Chemical Science

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: Y. Zhang and H. Guo, *Chem. Sci.*, 2026, DOI: 10.1039/D5SC07248D.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/chemical-science

View Article Online
DOI: 10.1039/D5SC07248D
Submitted to Chem. Sci., 09/17/2025, revised 12/10/2025

Node-Equivariant Message Passing for Efficient and Accurate Machine Learning Interatomic Potentials

Yaolong Zhang* and Hua Guo

Department of Chemistry and Chemical Biology, Center for Computational Chemistry, University of New Mexico, Albuquerque, New Mexico 87131, USA

* Corresponding author, email: ylzhangch@unm.edu

Abstract

Machine learned interatomic potentials, particularly equivariant message-passing (MP) models, have demonstrated high fidelity in representing first-principles data, revolutionizing computational studies in materials science, biophysics, and catalysis. However, these equivariant MP models still incur substantial computational and memory needs due to their expensive tensor product operations over edge space, significantly limiting their applicability in large-scale or long-time simulations. In this work, we propose a novel node-equivariant MP (NEMP) framework that performs equivariant operations between the central node and a virtual summed node encoding structure information of its neighbors. Crucially, NEMP maintains comparable or even superior accuracy across diverse test systems—including molecules, extended systems, and universal potential benchmarks—while achieving 1-2 orders of magnitude

View Article Online
DOI: 10.1039/D5SC07248D

reduction in memory and computational costs compared to edge equivariant MP models. In fact, NEMP reaches computational efficiency comparable to that of local descriptor-based models, and enabling previously inaccessible large-scale simulations.

Introduction

Molecular dynamics (MD) is an essential tool for investigating both equilibrium and non-equilibrium properties of systems in gaseous and condensed phases. The reliability of such simulations depends crucially on the accuracy of the interaction potential. In the meantime, it is also important to realize that the system size and timescale that can be practically realized by such simulations are limited by the computational efficiency of these potentials. Hence, there is high desire to construct accurate and efficient interaction potentials. Recent advances in machine learning (ML) have revolutionized this field by representing first-principles data with efficient and high-fidelity machine-learned interatomic potentials (MLIPs)[1-33], enabling applications across various fields of chemistry, physics, and materials science[17, 32, 34].

A widely successful class of MLIPs builds upon the atomistic neural network (NN) framework introduced by Behler and Parrinello in 2007[1], which decomposes the total interaction potential into atomic contributions dependent on local atomic environments within a cutoff radius. These environments are encoded using many-body atomic descriptors that preserve translational, rotational and permutational symmetries. In early models, only two and three-body terms are included[1, 11], based on the "near-sightedness" assumption for the interaction potential. However, such local descriptors are insufficiently sensitive to configurations with long-range differences. Studies have

View Article Online
DOI: 10.1039/D5SC07248D

shown that increasing the body order of interactions in the descriptors can enhance model accuracy[13, 35-37]. To systematically incorporate higher-order terms, several approaches have been proposed, including the Moment Tensor Potential (MTP)[5] and the Atomic Cluster Expansion (ACE)[12]. However, numerical costs increase rapidly with the rising body order, which significantly limits the practical use of explicit higher-order terms in large systems.

An alternative and increasingly popular strategy for enhancing expressiveness is the message-passing NNs (MPNNs), which describe atomic environments by iteratively exchanging information between atoms (nodes) through connections between them (edges) in molecular graphs[7, 10, 20, 21, 23-27, 29, 38, 39]. This end-to-end learning framework is capable of implicitly capturing many-body interactions and some non-local effects beyond the cutoff radius[23]. Consequently, they are generally more efficient than NNs with explicit high body-order terms in improving the description of local structures. Early MPNN models[7, 10, 40] were based on two-body message functions, but subsequent studies showed that explicitly incorporating three-body (or higher-order) interactions into the message function allows for a more effective discrimination of structures, which is crucial to increase the expressiveness of the MPNN model[21, 23, 41]. These strategies only pass symmetry-invariant scalar features and can be categorized as invariant MPNNs. Although invariant MPNNs can significantly improve model accuracy[21, 23, 42], they lack directional coupling between different atomic environments, which can lead to failure in certain systems where such interactions are critical[14, 27, 43]. This drawback can be easily remedied by passing equivariant features instead[24-26, 29, 39,

View Article Online
DOI: 10.1039/D5SC07248D

[43-51]. These equivariant MPNNs extend the traditional paradigm by learning and exchanging equivariant features that transform predictably under symmetry operations. Indeed, models like NequIP[25], Allegro[29] and MACE[26] employ equivariant tensor products via the Clebsch-Gordan coefficients, progressively and explicitly increasing body-order interactions with each MP layer.

Until now, all equivariant MPNNs perform equivariant message passing along edge connections, where tensor products are computed between edges and their corresponding nodes. Hence, we refer to such an approach as edge-equivariant message passing (EEMP). While EEMPs have achieved high fidelity across diverse benchmarks[25, 26, 29, 52], they require an enormous computational overhead and memory requirement. Moreover, MPNNs face inherent limitations in parallelization efficiency due to massive inter-environment communication. This constraint limits their scalability to moderate-scale systems (for example, simulating liquid water comprising thousands to tens of thousands of atoms on an NVIDIA A100 GPU with 80 GB of memory[39, 53]). Although some equivariant MPNNs reduce costs by replacing tensor products with direct summations of equivariant features[25, 43, 51], such simplifications often compromise model accuracy.

In this work, we introduce a novel equivariant MPNN framework that leverages node-equivariant message passing (NEMP) to improve computational efficiency without sacrificing accuracy. Our key innovation lies in the reformulation of equivariant message passing in the node space, which are significantly smaller than the edge space. Rather than computing tensor products between every edge and

View Article Online
DOI: 10.1039/D5SC07248D

corresponding neighbor node, we construct an expressive equivariant framework through a tensor product between the central node and a virtual summed node, where the neighbor features are summed with edge-dependent, highly expressive coefficients. These coefficients and features are iteratively refined during MP, preserving physical symmetries while maintaining or even increasing the model's expressive power. As a result, the expensive scaling in the equivariant MP step is reduced from with the number of neighboring pairs to with the number of central atoms. Extensive evaluations across molecular systems, gas-solid interfaces, liquids and solids demonstrate that our framework achieves high fidelity in representing the first-principles data while improving computational and memory efficiency by 1-2 orders of magnitude. This improvement allows not only reduction of simulation time, but also enable simulations of much larger systems.

Methods

Equivariant MPNNs have gained significant attention in atomic-scale ML for their remarkable fidelity and data efficiency[34]. In these models, molecular structures are considered as graphs, where atoms are treated as nodes and edges are defined as the connections between the central atom and its neighbors within a specified cutoff radius. The MP is realized iteratively in layers, in which each node gathers information from its surrounding atoms, typically encoded with symmetric features, and combines them through a learned aggregation function, which is termed massage. The aggregated information is then used to pass the message by updating the node features. Through

View Article Online
DOI: 10.1039/D5SC07248D

repeated layers, the network captures increasingly complex correlations and interactions across the molecular graph. The value of equivariant MPNNs lies in their ability to couple these node states through the tensor product of equivariant features of the central atom and its neighbors. This mechanism not only preserves equivariance but also introduces explicit incorporation of higher-order interactions, in contrast to invariant MPNNs, which capture such interactions only implicitly[23, 54].

The most critical component of equivariant MPNNs lies in how they perform equivariant message passing. In EEMP models, equivariant messages are passed by coupling equivariant features of each edge with those of the corresponding neighbor node, as illustrated in Fig. 1(a). This strategy results in two tensors of shape ($N_{atom} \times N_{neigh} \times N_{lmc} \times N_c$), where N_{atom} is the number of atoms in the system, N_{neigh} is the number of neighbors within the cutoff radius, N_{lmc} denotes the number of angular momentum combinations, and N_c is the number of channels (*vide infra*). In particular, N_{lmc} increases rapidly with the maximum angular momentum, exhibiting cubic scaling. This often results in the construction of extremely large tensors and, consequently, significant computational cost in both time and memory. As a result, the tensor product becomes the rate-determining step in MP, severely limiting the applicability of such models to large systems or long-time simulations. Below, we propose the NEMP to avoid some of the shortcomings of EEMP.

NEMP Architecture

The NEMP architecture is shown in Fig. 1(b). We first need to generate a

View Article Online
DOI: 10.1039/D5SC07248D

representation for each unique pair of atoms to enable its application in describing multi-element systems, which is crucial for building a universal potential across various elements. This can be achieved through an embedding NN, denoted as F_{emb}, which takes the combination of a pair of elements $\{Z_i, Z_j\}$ as input[29]:

$$\boldsymbol{\sigma}_{ij} = F_{emb}\left(\{Z_i, Z_j\}\right). \quad (1)$$

Here, Z_i and Z_j represent the element types of the central and neighbor atoms, respectively. The output of this embedding depends only on the combination of element types and is used to generate parameters that similarly depend only on $\boldsymbol{\sigma}_{ij}$ in subsequent processing:

$$\mathbf{c}(\boldsymbol{\sigma}_{ij}) = F_{coeff}\left(\boldsymbol{\sigma}_{ij}\right). \quad (2)$$

For the NEMP description of the local edge environment within a cutoff radius of the *i*th central atom, Bessel functions with optimizable parameters are used for the radial channels:

$$R_n(r_{ij}) = \frac{\sin\left(c_n^R\left(\boldsymbol{\sigma}_{ij}\right) r_{ij}\right)}{c_n^R\left(\boldsymbol{\sigma}_{ij}\right) r_{ij}} f_{cut}(r_{ij}; r_c). \quad (3)$$

Here, j denotes the surrounding atoms within the cutoff radius r_c, enforced by a polynomial cutoff function[55] $f_{cut}(r_{ij}; r_c)$. $c_n^R\left(\boldsymbol{\sigma}_{ij}\right)$ is the output of the NN (F_{coeff}) defined in the Eq. (2) for the radial functions (denoted by superscript R), with n ranging from 1 to N_c, where N_c is the number of channels (radial functions) in the initial layer.

