It is well known in group theory that SO(3) irreps are isomorphic to the irreps of SU(2) with integer spin labels \cite{17, 43, 44}. This relationship allows us to study both SO(3) and SU(2) at the same time. Depending on the mathematical context, vectors in $V_j$ might transform either by $W(j)$ (contravariant transformation) or by its complex conjugate (covariant transformation). In what follows we focus only on irreps and omit $\rho_j$ when we denote irreps. To distinguish these two cases, we denote components of any vector $v$ transforming contravariantly by raised index $v^m$ with $w_m$ being defined accordingly for the covariant case. With the notion of raised and lowered indices, one can contract vectors like $v^m w_m$, where the Einstein summation convention will be used consistently in this paper.

With the above basic notions clarified, let us formally define the Clebsch–Gordan product. Take two SU(2) irreps $(\rho_a, V_{j_a})$ and $(\rho_b, V_{j_b})$ of spin $j_a$ and $j_b$ respectively. We can then define the tensor product representation $(\rho_a \otimes \rho_b, V_{j_a} \otimes V_{j_b})$. As this is still an SU(2) representation, it can be decomposed into irreps...
labeled by spins. A Clebsch–Gordan decomposition is a matrix \( C^{(j_a, j_b, j)} \) which transforms the tensor product \( W^{j_a}(g) \otimes W^{j_b}(g) \) into \( W^{j}(g) \) for a prescribed spin \( j \) and any \( g \in SU(2) \). Formally,

\[
C^{(j_a, j_b, j)} \cdot (W^{j_a}(g) \otimes W^{j_b}(g)) \cdot C^{(j_a, j_b, j)} = W^{j}(g) \quad \text{for all } g \in SU(2).
\]

By definition, the Clebsch–Gordan decomposition is equivariant with respect to the action of SU(2) as well as SO(3). Formally, it can be understood as an element from the space of SU(2) equivariant maps \( \text{Hom}(V_j \otimes V_{j_a}, V_{j_b}) \), where \( V_j, V_{j_a}, V_{j_b} \) are the corresponding SO(3) irreps with the angular momenta labels \( j_1, j_2, j_3 \). In this case, we can write the Clebsch–Gordan product as a third-order tensor:

\[
C^{(j_a, j_b, j)}_{m_1, m_2, m_3} \in \text{Hom}(V_{j_1} \otimes V_{j_2}, V_{j_3}),
\]

where \( m_1, m_2, m_3 \) are the corresponding magnetic quantum numbers. There are well-established methods to compute \( C^{(j_a, j_b, j)} \) both theoretically and algorithmically, e.g. \([43, 45, 46]\). Summing in Einstein notation with lower and upper indices, we call it a Clebsch–Gordan product for input vectors \( \psi^{(j_a)}, \psi^{(j_b)} \):

\[
\psi^{(j_a) m_1} = C^{(j_a, j_b, j)}_{m_1, m_2, m_3} \psi^{(j_b) m_2} \psi^{(j) m_3}.
\]

We will also leave \( m \) as entry indices and write Clebsch–Gordan product among input vectors as inner products:

\[
\psi^{(j_a)} = \langle C^{(j_a, j_b, j)}, \psi^{(j_b)} \otimes \psi^{(j)} \rangle.
\]

### 3. Methods

Here we demonstrate how fusion diagrams can be used to design equivariant components that we call ‘neural fusion blocks.’ We present an explicit construction for transformations under SU(2) and SO(3). In each block, we apply a collection of fusion diagrams \( \mathcal{D} \) to our input tensors. Each incoming edge of the diagram is associated with an input to the block and each outgoing edge is associated with an output of the block. We denote the collection of input tensors associated with incoming edges as \( \{ \psi^{(in)}_1, \ldots, \psi^{(in)}_n \} \) and components corresponding to the spin label \( j_1, \ldots, j_m \) are denoted as \( \{ \psi_{\tau_1}^{(in, j_1)}, \ldots, \psi_{\tau_n}^{(in, j_m)} \} \), where \( \tau_1, \ldots, \tau_n \) are the channel dimension indices. We omit batch indices from our treatment for brevity. The fusion block then acts according to algorithm 1. More illustrations on the updating rule with diagrammatic examples can be found in section I in the SM. Due to the use of fusion diagrams, the resulting algorithm is guaranteed to be equivariant to rotations. It also serves an rotationally-equivariant universal approximator where we put the proof details in section II in the SM.