Before detailing the specific operations in each MP layer, we first introduce the notational convention used throughout the manuscript, where bold symbols denote

View Article Online
DOI: 10.1039/D5SC07248D

vectors with channel dimensions. The notations "I" and "J" refer to the central node (*i*) and a neighboring node (*j*), respectively, while "IJ" denotes the edge connecting them. The first Greek subscript indicates the specific type of feature, which include edge coefficients (ECs) (χ), equivariant (ψ), and invariant (ρ) features. The numerical superscript denotes the MP layer; when a quantity has no superscript, it represents an intermediate value within the layer that is not passed to the next MP layer. The Greek superscript (applicable only to the edge coefficients) specifies the particular subset used for a given operation. For example, $\mathbf{IJ}_{\chi}^{0}$ represents the edge coefficients of the 0th MP layer. Summation over J (capital J on the left, index *j* on the right) produces the node equivariant (ψ) or invariant (ρ) feature $\mathbf{I}_{\rho}^{0}=\sum_{j}\mathbf{IJ}_{\chi}^{0,\rho}$. Here, $\mathbf{IJ}_{\chi}^{0,\rho}$ denotes the subset of $\mathbf{IJ}_{\chi}^{0}$ used to compute the node invariant feature $\mathbf{I}_{\rho}^{0}$. Other symbols representing different layers or feature types can be understood in the same way.

By applying an NN to the combination of radial functions in Eq. (3) and the element-dependent coefficients $\mathbf{c}^{\chi}\left(\boldsymbol{\sigma}_{ij}\right)$ (Eq. (2)), we generate the environment dependent ECs of the initial layer:

$$\mathbf{IJ}_{\chi}^{0}=F_{edge}\left(\left\{\mathbf{c}^{\chi}\left(\boldsymbol{\sigma}_{ij}\right),\quad \mathbf{R}\left(r_{ij}\right)\right\}\right). \tag{4}$$

The corresponding node-equivariant features ($\mathbf{I}_{\psi,lm}^{0}$) for the central atom *i* can be obtained by directly summing the contributions of their neighboring atoms denoted by *j*:

$$\mathbf{I}_{\psi,lm}^{0}=\sum_{j}\mathbf{IJ}_{\chi}^{0,\psi}Y_{lm}(\mathbf{r}_{ij}). \tag{5}$$

Here, $Y_{lm}(\mathbf{r}_{ij})$ are the spherical harmonics, $\mathbf{IJ}_{\chi}^{0,\psi}$ are the initial layer ECs defined in

View Article Online
DOI: 10.1039/D5SC07248D

Eq. (4) related to the node-equivariant features. On the other hand, the node-invariant features ($\mathbf{I}_{\rho}^{0}$) in the initial layer can be generated by summation over the ECs:

$$\mathbf{I}_{\rho}^{0}=\sum_{j}\mathbf{IJ}_{\chi}^{0,\rho}\,. \tag{6}$$

Similarly, edge-invariant features are derived from node-equivariant features of the initial layer defined in Eq. (5):

$$\mathbf{IJ}_{\rho}=\sum_{l,m}\left(\mathbf{I}_{\psi,lm}^{0}\mathbf{IJ}_{\chi}^{0,\rho_i}Y_{lm}(\mathbf{r}_{ij})+\mathbf{J}_{\psi,lm}^{0}\mathbf{IJ}_{\chi}^{0,\rho_j}Y_{lm}(\mathbf{r}_{ij})\right). \tag{7}$$

These newly computed edge-invariant features $\mathbf{IJ}_{\rho}$ are then combined with the edge-invariant features from the initial layer $\left\{\mathbf{c}^{\rho}\left(\boldsymbol{\sigma}_{ij}\right),\ \mathbf{IJ}_{\chi}^{0}\right\}$ to produce refined edge-invariant features for the current layer $\mathbf{IJ}_{\rho}^{1}=\left\{\mathbf{c}^{\rho}\left(\boldsymbol{\sigma}_{ij}\right),\ \mathbf{IJ}_{\chi}^{0,\rho_{ij}},\ \mathbf{IJ}_{\rho}\right\}$. Using these updated features, the ECs ($\mathbf{IJ}_{\chi}^{1}$) for the first layer are generated using the same form as that in the initial layer as shown in Eq. (4). With these ECs, we perform a tensor contraction between the weighted spherical harmonic expansion and the node-equivariant features from the previous layer to obtain the new node-invariant features ($\mathbf{I}_{\rho}^{1}$):

$$\mathbf{I}_{\rho}^{1}=\sum_{l,m}\mathbf{I}_{\psi,lm}^{0}\sum_{j}\mathbf{IJ}_{\chi}^{1,\rho}Y_{lm}(\mathbf{r}_{ij})\,. \tag{8}$$

To overcome the bottleneck associated with EEMP, as previously discussed, we design our MP scheme such that the tensor product is performed only over the node space, which is significantly smaller than the edge space. For each central node, we first obtain its equivariant features ($\mathbf{I}_{\psi,lm}$):

$$\mathbf{I}_{\psi,lm}=\sum_{j}\mathbf{IJ}_{\chi}^{1,\psi}Y_{lm}(\mathbf{r}_{ij})\,. \tag{9}$$

View Article Online
DOI: 10.1039/D5SC07248D

To acquire information outside the cutoff sphere, we introduce a virtual summed node (~), as illustrated in Fig. 1(a). The corresponding equivariant feature ($\mathbf{I}^{1}_{\psi,lm,n}$), can be obtained by aggregating or summing over its neighboring nodes

$$\mathbf{I}^{1}_{\psi,lm,n} = \sum_{j} IJ^{1,\psi}_{\chi,n} \sum_{k=1} \omega^{1}_{j,k} \mathbf{J}^{0}_{\psi,lm,k}. \tag{10}$$

The message passing is then realized by performing a tensor product between the node-equivariant features of the central and virtual nodes:

$$\mathbf{I}^{1}_{\psi,l_f m_f} = \sum_{l_1,m_1,l_2,m_2} C_{l_1,m_1,l_2,m_2,l_f,m_f} \omega^{1}_{i,l_1,l_2,l_f} \mathbf{I}^{1}_{\psi,l_1 m_1} \mathbf{I}_{\psi,l_2 m_2}, \tag{11}$$

where, $C_{l_1,m_1,l_2,m_2,l_f,m_f}$ are the corresponding Clebsch–Gordan coefficients. This strategy produces two tensors of shape ($N_{atom} \times N_{lmc} \times N_c$). For comparison, the message-passing operation in EEMP model can be written as:

$$\mathbf{I}_{\psi,l_f m_f} = \sum_{j} \sum_{l_1,m_1,l_2,m_2} C_{l_1,m_1,l_2,m_2,l_f,m_f} \mathbf{R}(r_{ij}) Y_{lm}(\mathbf{r}_{ij}) \mathbf{J}_{\psi,l_2 m_2}, \tag{12}$$

which produces two intermediate tensors of shape ($N_{atom} \times N_{neigh} \times N_{lmc} \times N_c$). As a result, it is significantly more efficient in terms of both time and memory, removing the bottleneck in EEMP. Considering the case where the central atom has 10 neighboring atoms within the cutoff radius, a situation not uncommon in extended systems, NEMP would in principle be one order of magnitude faster than EEMP.

The final node-equivariant features for the first layer are obtained through a ResNet-style connection[25, 56]:

$$\mathbf{I}^{1}_{\psi,lm,k} = \sum_{n=1} \omega^{0}_{i,k} \mathbf{I}^{0}_{\psi,lm,n} + \sum_{n=1} \omega^{1}_{i,k} \mathbf{I}^{1}_{\psi,lm,n}. \tag{13}$$

In Eqs. (10) and (13), ω represents different optimizable parameters, where the subscripts indicate their dependencies. For example, $\omega^{1}_{i,k}$ in Eq. (10) denotes that the

View Article Online
DOI: 10.1039/D5SC07248D

optimizable parameter depends on the element species (i) and the channel coupling (ν). Importantly, in Eqs. (10) and (13), we use element-dependent coefficients for basis contraction before further product or sum coupling. This approach facilitates information exchange between different channels, similar to basis contraction in quantum chemistry, and has been shown to enhance expressiveness[57].

Subsequently, the node-equivariant features are employed to generate the edge-invariant features using the same form in first layer as shown in Eq. (7), allowing the message-passing process to continue as shown in Fig. 1(b) and (c). After T MP iterations, the refined edge-invariant features converge to a sufficiently expressive representation of the local edge environment, naturally producing fully resolved coefficients. In this context, the aggregations in Eqs. (9) and (10) can be viewed as locally information-preserving when the number of channels is sufficiently large. For example, the mapping from edge state to the node state forms a small linear system whose coefficient matrix (the ECs) can be inverted in practice, enabling the original edge information to be recovered. Consequently, through the tensor product in Eq. (11), our model does not sacrifice expressive power, but with much lower computational cost and memory consumption than the EEMP. Similarly, the node-invariant features develop a faithful representation of the entire system structure. The model aggregates (e.g., concatenates or integrates) node-invariant features of each layer and processes them through an NN to predict atomic energies. The total energy of the system is finally obtained by summing these atomic contributions $E=\sum_{i}^{N} F_{readout}\left(\left\{\mathbf{I}_{\rho}^{0}, \quad \mathsf{L} \quad , \quad \mathbf{I}_{\rho}^{T}\right\}\right)$.

Furthermore, the sum operation can introduce additional interactions, as shown in

View Article Online
DOI: 10.1039/D5SC07248D

Fig. 1(a) and (c). In the EEMP model, each edge-equivariant feature (spherical harmonics) only interacts with the corresponding neighbor's node-equivariant features. For example, using the local environment centered at atom *I* depicted in Fig. 1(a), the interaction of an EEMP model is expressed as

$$\mathbf{IJ}_{\chi}(r_{ij})\mathbf{Y}(\mathbf{r}_{ij})\otimes\mathbf{J}_{\psi}+\mathbf{IK}_{\chi}(r_{ik})\mathbf{Y}(\mathbf{r}_{ik})\otimes\mathbf{K}_{\psi}+\mathbf{IH}_{\chi}(r_{ih})\mathbf{Y}(\mathbf{r}_{ih})\otimes\mathbf{H}_{\psi}, \quad (14)$$

Here, *J, H,* and *K* denote the neighboring nodes. Similarly, in Fig. 1(c), the interaction in the NEMP model is defined as follows:

$$\left(\mathbf{IJ}^{\psi}_{\chi}\mathbf{Y}(\mathbf{r}_{ij})+\mathbf{IK}^{\psi}_{\chi}\mathbf{Y}(\mathbf{r}_{ik})+\mathbf{IH}^{\psi}_{\chi}\mathbf{Y}(\mathbf{r}_{ih})\right)\otimes\left(\mathbf{IJ}^{\psi}_{\chi}\mathbf{J}_{\psi}+\mathbf{IK}^{\psi}_{\chi}\mathbf{K}_{\psi}+\mathbf{IH}^{\psi}_{\chi}\mathbf{H}_{\psi}\right), \quad (15)$$

There are more direct interactions and many-body based coefficients in our model since coefficients in EEMPs only depend on distances between the central atom and its neighbors, while our NEMP model depends on iteratively refined many-body edge features. The ECs are learned from data to automatically determine interaction significance. Numerical results presented below show that our model can achieve state-of-the-art accuracy across diverse systems, while offering substantially higher efficiency and requiring fewer parameters (approximately 50k to 500k approximately 50k–500k; see Supplementary Information (SI) for details) compared with EEMP models[25, 52, 58].

Results and Discussion

Single-component systems

3BPA

View Article Online
DOI: 10.1039/D5SC07248D

For MLIP models, transferability is a crucial measurement for assessing model performance. To evaluate the transferability of NEMP, we employ the flexible drug-like molecule 3-(benzyloxy)pyridin-2-amine (3BPA) dataset, sampled from MD simulations at 300, 600, and 1200 K[59]. The reference energies and forces were computed using density functional theory (DFT) with the *w*B97X functional and the 6-31G basis set, as described in ref. [59]. Following the same protocol, our model uses the same training set as all models listed in Table I to ensure a fair comparison. Specifically, 500 structures from the 300 K dataset were employed and randomly split into training and validation sets with a 9:1 ratio. In addition, configurations sampled at 300, 600, and 900 K were used to construct three separate test datasets. Given the substantial energy range difference between the high-temperature test datasets (~4.5 eV) and the training dataset (1.1 eV), we adopt a linear mapping function (F_{edge} in Eq. (4)) to transform the edge-invariant features $\mathbf{IJ}_{\rho}^{0}$ into edge coefficients $\mathbf{IJ}_{\chi}^{0}$ in each MP layer to enhance transferability.

Comparisons with several existing MLIP models, including ACE[59], SGDML[60], ANI-2x[61], CACE[39], NequIP[25], Allegro[29], and MACE[26], are summarized in Table I. Notably, NEMP achieves superior fidelity compared to all models except MACE, for which the errors are practically comparable. This supports our claim that NEMP is effectively equivalent to the EEMP framework in terms of accuracy. We note that MACE incorporates additional many-body interactions by leveraging node-equivariant features as a basis after each MP layer, enabling higher-order representations, so it is not surprising that it achieves better accuracy than other MLIPs. As demonstrated in

View Article Online
DOI: 10.1039/D5SC07248D

examples discussed below, however, introducing an NN to map the edge-invariant features to ECs in each layer can mitigate this issue, yielding comparable or even better performance than MACE.

Liquid water

To validate the performance of the NEMP model on complex liquids, we benchmark it against ab initio reference data generated by Cheng et al.[35] The dataset comprises 1,593 structures of 64 water molecules in a periodic box, computed using DFT with the revPBE0-D3 functional. This level of theory has been demonstrated to provide a reliable description of water's structure and dynamics across a range of pressures and temperatures. In Table II, we compare the performance of several models, including (1) local descriptor-based atomic NNs, (2) invariant MPNNs, and (3) equivariant MPNNs. Among these, two and three-body descriptor-based NNs like the Behler-Parrinello NN (BPNN)[35] and the embedded atom neural network (EANN)[11, 62] give the largest errors, as expected. REANN (invariant MPNN)[57] significantly reduces the errors of local descriptor-based models, reaching comparable accuracy with equivariant models NequIP[25] and CACE[39]. Once again, MACE[58] demonstrated superior fidelity, achieving a root mean square error (RMSE) of 1.9 meV/H_2O for energy and 36.2 meV/Å for forces, outperforming other EEMP models. Notably, our NEMP model, which incorporates an NN as the edge-feature update function F_{edge}, achieves errors of to 2.4 meV/H_2O and 36.8 meV/Å while using significantly fewer parameters (see SI for details). This result supports our claim that introducing NN-mapped many-body ECs enhances the model's expressive power.

View Article Online
DOI: 10.1039/D5SC07248D

To further evaluate the reliability of the NEMP potential, we performed Nose-Hoover thermostat-based NVT MD simulations of the liquid water system consisted of 64 H_2O molecules in a periodic cubic box, equilibrated at 300 K. The simulations were carried out for 50 ps with a time step of 0.1 fs. Fig. 2 compares the oxygen-oxygen radial distribution function (RDF) obtained from the NEMP potentials with that from DFT using the same exchange–correlation functional and identical simulation conditions[63]. The close agreement further demonstrates the accuracy in reproducing the reference DFT energies and interatomic forces. A substantially longer 1 ns NVT simulation of a 512-molecule H_2O system further confirms the stability and robustness of our potential. Crucially, both simulation results demonstrate a good agreement with experimental measurements. The consistency across different system sizes confirms the convergence of our MD simulations and indicates the long-term stability and scalability of our potential to larger systems.

Universal Potential Applications

ANI-1x

After demonstrating the high fidelity of the NEMP model for single-component systems, we further evaluate its generalization ability for constructing universal potentials. To this end, we benchmark the model using the ANI-1x dataset—a dataset for training transferable universal potentials for organic molecules with H, C, N, and O elements[64]. Following the protocol adopted in MACE, we trained our model on a subset of ~490k data points (approximately 10% of the full dataset) and evaluated its

View Article Online
DOI: 10.1039/D5SC07248D

performance on the COMP6 benchmark dataset[65]. As presented in Table III, the NEMP model outperforms all existing approaches (including ANI-1x[66], equivariant transformer based model TrIP, EEMP model MACE[58] and TensorNet[67]), achieving a total mean absolute error (MAE) 0.39 kcal/mol for energy and 0.34 kcal/mol/Å for forces. It should be noted that the comparison with the transformer model might not be fully conclusive regarding the effect of the transformer alone, as the transformer and equivariant components were not implemented on an identical base architecture. More rigorous comparisons in prior studies[49, 68] have demonstrated that the equivariant framework contributes more to performance gains than the transformer architecture alone. Moreover, integrating both components yielded the superior overall results. More importantly, comparisons with the EEMP model are consistent with the findings above, confirming that NN-mapped many-body ECs can enhance the model's expressive power. Notably, while the training data only included systems with $\leq$50 atoms, the model successfully extrapolates to larger molecular systems (up to 300 atoms). These results highlight NEMP as a promising and scalable framework for developing general-purpose potentials.