As our focus is on the construction of specific components in an SO(3)-equivariant architecture rather than on proposing an entirely new architecture, we demonstrate the potential of our formalism by incorporating it into existing neural networks. Specifically, we choose to augment the Cormorant architecture proposed in [35] and the recent state-of-art model [47] with one additional three-body fusion block that replaces the conventional node-edge two-body interaction, with the aim of capturing inter-atomic interactions in a more faithful manner. Capturing three-body interactions in an SO(3) equivariant way with edge features could lead to a large overhead on computational resources. Applying fusion blocks to point clouds also requires ensuring that the resulting neural network obeys permutation symmetry. Since each fusion diagram has a single output, we can reinforce the permutation equivariance by passing these outputs through an aggregation function and incorporate them into existing message-passing-like mechanisms. It is worth mentioning that except employing Clebsch–Gordan products, there are other efficient architectures like using spherical coordinates of neighboring atoms and leveraging spherical harmonics to encode angular momentum information into higher dimensional representation of SO(3) and filtering through the spherical representations [5].
3.1. Cormorant with Fusion Diagrams (CoFD)

Cormorant is one of the first equivariant neural networks that utilize the group equivariance and designed to learn molecular dynamics simulations and ground-state molecular properties. A neuron in Cormorant layer \( s \) operates as follows:

\[
F_{i}^{-1} = \left[ F_{i}^{1} \oplus \left( F_{i}^{1} \odot_{cg} F_{i}^{1} \right) \oplus \sum_{j} \left( Y(x_{ij}) \odot_{cg} F_{j}^{1} \right) g_{ij} \right] \cdot W_{\text{vertex}}^{s,\ell}.
\]  

(10)

Here \( j \) sums over atom \( j \)'s local neighborhood, and \( g_{ij} \) is a learned rotationally invariant function that takes \( F_{i}^{1} \) and \( F_{j}^{1} \) as input in addition to other rotationally invariant features. \( \odot_{cg} \) denotes the channel-wise Clebsch–Gordan products and \( Y(x_{ij}) \) the spherical harmonics with input the relative displacement vectors between \( i \)th and \( j \)th atoms (we refer the reader to [35] for the precise functional form of \( g_{ij} \)). Each of these terms corresponds to the two-body diagram on the left below, while the product with \( g_{ij} \) is in a three-way product, it never has a covariant or contravariant component.

In particular, we observe that this layer has no equivariant interaction between equivariant parts of the activation atom \( i \) and the activation for atom \( j \). Instead, their activations only interact through the rotationally invariant function \( g \). Instead, in the present paper we employ our fusion diagrams to add an additional term to (10) that fully integrates all of the information between atom \( i \) and atom \( j \). This corresponds to the fusion diagram on the right. The resulting fusion block has three inputs: the input activation for atom \( i \), the input activation for atom \( j \), and the collection of spherical harmonics evaluated on their relative displacements, and one output: a new feature for atom \( i \). Consequently, we require a fusion diagram with three ingoing edges and one outgoing edge. Going from left to right, we input the representation of atom \( i \), the representation of atom \( j \), and the spherical harmonic representation of the edge connecting the two. We then incorporate this as a new term in (10), giving the following functional form for our modified Cormorant layer.

\[
F_{i}^{-1} = \left[ F_{i}^{1} \oplus \left( F_{j}^{1} \odot_{cg} F_{i}^{1} \right) \oplus \sum_{i} \left( Y(x_{ij}) \odot_{cg} F_{j}^{1} \right) g_{ij} \oplus F_{i}^{\text{fusion}} \right] \cdot W_{\text{vertex}}^{s,\ell},
\]  

(11)

where \( F_{i}^{\text{fusion}} \) is the output of the fusion block where inputs came from atom \( i \) and atom \( j \) within the coming legs chosen to be \( i \)th atom, \( j \)th atom, and their connecting edge. In other words, we use fusion diagrams to efficiently combine the atom-level messaging passing and edge-level message passing.