HME21

To further test the performance of the NEMP model for materials science applications, we choose to work with the HME21 dataset[69], which includes regular and disordered crystals spanning 37 elements. This diverse chemical space makes HME21 an ideal benchmark for evaluating the generalizability of MLIPs for predicting structures and energies of materials. Accurately modeling systems with such a large

View Article Online
DOI: 10.1039/D5SC07248D

elemental variety has long been considered a key challenge in MLIP construction[41, 50, 70-73]. For fair comparison, we use the same data split as in the original test. As shown in Table IV, NEMP achieves accuracy comparable to MACE[58] while outperforming NequIP[25, 58] and TeaNet[47, 58]. The consistently low errors confirm NEMP's ability to handle extensive elemental diversity without sacrificing precision in energy and force predictions.

EMLP dataset

The Element-based Machine Learning Potential (EMLP)[74] is a recent NequIP trained MLIP encompassing Ag, Pd, C, H, and O, designed for applications in heterogeneous catalysis and beyond. Comprising of 116,516 DFT-calculated data points generated through random chemical space exploration, the dataset prioritizes diverse local atomic interactions over extensive collections of structurally similar configurations to ensure generalizability. EMLP demonstrates accurate property predictions across solid, liquid, gas phases and gas-surface systems without requiring system-specific sampling. Perhaps most importantly, it demonstrates a superior ability in predicting chemical transformation barriers (including transition state geometries and energies), which is often lacking in other MLIPs. Furthermore, it has also been demonstrated to correctly describe liquid molecular systems with only limited data. The inherent complexity and demonstrated transferability of the EMLP dataset presents it as an excellent benchmark for evaluating performance of MLIPs.

Following the same training strategy, we developed an NEMP model using the EMLP dataset and evaluated its performance by comparing predicted energetic profiles

View Article Online
DOI: 10.1039/D5SC07248D

for CO oxidation on various Pd surfaces against reference DFT calculations. As shown in Fig. 3, both models show good agreement with DFT results, with NEMP achieving comparable or even superior accuracy to EMLP in certain cases, consistent with our observations in other systems discussed above.

Beyond gas–surface systems, we further validated this universal potential through MD simulations of liquid methanol. We carried out Nose–Hoover thermostated NVT simulations, with a system comprising 32 CH_3OH molecules in a periodic cubic cell (side length: 12.93 Å) equilibrated at 300 K. The simulation was run for 50 ps with a time step of 0.1 fs. Fig. 4 compares the C–C, C–O, and O–O RDFs predicted by the NEMP potential with those from ab initio molecular dynamics (AIMD), EMLP, and experimental results[75]. The NEMP predictions are obviously in closer agreement with AIMD than those of EMLP, which underscores the accuracy of NEMP in reproducing reference energies and interatomic forces, further confirming its robust transferability across different phases and systems.

Computational Efficiency and Scalability of NEMP

By design, the NEMP represents significant advantages in computational efficiency over existing equivariant models. As displayed in Table I for the 3BPA molecule, NEMP reduces the computational cost by approximately 1-2 order of magnitude compared to EEMP baselines (MACE/NequIP/Allegro) while maintaining comparable accuracy.

To evaluate CPU scalability, we measured the computational cost per MD step in NVT simulations of liquid water (300 K) using the JAX_MD package[76]. Fig. 5(a)

View Article Online
DOI: 10.1039/D5SC07248D

demonstrates linear scaling with system size on a single Intel® Xeon Gold 6438Y core. The NEMP potential shows marginally higher computational cost than BPNN[1, 62] (an efficient local-descriptor atomistic model). For a more detailed comparison, we reduced the number of hyperparameters in our model, creating an NEMP-small variant with RMSE of 3.1 meV/H_2O for energy and 47.9 meV/Å for forces, maintaining accuracy comparable to EEMP models. This model achieves a speed of 2.5×10^{-4} s/atom/step, representing approximately 2.0× acceleration over the standard NEMP and 1.5× faster performance than the efficient local-descriptor-based BPNN model. Notably, while the REANN (invariant MPNN) design theoretically promises higher speed than EEMP, both NEMP variants demonstrate similar performance in our CPU benchmarks.

For GPU benchmarks, memory emerges as a critical constraint due to hardware limitations. Our solution is a novel asynchronous architecture (Fig. 6) that decouples neighbor-list (NL) construction on the CPU from force evaluation on the GPU. The implementation combines JAX's non-blocking execution with a skin-algorithm cutoff buffer to create pipelined updates: the GPU computes forces using historical NLs (time *t-n*) while the CPU simultaneously generates new NLs based on current positions (time *t*) through optimized Fortran code, with asynchronous NL transfers to the GPU every *n* steps. This design eliminates GPU waiting time by masking NL latency and overcomes the high memory consumption of NL calculations for general lattice parameters in JAX_MD required for just-in-time (JIT) compilation compatibility (limited to ~50k atoms), enabling nearly 800K atoms simulations. This strategy can be readily adapted to other ML models for enhanced memory and computational efficiency for efficient

View Article Online
DOI: 10.1039/D5SC07248D

GPU models.

Benchmarks on NVIDIA A100/H100 GPUs (Fig. 5(b)) demonstrate the computational and memory efficiency of our model. Comparison between Figs. 5(a) and 5(b) reveals GPU/CPU acceleration ratios that scale with system size, reaching a maximum of ~200 × for systems exceeding ~10k atoms. NEMP-small successfully simulates nearly 800K atoms within 80 GB memory, in contrast to EEMP models (MACE/NequIP/CACE) which are limited to 10k ~ 24k atoms with a speed around $0.9\times10^{-4} \sim 1.5\times10^{-4}$ s/atom/step on A100 GPUs with 80 GB memory. Notably, NEMP-small achieves 5× faster evaluation speeds than REANN. Compared to EEMP models, NEMP-small reduces computational costs by two orders of magnitude while improving memory efficiency by 30-70× on A100 GPUs, further achieving 0.8 μs/atom/step on H100 GPUs. Benchmark results validate our model's high computational efficiency and low memory consumption.

It is important to note that this asynchronous strategy is most effective when the evaluation time of the ML model is comparable to, or shorter than, the overhead associated with CPU-GPU data transfer, or when using frameworks that require JIT compilation, such as JAX-MD, where constructing NLs for general lattice parameters incurs substantial memory overhead. In contrast, when the cost of force evaluation significantly exceeds these overheads, a conventional sequential setup, where NL construction and MD steps are performed entirely on either the CPU or the GPU, can also achieve near-optimal performance.

Conclusions

View Article Online
DOI: 10.1039/D5SC07248D

We present here a novel node-based equivariant MPNN framework that achieves high efficiency and accuracy in modeling MLIPs across diverse chemical systems. Central to our approach is the tensor product operation between node states and virtual summed node states, weighted by many-body ECs generated through NNs within each MP layer. This design significantly reduces computational costs while enhancing expressive power.

Benchmark results demonstrate that our model achieves accuracy comparable to or exceeding prevailing EEMP models across datasets designed for gas-phase molecules, extended systems, and universal potentials. MD simulations confirm 1~2 orders of magnitude acceleration relative to existing EEMP models while maintaining accuracy, achieving speeds surpassing even efficient local-descriptor methods like BPNN.

Our asynchronous GPU-CPU algorithm resolves memory constraints by decoupling NL construction on CPUs from force evaluation on GPUs. Crucially, standard implementations of GPU-based NL building in the JAX_MD package must satisfy JIT compilation constraints, incurring prohibitive memory overhead for general lattice parameters. By shifting NL operations to CPUs with optimized Fortran codes, we enable simulations of nearly 800K atoms compared to the 50k-atom limit imposed by the JAX_MD implementation without any extra overhead. This is significant as the NEMP can be used to simulate much larger systems than permitted by the existing EEMPs.

The current computational bottleneck lies in the EC calculation, where an NN processes edge-invariant input features. Although the channel contractions in Eqs. (10)

View Article Online
DOI: 10.1039/D5SC07248D

and (13) exhibit quadratic scaling, these operations remain subdominant as they only operate on the node dimensions. The EC evaluation, which scales linearly with the number of channels, can be further accelerated using TensorFloat-32 optimization. When combined with the parallel algorithms for MPNN frameworks, such as that proposed recently by Xia and Jiang[77], our approach can be further extended to larger-scale systems. The NEMP model's unique features make it particularly suited for advancing MD simulations and for the development of accurate, efficient universal potentials for chemical, biophysical, and materials studies at ab initio accuracy. Furthermore, the exceptional performance demonstrates that our strategy provides a simple and general approach to upgrading other EEMP models, improving efficiency in both computational time and memory usage.

Data availability:

The NEMP Package is available from https://github.com/zhangylch/NEMP. All datasets used for model training and validation in this study are publicly accessible from previously published sources.

Author contributions:

YZ and HG designed the project. YZ wrote the code and performed all calculations. YZ and HG discussed results and wrote the paper.

Conflicts of interest:

View Article Online
DOI: 10.1039/D5SC07248D

There are no conflicts to declare.

Acknowledgements:

This work was supported by Department of Energy (Grant No. DE-SC0015997 to H.G.). The computation was performed at the Center for Advanced Research Computing (CARC) at UNM. We gratefully acknowledge Dr. Peijun Hu and Dr. Wenbo Xie for providing the EMLP dataset and the corresponding validation data for the potential.

Reference

1. J. Behler and M. Parrinello, *Phys. Rev. Lett.*, 2007, **98**, 146401.
2. A. P. Bartók, M. C. Payne, R. Kondor and G. Csányi, *Phys. Rev. Lett.*, 2010, **104**, 136403.
3. B. Jiang and H. Guo, *J. Chem. Phys.*, 2013, **139**, 054112.
4. K. Shao, J. Chen, Z. Zhao and D. H. Zhang, *J. Chem. Phys.*, 2016, **145**, 071101.
5. A. V. Shapeev, *Multiscale Model. Sim.*, 2016, **14**, 1153-1173.
6. S. Chmiela, A. Tkatchenko, H. E. Sauceda, I. Poltavsky, K. T. Schutt and K. R. Muller, *Sci. Adv.*, 2017, **3**, e1603015.
7. K. T. Schütt, H. E. Sauceda, P.-J. Kindermans, A. Tkatchenko and K.-R. Müller, *J. Chem. Phys.*, 2018, **148**, 241722.
8. L. Zhang, J. Han, H. Wang, R. Car and W. E, *Phys. Rev. Lett.*, 2018, **120**, 143001.
9. K. Yao, J. E. Herr, David W. Toth, R. McKintyre and J. Parkhill, *Chem. Sci.*, 2018, **9**, 2261-2269.
10. O. T. Unke and M. Meuwly, *J. Chem. Theory Comput.*, 2019, **15**, 3678-3693.

View Article Online
DOI: 10.1039/D5SC07248D

11. Y. Zhang, C. Hu and B. Jiang, *J. Phys. Chem. Lett.*, 2019, **10**, 4962-4967.

12. R. Drautz, *Phys. Rev. B*, 2019, **99**, 014104.

13. V. Zaverkin and J. Kästner, *J. Chem. Theory Comput.*, 2020, **16**, 5410-5421.

14. O. T. Unke, S. Chmiela, H. E. Sauceda, M. Gastegger, I. Poltavsky, K. T. Schütt, A. Tkatchenko and K.-R. Müller, *Chem. Rev.*, 2021, **121**, 10142-10186.

15. F. Musil, A. Grisafi, A. P. Bartók, C. Ortner, G. Csányi and M. Ceriotti, *Chem. Rev.*, 2021, **121**, 9759-9815.

16. J. Behler, *Chem. Rev.*, 2021, **121**, 10037-10072.

17. M. Meuwly, *Chem. Rev.*, 2021, **121**, 10218-10239.

18. B. Huang and O. A. von Lilienfeld, *Chem. Rev.*, 2021, **121**, 10001-10036.

19. V. L. Deringer, A. P. Bartók, N. Bernstein, D. M. Wilkins, M. Ceriotti and G. Csányi, *Chem. Rev.*, 2021, **121**, 10073-10141.

20. R. Zubatyuk, S. Smith Justin, J. Leszczynski and O. Isayev, *Sci. Adv.*, 2021, **5**, eaav6490.

21. O. T. Unke, S. Chmiela, M. Gastegger, K. T. Schütt, H. E. Sauceda and K.-R. Müller, *Nat. Commun.*, 2021, **12**, 7273.

22. P. O. Dral, F. Ge, B.-X. Xue, Y.-F. Hou, M. Pinheiro, J. Huang and M. Barbatti, *Top. Curr. Chem.*, 2021, **379**, 27.

23. Y. Zhang, J. Xia and B. Jiang, *Phys. Rev. Lett.*, 2021, **127**, 156002.

24. K. Schütt, O. Unke and M. Gastegger, Proceedings of the 38th International Conference on Machine Learning, 2021. **139**, 9377-9388.

25. S. Batzner, A. Musaelian, L. Sun, M. Geiger, J. P. Mailoa, M. Kornbluth, N. Molinari, T.

View Article Online
DOI: 10.1039/D5SC07248D

E. Smidt and B. Kozinsky, *Nat. Commun.*, 2022, **13**, 2453.

26. I. Batatia, D. P. Kovacs, G. Simm, C. Ortner and G. Csanyi, Advances in Neural Information Processing Systems, 2022. **35**, 11423-11436.

27. J. Thorben Frank, O. T. Unke and K.-R. Müller, Advances in Neural Information Processing Systems, 2022. **35**, 29400-29413.

28. Y. Zhang and B. Jiang, *Nat. Commun.*, 2023, **14**, 6424.

29. A. Musaelian, S. Batzner, A. Johansson, L. Sun, C. J. Owen, M. Kornbluth and B. Kozinsky, *Nat. Commun.*, 2023, **14**, 579.

30. Y. Zhang, Q. Lin and B. Jiang, *WIREs Comput. Mol. Sci.*, 2023, **13**, e1645.

31. I. Batatia, L. L. Schaaf, H. Chen, G. Csányi, C. Ortner and F. A. Faber, 2023, arXiv:2310.10434.

32. X.-T. Xie, Z.-X. Yang, D. Chen, Y.-F. Shi, P.-L. Kang, S. Ma, Y.-F. Li, C. Shang and Z.-P. Liu, *Precis. chem.*, 2024, **2**, 612-627.

33. Q. Yu, R. Ma, C. Qu, R. Conte, A. Nandi, P. Pandey, P. L. Houston, D. H. Zhang and J. M. Bowman, *Nat. Comput. Sci.*, 2025, **5**, 418-426.

34. J. Xia, Y. Zhang and B. Jiang, *Chem. Soc. Rev.*, 2025, **54**, 4790-4821.

35. S. D. Huang, C. Shang, P. L. Kang and Z. P. Liu, *Chem. Sci.*, 2018, **9**, 8644-8655.

36. S. N. Pozdnyakov, M. J. Willatt, A. P. Bartók, C. Ortner, G. Csányi and M. Ceriotti, *Phys. Rev. Lett.*, 2020, **125**, 166001.

37. Z. Fan, Z. Zeng, C. Zhang, Y. Wang, K. Song, H. Dong, Y. Chen and T. Ala-Nissila, *Phys. Rev. B*, 2021, **104**, 104309.

38. O. Unke, M. Bogojeski, M. Gastegger, M. Geiger, T. Smidt and K.-R. Müller, Advances

View Article Online
DOI: 10.1039/D5SC07248D

in Neural Information Processing Systems, 2021. **34**, 14434–14447.

39. B. Cheng, *npj Comput. Mater.*, 2024, **10**, 157.

40. K. T. Schütt, F. Arbabzadah, S. Chmiela, K. R. Müller and A. Tkatchenko, *Nat. Commun.*, 2017, **8**, 13890.

41. C. Chen and S. P. Ong, *Nat. Comput. Sci.*, 2022, **2**, 718-728.

42. R. Zubatyuk, J. S. Smith, B. T. Nebgen, S. Tretiak and O. Isayev, *Nat. Commun.*, 2021, **12**, 4870.

43. Y. Wu, J. Xia, Y. Zhang and B. Jiang, *J. Phys. Chem. A*, 2024, **128**, 11061-11067.

44. N. Thomas, T. Smidt, S. Kearnes, L. Yang, L. Li, K. Kohlhoff and P. Riley, *arXiv*, 2018, arXiv:1802.08219.

45. B. Anderson, T. S. Hy and R. Kondor, Advances in Neural Information Processing Systems 32, 2019, 14537-14546.

46. J. Brandstetter, R. Hesselink, E. van der Pol, E. J. Bekkers and M. Welling, *arXiv*, 2021, arXiv:2110.02905.

47. S. Takamoto, S. Izumi and J. Li, *Comput. Mater. Sci.*, 2022, **207**, 111280.

48. M. Haghighatlari, J. Li, X. Guan, O. Zhang, A. Das, C. J. Stein, F. Heidar-Zadeh, M. Liu, M. Head-Gordon, L. Bertels, H. Hao, I. Leven and T. Head-Gordon, *Digit. Discov.*, 2022, **1**, 333-343.

49. S. Gong, Y. Zhang, Z. Mu, Z. Pu, H. Wang, X. Han, Z. Yu, M. Chen, T. Zheng, Z. Wang, L. Chen, Z. Yang, X. Wu, S. Shi, W. Gao, W. Yan and L. Xiang, *Nat. Mach. Intell.*, 2025, **7**, 543-552.

50. Y.-L. Liao, B. Wood, A. Das and T. Smidt, *arXiv*, 2023, arXiv:2306.12059.

View Article Online
DOI: 10.1039/D5SC07248D

51. M. Wen, W.-F. Huang, J. Dai and S. Adhikari, *npj Comput. Mater.*, 2025, **11**, 128.

52. I. Batatia, S. Batzner, D. P. Kovács, A. Musaelian, G. N. C. Simm, R. Drautz, C. Ortner, B. Kozinsky and G. Csányi, *Nat. Mach. Intell.*, 2025, **7**, 56-67.

53. C. W. Tan, M. L. Descoteaux, M. Kotak, G. d. M. Nascimento, S. R. Kavanagh, L. Zichi, M. Wang, A. Saluja, Y. R. Hu, T. Smidt, A. Johansson, W. C. Witt, B. Kozinsky and A. Musaelian, *arXiv*, 2025, arXiv:2504.16068.