3.2. MACE with Fusion Diagrams (MoFD)

3.2.1. Implementation of 3-body fusion blocks

In our modified MACE architecture we use fusion diagrams (figure 1), a local neighborhood is defined by a cut-off radius \( r \), the information on the central particle is \( \Psi_{i}^{(in)} \), adjacent particle \( \Psi_{j}^{(in,j)} \), incident of edge \( i \in \mathcal{N}(o) \), and the incident edges \( \Phi_{i}^{(in,j)} \). In particular, this information is passed to a \( l \)-element sequence of independently 3-body fusion interactions given by a sequence of different internal spin spin configurations \( \{k_{1}, k_{2}, \ldots, k_{l}\} \), with the \( l \)-element sequence of outgoing activation on the center particle:

\[
\Psi_{\omega,k_{s}}^{(out,k_{s})} := \sum_{i \in \mathcal{N}(o)} \text{FusionBlock} \left( \Psi_{i}^{(in,j)} \Phi_{i}^{(in,j)} \Psi_{i}^{(in,k_{s})} \right).
\]  

(12)
Figure 1. Schematic illustration of the implementation of fusion blocks in the MACE architecture. For each atom the fusion block first fuses all the neighboring atoms for a given radius cut-off by pre-selected fusion diagram templates. Specifically, for each neighboring atom, we fuse the information from the root, neighbor atom, and their connecting edge. Then the fusion block applies an aggregation method: in the present work, we simply sum all the neighbors.

The final permutation invariant update to the center node information is obtained by concatenating \( \{ \Psi_{a,k_\alpha}^{(\text{out,j})} \}_{\alpha=1,...,l} \), followed by a linear mixing layer along the new concatenated axis. Note that in the sparse implementation, the feature dimension for each incoming activation never gets updated. Each time the internal spin configuration is only specified by a single internal spin label \( k_\alpha \), thus sparsifying the three-body information flow. For each internal spin value \( k \), the 3-body interaction fuses into a single \( SO(3) \)-equivariant tensor (this fusion corresponds to the figure 1 in the SM), while the final messaging passing aggregates neighboring edges and nodes information to the center node. In this implementation based on MACE architectures, we found our fusion block would only marginally increase the number of trainable parameters given the same channel width.

Fusion Block can also be initialized with significantly more trainable parameters than the original Mace does, which we denote as the dense implementation. The key difference to the sparse implementation is the inclusion of multi-partite internal spins to create a nested FusionBlock module. More specifically, given a \( l \)-sequence of internal spins \( \{ k_\alpha \} \), we can choose a tuple \((k_{\alpha_1}, k_{\alpha_2})\), a triple \((k_{\alpha_1}, k_{\alpha_2}, k_{\alpha_3})\), and beyond instead of specifying a single choice of the internal spin in the sparse implementation. Hence, a total of \( l! \) selections can be made resulting in a significant boost to the model size and number of trainable parameters. In our explicit implementation, we feed all choices of internal spins at once, resulting in a typical 10X boost of the trainable parameter size. As a result, we do not need to additionally pass a linear layer to reshape the channel width. The dense model could often outperform or be on par with the sparse implementation only with half the channel width. As an overall observation, the choices of internal spins are vital to our numerical performance. In our practice, the internal spins are chosen to range from \( j = 0, 1, 2 \), and sometimes with both parties.

4. Results

We describe three well-rounded benchmarks to test CoFD and MoFD, including QM-9 [36] molecular property prediction, MD-17 [6] small molecular dynamics, non-adiabatic molecular dynamics of stilbene. Our results are summarized in figure 2, tables 1 and 2.

4.1. Molecular properties and MD-17 molecular dynamics datasets

We first implement the fusion diagram on Cormorant architecture [35]. The standard QM-9 benchmark dataset [36] is used to test the performance of the CoFD model to predict molecular quantum properties of roughly 130 000 molecules in equilibrium, which contains multiple tasks of scalar value regression including atomization enthalpy, free energy, etc. In contrast, the MD-17 dataset [6] involves learning the ground-state PES and its gradient, for eight small organic molecules at room temperature from reference DFT calculations.