54. J. Han, *arXiv*, 2024, arXiv:2407.11756.

55. K. Johannes, G. Janek and G. Stephan, International Conference on Learning Representations, 2020.

56. K. He, X. Zhang, S. Ren and J. Sun, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016. **1512**, 770-778.

57. Y. Zhang, J. Xia and B. Jiang, *J. Chem. Phys.*, 2022, **156**, 114801.

58. D. P. Kovács, I. Batatia, E. S. Arany and G. Csanyi, *J. Chem. Phys.*, 2023, **159**.

59. D. P. Kovács, C. v. d. Oord, J. Kucera, A. E. A. Allen, D. J. Cole, C. Ortner and G. Csányi, *J. Chem. Theory Comput.*, 2021, **17**, 7696-7711.

60. S. Chmiela, H. E. Sauceda, K.-R. Müller and A. Tkatchenko, *Nat. Commun.*, 2018, **9**, 3887.

61. X. Gao, F. Ramezanghorbani, O. Isayev, J. S. Smith and A. E. Roitberg, *J. Chem. Info. Model.*, 2020, **60**, 3408-3415.

62. Y. Zhang, C. Hu and B. Jiang, *Phys. Chem. Chem. Phys.*, 2021, **23**, 1815-1821.

63. O. Marsalek and T. E. Markland, *J. Phys. Chem. Lett.*, 2017, **8**, 1545-1551.

64. J. S. Smith, R. Zubatyuk, B. Nebgen, N. Lubbers, K. Barros, A. E. Roitberg, O. Isayev

View Article Online
DOI: 10.1039/D5SC07248D

and S. Tretiak, *Sci. Data.*, 2020, **7**, 134.

65. J. S. Smith, B. Nebgen, N. Lubbers, O. Isayev and A. E. Roitberg, *J. Chem. Phys.*, 2018, **148**, 241733.

66. V. Zaverkin, D. Holzmüller, L. Bonfirraro and J. Kästner, *Phys. Chem. Chem. Phys.*, 2023, **25**, 5383-5396.

67. G. Simeon and G. De Fabritiis, *Advances in Neural Information Processing Systems*, 2023, **36**, 37334-37353.

68. B. E. Hedelius, D. Tingey and D. Della Corte, *J. Chem. Theory Comput.*, 2024, **20**, 199-211.

69. S. Takamoto, C. Shinagawa, D. Motoki, K. Nakago, W. Li, I. Kurata, T. Watanabe, Y. Yayama, H. Iriguchi, Y. Asano, T. Onodera, T. Ishii, T. Kudo, H. Ono, R. Sawada, R. Ishitani, M. Ong, T. Yamaguchi, T. Kataoka, A. Hayashi, N. Charoenphakdee and T. Ibuka, *Nat. Commun.*, 2022, **13**, 2991.

70. I. Batatia, P. Benner, Y. Chiang, A. M. Elena, D. P. Kovács, J. Riebesell, X. R. Advincula, M. Asta, M. Avaylon and W. J. Baldwin, *arXiv*, 2023, arXiv:2401.00096.

71. D. Zhang, X. Liu, X. Zhang, C. Zhang, C. Cai, H. Bi, Y. Du, X. Qin, A. Peng, J. Huang, B. Li, Y. Shan, J. Zeng, Y. Zhang, S. Liu, Y. Li, J. Chang, X. Wang, S. Zhou, J. Liu, X. Luo, Z. Wang, W. Jiang, J. Wu, Y. Yang, J. Yang, M. Yang, F.-Q. Gong, L. Zhang, M. Shi, F.-Z. Dai, D. M. York, S. Liu, T. Zhu, Z. Zhong, J. Lv, J. Cheng, W. Jia, M. Chen, G. Ke, W. E, L. Zhang and H. Wang, *npj Comput. Mater.*, 2024, **10**, 293.

72. K.-i. Nomura, S. Hattori, S. Ohmura, I. Kanemasu, K. Shimamura, N. Dasgupta, A. Nakano, R. K. Kalia and P. Vashishta, *J. Phys. Chem. Lett.*, 2025, **16**, 6637-6644.

View Article Online
DOI: 10.1039/D5SC07248D

73. T. Shiota, K. Ishihara, T. M. Do, T. Mori and W. Mizukami, *arXiv*, 2024, arXiv:2412.13088.

74. C. Yang, C. Wu, W. Xie, D. Xie and P. Hu, *Nature Catalysis*, 2025, **8**, 891-904.

75. T. Yamaguchi, K. Hidaka and A. K. Soper, *Mol. Phys.*, 1999, **96**, 1159-1168.

76. S. Schoenholz and E. D. Cubuk, *Advances in Neural Information Processing Systems*, 2020, **33**, 11428-11441.

77. J. Xia and B. Jiang, *arXiv*, 2025, arXiv:2505.06711.

78. L. B. Skinner, C. J. Benmore, J. C. Neuefeind and J. B. Parise, *J. Chem. Phys.*, 2014, **141**, 214507.

View Article Online
DOI: 10.1039/D5SC07248D

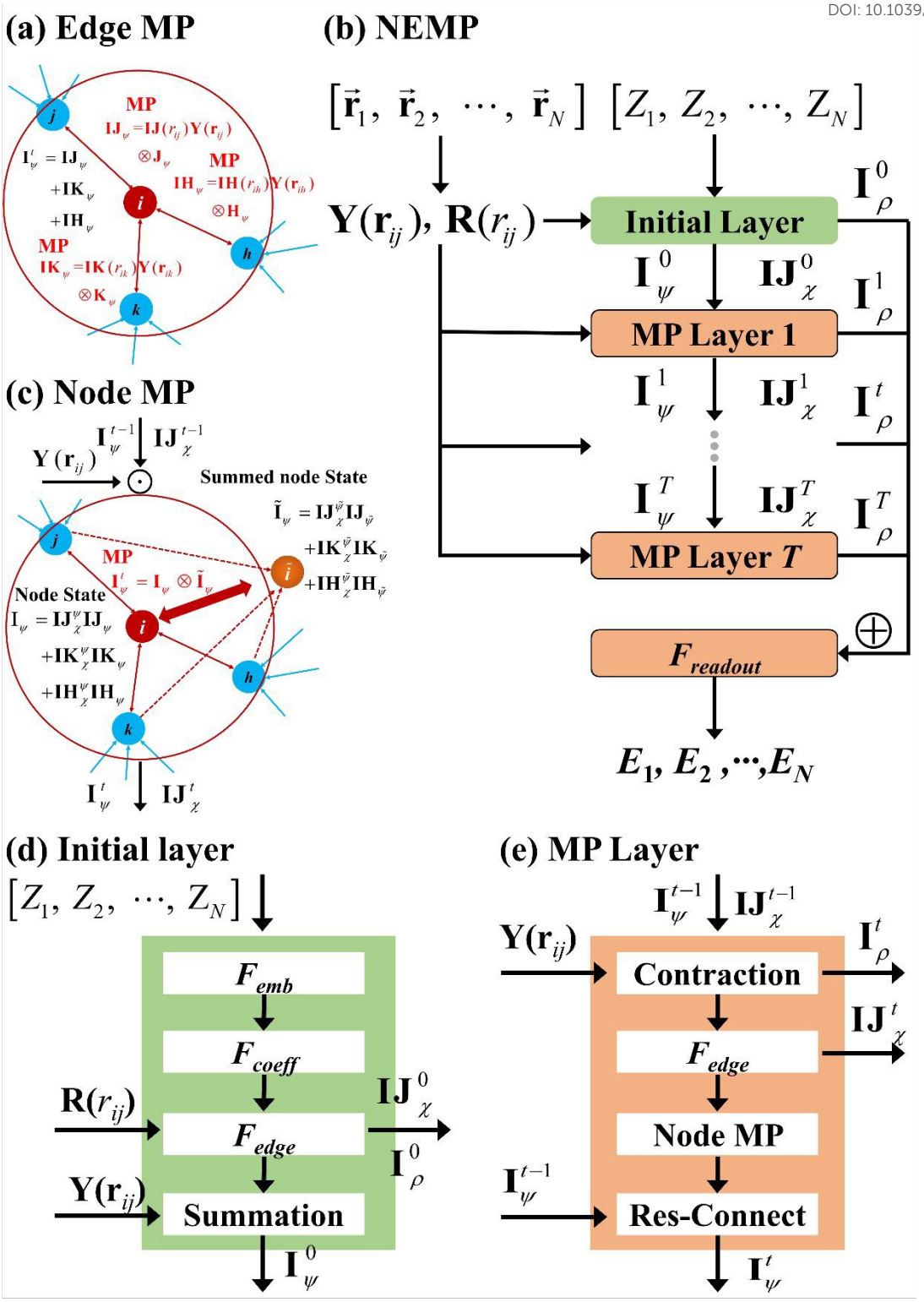

Fig. 1. Schematic illustrations of (a) the MP step for the EEMP, (b) the NEMP module, and (c) the overall NEMP architecture including the initial layer and MP layer. Panels (d) and (e) respectively present the initial layer and the MP layer, which contains the

View Article Online
DOI: 10.1039/D5SC07248D

node MP step whose detailed structure is shown in panel (b). Here, $\otimes$ denotes the tensor product; 'i' represents the central node, and 'j, k, h' denote the neighboring nodes. "Res-connect" refers to the residual connection applied to the node equivariant features, as defined in Eq. (13). "Contraction" refers to the computation of node and edge invariant features via tensor contraction, as described in Eq. (7) and (8).