We compare the CoFD model and the original Cormorant model. The fusion diagram reduces the number of parameters in our networks, ensuring that we are not simply improving performance by adding additional parameters: for MD17, the networks with fusion diagrams have 135 393 parameters compared to...
Figure 2. (a) Illustration of photo-induced cis-trans isomerization of stilbene (b) initial and end configurations of three representative trajectories, which are Wigner-sampled [53]. (c) The one-dimensional cut of stilbene ground/ excited-state PESs by rotating the carbon-carbon bond as illustrated in (a), which illustrates the energetic diagram of stilbene isomerization process.

Table 1. Mean absolute error of various prediction targets on QM-9 (top) and conformational energies (in units of kcal mol$^{-1}$) on MD-17 (bottom), for both the original Cormorant architecture and our modified version that incorporates a fusion block. It should be noted that the CoFD models have significantly fewer parameters than the original Cormorant. We report the mean and standard deviation from multiple runs. In comparison, the model with lower predictive error has been bolded.

|                  | Cormorant | CoFD    |
|------------------|-----------|---------|
| $\alpha$ (bohr$^2$) | 0.085     | 0.088   |
| $\Delta \varepsilon$ (eV) | 0.061 | 0.062   |
| $\varepsilon_{\text{HOMO}}$ (eV) | 0.034 | 0.0391  |
| $\varepsilon_{\text{LUMO}}$ (eV) | 0.038 | 0.0347  |
| $\mu$ (D)         | 0.038     | 0.035   |
| $C_v$ (cal/mol K) | 0.026     | 0.0272  |
| $G$ (eV)          | 0.020     | 0.0135  |
| $H$ (eV)          | 0.021     | 0.0132  |
| $R^2$ (bohr$^2$) | 0.961     | 0.50    |
| $U$ (eV)          | 0.021     | 0.0130  |
| $U_0$ (eV)        | 0.022     | 0.0133  |
| ZPVE (meV)        | 2.027     | 1.43    |

|                  | Cormorant | CoFD    |
|------------------|-----------|---------|
| Aspirin          | 0.098     | 0.0951  |
| Ethanol          | 0.027     | 0.0241  |
| Malonaldehyde    | 0.041     | 0.0380  |
| Naphthalene      | 0.029     | 0.0321  |
| Salicylic Acid   | 0.066     | 0.0608  |
| Toluene          | 0.034     | 0.0316  |
| Uracil           | 0.023     | 0.0297  |

154 241 in the original Cormorant [35], and our QM9 neural network has 121 872 parameters compared to 299 808 in the original [35]. We report that the total time of training QM9 (resp. MD17) use 20 (resp. 12) hours with 256 Epoches, each with a mini-batch size of 64. Hence each epoch costs 281 (resp. 169) seconds. Code for our modified network can be found at https://github.com/ehthiede/diagram_corm. To be noted, the fusion block used in the CoFD to predict QM-9 and MD-17 is a sparse implementation. We did not use the dense implementation in predicting the QM-9 and MD-17 properties due to the large computational expense. However, it would be an interesting future direction to reduce the recourse overhead in the dense implementation, which would enable more subsequent experiments.
Table 2. Comparative analysis of MACE and MoFD models in dense and sparse implementations, evaluated on single and multiple independent non-adiabatic trajectories of cis-stilbene. The table presents the feature dimension, number of parameters (Num. of Param.), training set size (Train Size), and results for the ground state, first excited state, and second excited state. Results include energy values in milli-Hartree (mHartree) and forces in milli-Hartree per Angstrom (mHartree/Å). The bold figures represent the best performance in each category.

| Model            | Feature dimension | Num. of Param. | Train size | Ground state (Energy, Forces) | First Excited State (Energy, Forces) | Second excited state (Energy, Forces) |
|------------------|-------------------|---------------|------------|-------------------------------|-------------------------------------|--------------------------------------|
| MACE             | 64                | 330 320       | 285        | (19.15, 0.70)                 | (9.88, 1.25)                        | (21.80, 1.03)                       |
| MoFD-sparse      | 16                | 66 784        | 285        | (19.74, 0.62)                 | (13.43, 1.12)                       | (23.42, 1.07)                       |
| MACE             | 128               | 979 088       | 950        | (26.44, 1.30)                 | (29.07, 3.56)                       | (48.77, 3.05)                       |
| MoFD-sparse      | 32                | 141 168       | 950        | (28.84, 1.40)                 | (36.69, 3.55)                       | (55.08, 3.11)                       |
| MoFD-dense       | 16                | 690 976       | 950        | (27.59, 1.14)                 | (32.64, 3.31)                       | (54.76, 2.65)                       |

4.2. Stilbene Non-adiabatic molecular dynamics

Non-adiabatic MD (NAMD) [48] is a powerful approach for predicting photo-induced chemical processes, including photo-catalytic reactivity [49], photo-induced DNA damage [50], and the performance of Sun-screening products [51]. Unlike ground-state dynamics, NAMD involves evaluating multiple PESs and their gradients simultaneously. However, studying excited-state dynamics requires higher accuracy electronic structure methods than DFT [52], resulting in significantly higher computational costs. Thus, there is motivation to test our model’s ability to study multiple PESs that are not generated by DFT.

In this study, we explore the photo-induced cis-trans isomerization process of stilbene, a phenomenon first reported by Syage et al. [54]. Our approach utilizes the complete active space self-consistent field theory [55], specifically targeting the conjugated π orbital localized on the carbon-carbon double bond and its anti-bonding counterpart. This selection forms our active space, characterized as two electrons in two orbitals (2e,2o), and all calculations are conducted using the 6-31G* basis set. To accurately capture the quantum effects inherent in photoisomerization, we adopt a quantum-classical approximation through trajectory surface hopping, as implemented in the SHARC package [56]. This method integrates both quantum and classical dynamics, crucial for studying processes like isomerization. Wigner sampling [45] is employed to generate a variety of initial configurations, initiating the molecular trajectories under study.

A stringent criterion is applied to ensure the quality of the data: only trajectories maintaining total energy conservation within 0.2 eV were considered valid and included in the dataset. This threshold ensures the physical relevance of the trajectories by excluding those that do not adhere to energy conservation principles. The resultant dataset, therefore, comprises multiple molecular trajectories of stilbene, predominantly initiated in an excited state. These trajectories provide a comprehensive view of the isomerization process, offering valuable insights into the dynamics of this photochemical reaction. Detailed computational specifications and a more thorough introduction to the methods employed are available in the SM.

The widely-adopted MD17 dataset [6] comprises adiabatic dynamic trajectories using the PBE functional, though the spin polarization, basis set, and computational grid information are absent from the literature, near equilibrium, where molecular movements are trivial. As a result, MD17 is heavily biased towards sampling the reactant region of the PES without considering the driven non-equilibrium forces [57]. However, a meaningful chemical reaction typically involves three parts on the PES: reactant, product, and transition state. It is important to note that the accuracy of common density functionals is usually a few kcal mol⁻¹ when compared to higher levels of theory. For example, the PBE functional used in the MD17 dataset has an average error of more than 9 kcal mol⁻¹ (roughly 0.4 eV) when predicting reaction barriers [58]. In contrast, the trajectories we sampled visited the reactant, product, and transition state regions of multiple PESs, as illustrated in figure 2.