View Article Online
DOI: 10.1039/D5SC07248D

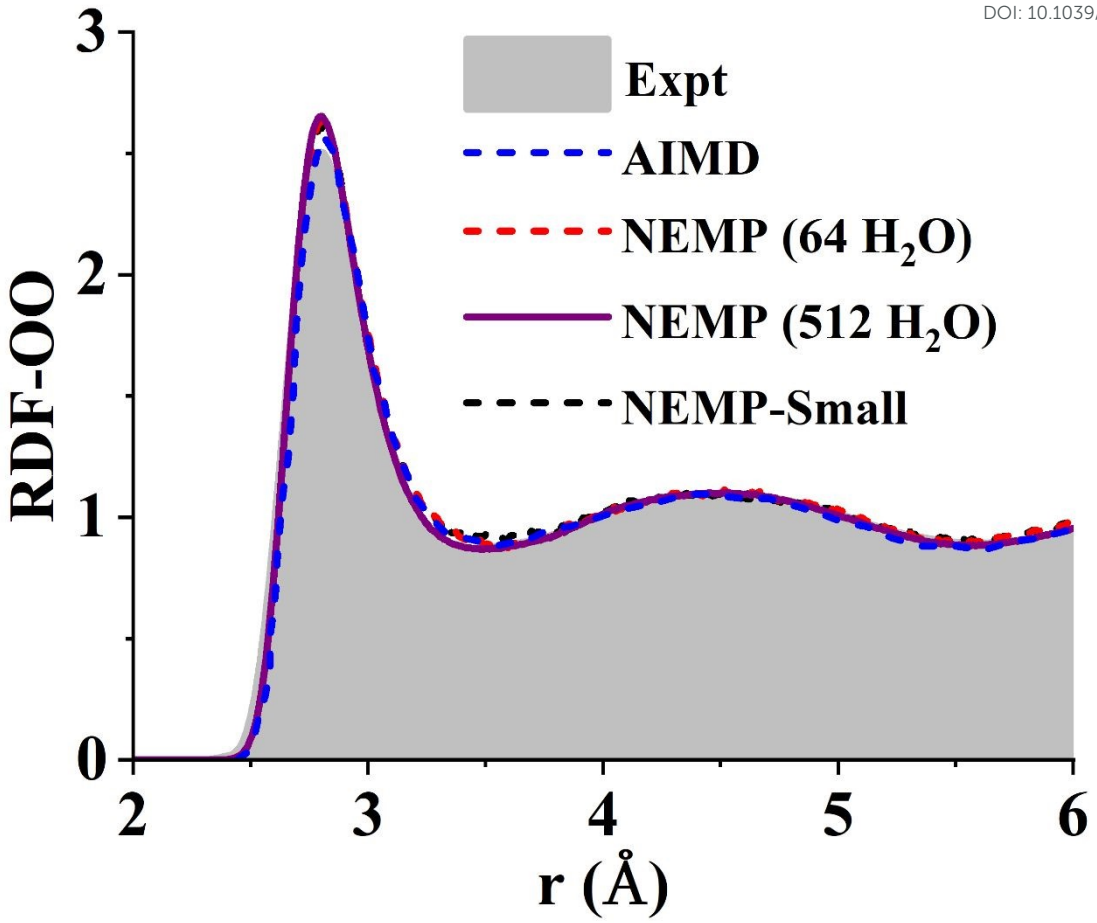

Fig. 2. Comparison of experimental[78] and theoretical O–O radial distribution functions of liquid water at 300 K. The theoretical results are obtained from DFT-based ab initio molecular dynamics[63] (AIMD) and MD simulations using the NEMP and NEMP-Small potentials for system sizes of 64 and 512 H_2O molecules.

View Article Online
DOI: 10.1039/D5SC07248D

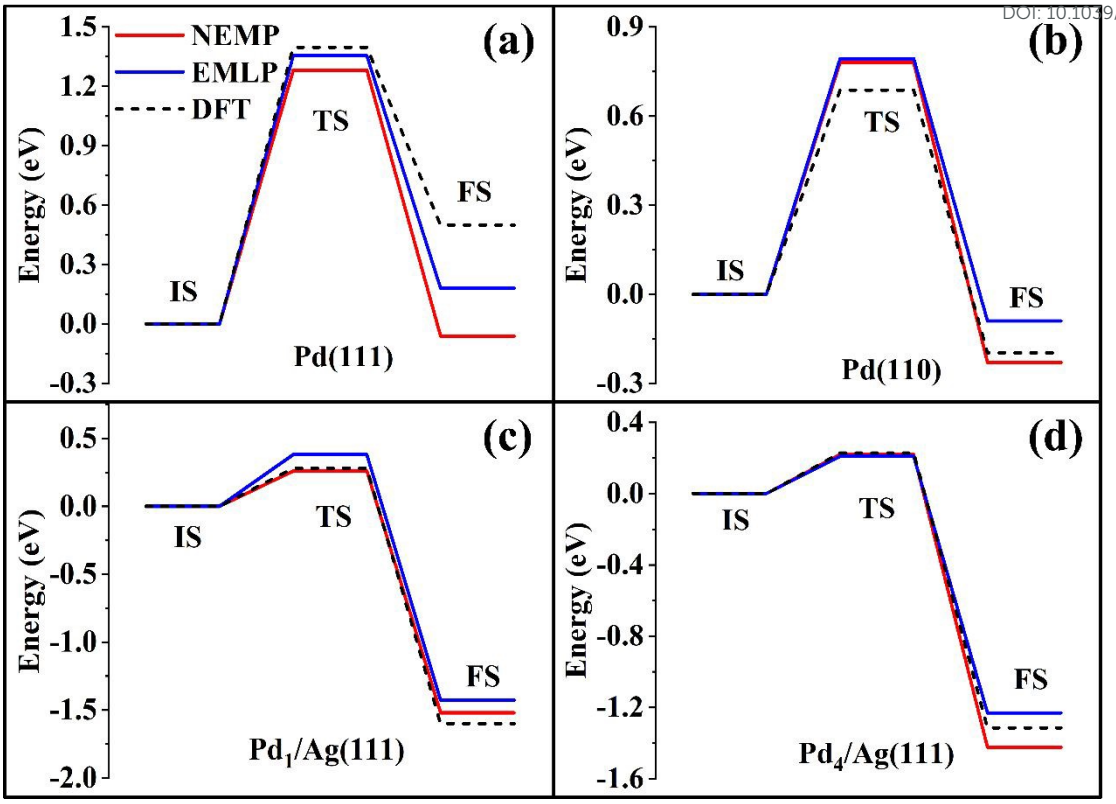

Fig. 3. CO oxidation reaction path on (a) Pd(111), (b) Pd(110), (c) Pd(111), and (d) Pd_4/Ag(111), comparing DFT calculations with EMLP/NEMP predictions.