To compare the performance of MACE and MoFD with sparse implementation, we selected one reactive trajectory and employed the MACE model with a feature channel dimension of 64 and high-order equivariant features with l = 0, 1, 2, 3. For MoFD, we maintain the same feature angular momentum and set the feature channel dimension to 16, resulting in a model with only 66 784 parameters, an order of magnitude smaller than that of MACE. Given the increased difficulty in predicting atomic forces, we adjust the training loss on energy and forces with a ratio of 1:1000, as recommended in previous literature [11, 12]. As the loss is disproportionately weighted towards the force, we concentrate on the force regression performance. The models are trained on 285 samples and tested on a separate hold-out test set of 428 samples. The models were trained in a state-specific fashion, which means each model regresses single state’s PES and
forces for comparison purposes. Our findings indicate that MoFD with sparse implementation has a decent performance in force prediction for the first two states, while MACE fits better when predicting the energy. We further assess the generalization ability of our models across different trajectories by incorporating two additional independent trajectories into the dataset, resulting in a total of 950 training samples and 1395 hold-out testing samples. We increase the complexity of MACE by expanding its feature channel width to 128, leading to a total of 690 976 parameters. Concurrently, we double the feature dimension of MoFD to 32, making it only as large as MACE. In terms of runtime, each epoch requires 52 s in the MACE model compared to 31 s in the MoFD model, attributed to the utilization of lower-dimensional angular momentum features as inputs. The MoFD model with the dense implementation surpass MACE in the force prediction tasks, while the MoFD model with the sparse implementation remains comparable to MACE’s accuracy as indicated in table 2. Nonetheless, it is crucial to note that the performance of all models decrease when learning excited states due to the less well-defined topologies of excited-state PESs [52].

5. Discussion

In this work, we have introduced a new method for constructing equivariant blocks for rotation-equivariant layers based on fusion diagrams. Previous work has shown that tensor products can be used to construct neurons for rotation-equivariant neural networks. Moreover, prior research has observed that neural network ansatzes for the quantum system can be unified with spin network ansatzes. Our work is the first to employ these connections in the opposite direction: by employing diagrammatic methods used in physics, we construct new components that can be incorporated into equivariant neural networks.

Using classic results from invariant theory, we show that neural networks built from using fusion blocks are capable of approximating any continuous SU(2)-equivariant functions. To demonstrate the practical utility of fusion blocks, we perturb existing SO(3) equivariant neural network architectures, such as Cormorant [35] and MACE [12], by incorporating a fusion block in each layer. The modified architectures generally achieves better performance for a smaller number of parameters. Indeed, the idea of using equivariance and symmetry to prune neural networks has been applied [59] in the quantum setting. We believe this indicates that fusion blocks can be a useful addition to group-equivariant neural networks.

To test the performance of the fusion block approach, we apply the revised CoFD and MoFD models not only to the standard benchmark datasets QM-9 [36] and MD-17 [6], but also novel applications such as non-adiabatic molecular dynamics. We find that the addition of the fusion blocks improved the performance of the models.

In future work, we hope to use fusion blocks to improve the interpretability of equivariant neural networks. In theoretical physics, fusion diagrams represent physical processes that correspond to many-body interactions. Furthermore, physicists often manipulate fusion diagrams through internal permutations through a process known as recoupling. Recouplings relate to the physical properties of different fusion diagrams and can show symmetries present in the products that may not be immediately apparent by inspection. Employing the formalism of recoupling may highlight hidden symmetries in the network architecture, indicating new ways to save computational effort. Employing the language of fusion diagrams in these settings could help unify our physical picture of fusion diagrams with computational realities.

Finally, fusion diagrams are graphical representations of ways in which local atoms are being fused. It is of interest to consider the effect of the local subgraph topology on the corresponding fusion blocks; in particular, whether fusion diagrams serve as a general principle towards building more expressive graph neural nets with 3D equivariance specific to chemical applications. We leave addressing these questions as future research opportunities.

Data availability statement

The data cannot be made publicly available upon publication because they are not available in a format that is sufficiently accessible or reusable by other researchers. The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments

J L is supported in part by International Business Machines (IBM) Quantum through the Chicago Quantum Exchange, and the Pritzker School of Molecular Engineering at the University of Chicago through AFOSR MURI (FA9550-21-1-0209).
ORCID iDs

Zimu Li  https://orcid.org/0000-0003-1473-6492
Junyu Liu  https://orcid.org/0000-0003-1669-8039

References

[1] Wang H, Zhang L, Han J and Weinan E 2018 Deepmd-kit: a deep learning package for many-body potential energy representation and molecular dynamics Comput. Phys. Commun. 228 178–84
[2] Schoenholz S and Cubuk E D 2020 JAX MDE: a framework for differentiable physics Advances in Neural Information Processing Systems vol 33 pp 11428–41
[3] Atz K, Grisson F and Schneider G 2021 Geometric deep learning on molecular representations Nat. Mach. Intell. 3 1023–32
[4] Walters W P and Barzilay R 2020 Applications of deep learning in molecule generation and molecular property prediction Acc. Chem. Res. 54 263–70
[5] Shuaibi M et al 2021 Rotation invariant graph neural networks using spin convolutions (arXiv:2106.09575)
[6] 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 Sci. Adv. 3 e1603015
[7] Esteves C 2020 Theoretical aspects of group equivariant neural networks (arXiv:2004.05154)
[8] Cohen T S et al 2021 Equivariant convolutional networks (available at: https://arxiv.org/abs/1602.07576)
[9] García Soto V, Hoogeboom E and Welling M 2021 E(6) equivariant graph neural networks (arXiv:2102.09844)
[10] Du W et al 2021 SE(3) equivariant graph neural networks with complete local frames (arXiv:2110.14811)
[11] Batatia I et al 2021 E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials (arXiv:2101.03164)
[12] Batatia I, Kovács D P, Simm G N, Ortner C and Csányi G 2022 Mace: higher order equivariant message passing neural networks for fast and accurate force fields (arXiv:2206.07697)
[13] Kondor R, Lin Z and Trivedi S 2018 Clebsch-Gordan Net: a fully Fourier space spherical convolutional neural network Advances in Neural Information Processing Systems (NeurIPS)
[14] Kondor R and Trivedi S 2018 On the generalization of equivariance and convolution in neural networks to the action of compact groups (arXiv:1802.03690)
[15] Olver P J 1999 Classical Invariant Theory (London Mathematical Society Student Texts) (Cambridge University Press)
[16] Procesi C 2007 Lie Groups: An Approach Through Invariants and Representations (University Press)
[17] Goodman R and Wallach N R 2009 Symmetry, Representations and Invariants (Springer)
[18] White S R 1992 Density matrix formulation for quantum renormalization groups Phys. Rev. Lett. 69 2363
[19] Rommer S and Oßlund S 1997 Class of ansatz wave functions for one-dimensional spin systems and their relation to the density matrix renormalization group Phys. Rev. B 55 2164
[20] Orús R 2019 Tensor networks for complex quantum systems Nat. Rev. Phys. 1 388–50
[21] Fannes M, Nachtergaele B and Werner R F 1992 Finite correlated states on quantum spin chains Commun. Math. Phys. 144 443–90
[22] Verstraete F and Cirac I J 2004 Renormalization algorithms for quantum-many body systems in two and higher dimensions (available at: https://arxiv.org/abs/cond-mat/0407066)
[23] Vidal G 2007 Entanglement renormalization Phys. Rev. Lett. 99 220405
[24] Huang C et al 2021 Efficient parallelization of tensor network contraction for simulating quantum computation Nat. Comput. Sci. 1 578–87
[25] Pan F and Zhang P 2021 Simulating the symmetric quantum supremacy circuits (available at: https://arxiv.org/abs/2103.03074)
[26] Stoudenmire E and Schwab D J 2016 Supervised learning with tensor networks Advances in Neural Information Processing Systems vol 29
[27] Efthymiou S, Hilary J and Leichenauer S 2019 Tensor network for machine learning (available at: https://arxiv.org/abs/1906.06329)
[28] Roberts C et al 2019 Tensor-network: a library for physics and machine learning (arXiv:1905.01330)
[29] Swingle B 2012 Entanglement renormalization and holography Phys. Rev. D 86 065007
[30] Pastawski F, Yoshida B, Harlow D and Preskill J 2015 Holographic quantum error-correcting codes: toy models for the bulk/boundary correspondence J. High Energy Phys. JHEP06(2015)149
[31] Måkinen I 2019 Introduction to SU(2) recoupling theory and graphical methods for loop quantum gravity (arXiv:1910.06821)
[32] Singh S and Vidal G 2012 Tensor network states and algorithms in the presence of a global SU(2) symmetry Phys. Rev. B 86 195114
[33] Schmoll P, Singh S, Rizzi M and Orús R 2020 A programming guide for tensor networks with global SU(2) symmetry Ann. Phys., NY 419 168232
[34] Schmoll P and Orús R 2020 Benchmarking global SU(2) symmetry in two-dimensional tensor network algorithms Phys. Rev. B 102 241101
[35] Anderson B, Hy T-S and Kondor R 2019 Cormorant: covariant molecular neural networks (arXiv:1906.04015)
[36] Ramakrishnan R, Durlu P O, Rupp M and von Lilienfeld O A 2014 Quantum chemistry structures and properties of 134 kilo molecules Sci. Data 1 140022
[37] Zaheer M et al 2017 Deep sets (arXiv:1703.06114)
[38] Segol N and Lipman Y 2019 On universal equivariant set networks (arXiv:1910.02421)
[39] Maron H, Litany O, Chechik G and Fetaya E 2020 On learning sets of symmetric elements Int. Conf. on Machine Learning (PMLR) pp 6734–44
[40] Cohen T S and Welling M 2016 Group equivariant convolutional networks (arXiv:1602.07576)
[41] Maron H, Ben-Hamu H, Shamir N and Lipman Y 2019 Invariant and equivariant graph networks Int. Conf. on Learning Representations
[42] Thiede E H, Hy T S and Kondor R 2020 The general theory of permutation equivariant neural networks and higher order graph variational encoders (arXiv:2004.03900)
[43] Varshalovich D A, Moskalev A N and Khersonskii V K 1988 Quantum Theory of Angular Momentum (World Scientific)
[44] Lindgren I and Morrison J 1986 Atomic Many-Body Theory (Springer)
[45] Biedenharn L C, Giovannini A and Louck J D 1967 Canonical definition of wigner coefficients in un J. Math. Phys. 8 691–700
[46] Gao M-J and Chen J-Q 1985 The Clebsch-Gordan coefficients of permutation groups S(2)–S(6) J. Phys. A 18 189–213
[47] Batatia I et al 2022 The design space of E(3)-equivariant atom-centered interatomic potentials (arXiv:2205.06643)
[48] Tully J C 2012 Perspective: nonadiabatic dynamics theory J. Chem. Phys. 137 22A301
[49] Tong H, Ouyang S, Bi Y, Umezawa N, Oshikiri M and Ye J 2012 Nano-photocatalytic materials: possibilities and challenges Adv. Mater. 24 229–51
[50] Dougherty T J, Gomer C J, Henderson B W, Jori G, Kessel D, Korbelik M, Moan J and Peng Q 1998 Photodynamic therapy J. Natl Cancer Inst. 90 889–905
[51] Baker L A, Marchetti B, Karsili T N, Stavros V G and Ashfold M N 2017 Photoprotection: extending lessons learned from studying natural sunscreens to the design of artificial sunscreen constituents Chem. Soc. Rev. 46 3770–91
[52] Mai S, Atkins A J, Plasser F and González L 2019 The influence of the electronic structure method on intersystem crossing dynamics. The case of thioformaldehyde J. Chem. Theory Comput. 15 3470–80
[53] Dahl J P and Springborg M 1988 The morse oscillator in position space, momentum space and phase space J. Chem. Phys. 88 4535–47
[54] Syage J, Lambert W R, Felker P, Zewail A and Hochstrasser R 1982 Picosecond excitation and trans-cis isomerization of stilbene in a supersonic jet: dynamics and spectra Chem. Phys. Lett. 88 266–70
[55] Roos B O, Taylor P R and Sigbahn P E 1980 A complete active space SCF method (CASSCF) using a density matrix formulated super-Cl approach Chem. Phys. 48 157–73
[56] Mai S, Marquetand P and González L 2018 Nonadiabatic dynamics: the sharc approach Wiley Interdiscip. Rev.-Comput. Mol. Sci. 8 e1370
[57] Pengmei Z, Liu J and Shu Y 2023 xxmd: benchmarking neural force fields using extended dynamics beyond equilibrium (arXiv:2308.11155)
[58] Mardirossian N and Head-Gordon M 2017 Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals Mol. Phys. 115 2315–72
[59] Wang X et al 2022 Symmetric pruning in quantum neural networks (arXiv:2208.14057)