View Article Online
DOI: 10.1039/D5SC07248D

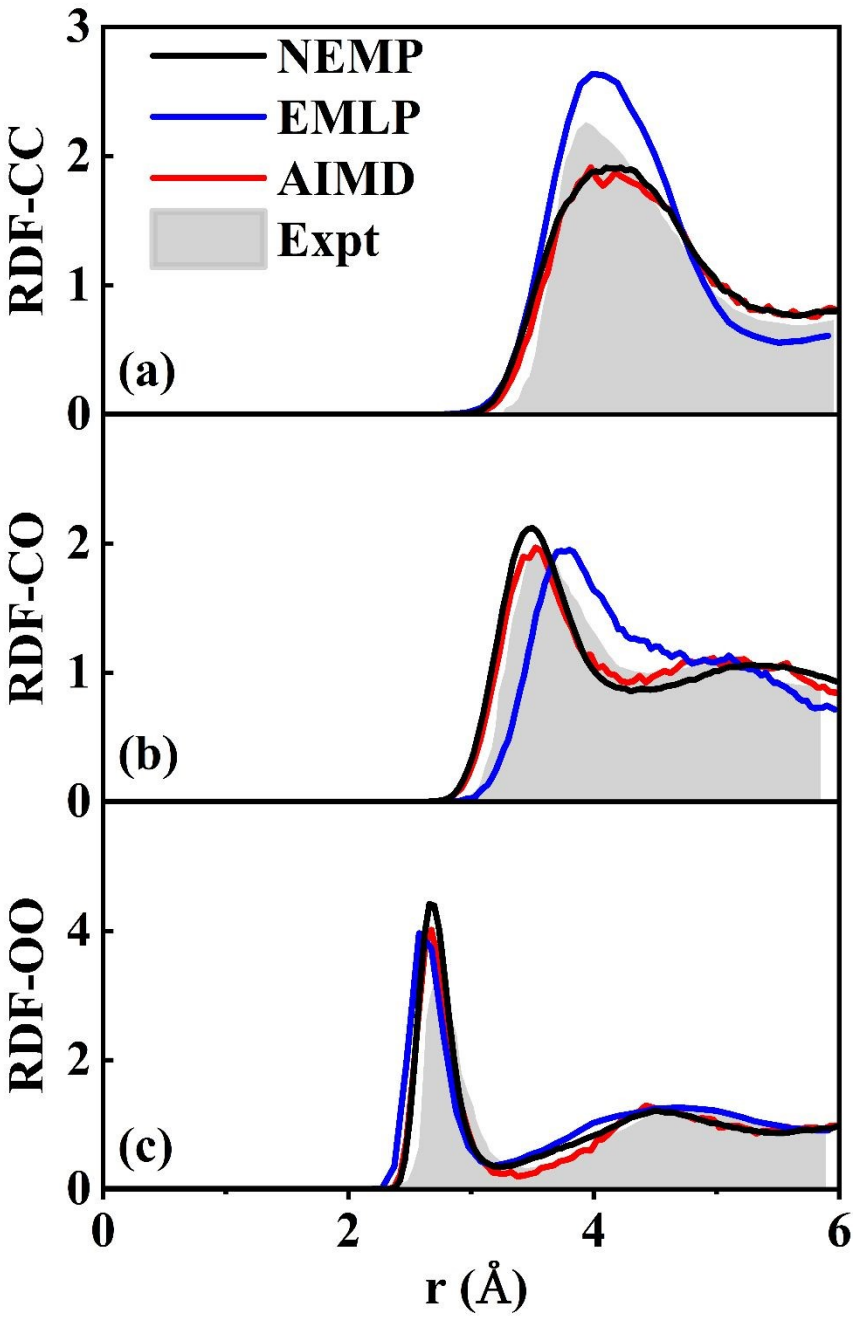

Fig. 4. Comparison of experimental[75] and theoretical RDFs of liquid methanol at 300 K for (a) C – C, (b) C – O, and (c) O – O pairs, obtained using DFT, and EMLP, and NEMP potentials.

View Article Online
DOI: 10.1039/D5SC07248D

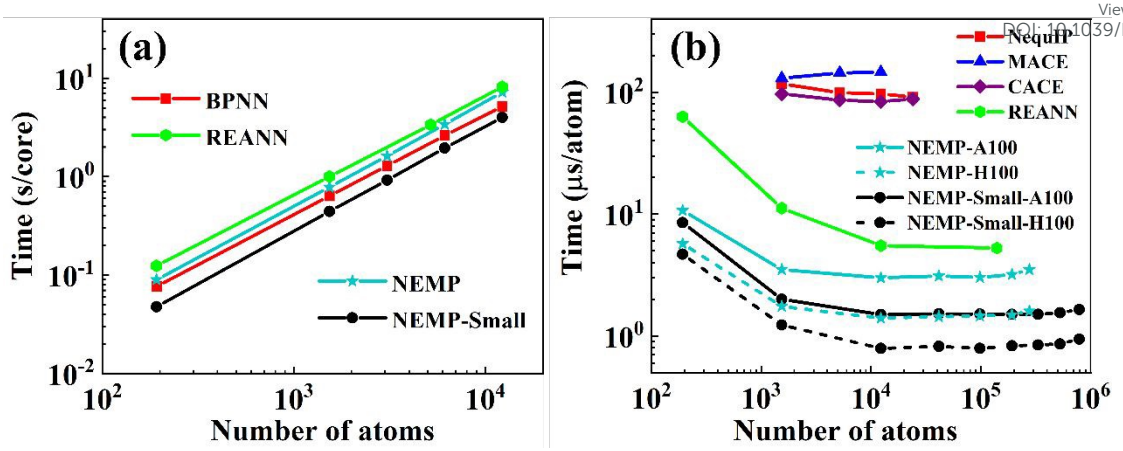

Fig. 5. Computational cost per MD step in liquid water simulations for: (a) NEMP and NEMP-Small, evaluated on a single core of Intel® Xeon Gold 6438Y CPUs and, compared with BPNN and REANN, whose results are taken from Refs. [57, 62], were computed on a single core of Intel® Xeon 6132 CPU; and (b) NEMP, NEMP-Small (A100/H100 GPUs) versus REANN, CACE, NequIP, and MACE (A100 GPUs). Note that all calculations of the NEMP model are conducted using float32 precision.

View Article Online
DOI: 10.1039/D5SC07248D

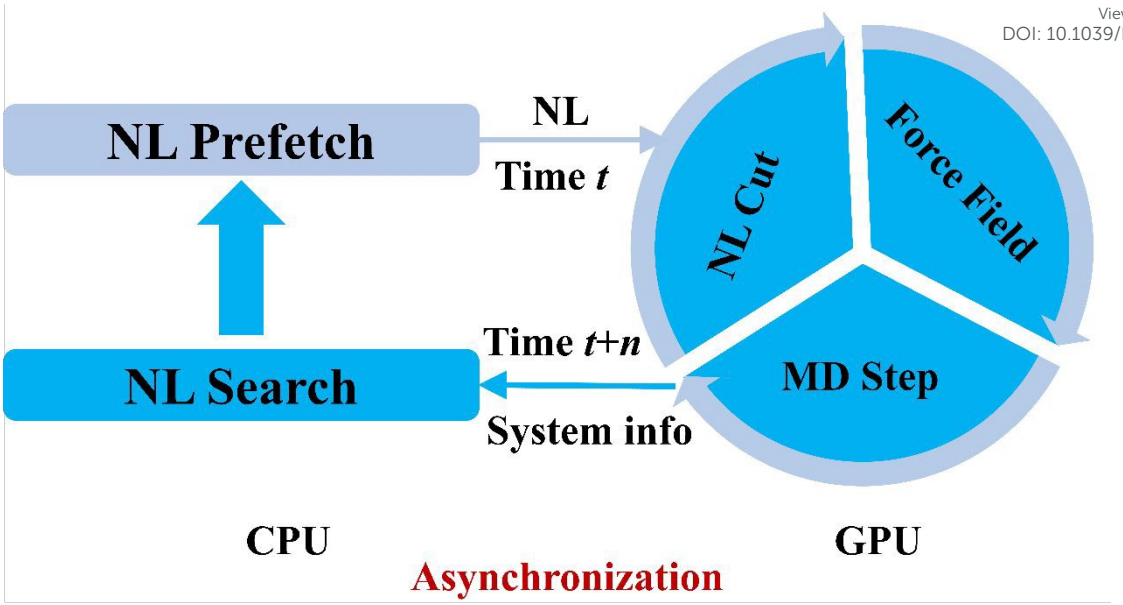

Fig. 6. Schematic of the asynchronous computation workflow showing NL generation on CPU concurrent with MD propagation on GPU.

View Article Online
DOI: 10.1039/D5SC07248D

Table I: Prediction errors (RMSEs) for energy (meV) and atomic forces (meV/Å) of the 3BPA molecule across different MLIPs. The boldfaced number indicates the model with the lowest error. 'E' and 'F' in the first column denote energy and forces, respectively. The numbers in parentheses represent the error bars obtained from four independent training runs with different initial parameters.

Dataset	ACE	sGDML	ANI-2x	NequIP	CACE	MACE	NEMP
300 K, E	7.1	9.1	38.6	3.3	6.3	**3.0**	3.4 (0.17)
300 K, F	27.1	46.2	84.4	11.3	21.4	**8.8**	10.7 (0.18)
600 K, E	24.0	484.8	54.5	11.2	18.0	**9.7**	**9.7** (0.23)
600 K, F	64.3	439.2	102.8	27.3	45.2	**21.8**	24.4 (0.26)
1200 K, E	85.3	774.5	88.8	40.8	58.0	**29.8**	32.5 (0.13)
1200 K, F	187.0	711.1	139.6	86.4	113.8	**62.0**	69.2 (1.1)
Time	\	\	\	103.5	\	24.5	**2.4**

View Article Online
DOI: 10.1039/D5SC07248D

Table II: RMSEs for energy per water molecule (meV/H_2O) and atomic forces (meV/Å) across different MLIPs for liquid water. The boldfaced number indicates the model with the lowest error. The numbers in parentheses represent the error bars obtained from four independent training runs with different initial parameters.

	BPNN	EANN	REANN	NequIP	CACE	MACE	NEMP
Energy	7.0	6.3	2.0	2.8	**1.8**	1.9	2.4 (0.1)
Forces	120.0	129.0	47.0	45.0	47.0	**36.2**	36.8 (0.5)

View Article Online
DOI: 10.1039/D5SC07248D

Table III: Performance comparison of MAEs for total energies (kcal/mol) and atomic forces (kcal/mol/Å) across different MLIPs for the COMP6 benchmark dataset. Note that the ANI-1x model utilized 10× more training data than other models. The boldfaced numbers indicate the model with the lowest error. 'E' and 'F' in the second column denote energy and forces, respectively. The numbers in parentheses represent the error bars obtained from four independent training runs with different initial parameters.

Dataset		ANI-1x	TrIP	TensorNet	MACE	**NEMP**
ANI-MD	E	3.4	\	1.61	3.25	**1.42 (0.098)**
	F	2.68	\	0.82	0.62	**0.42 (0.012)**
DrugBank	E	2.65	\	0.98	0.73	**0.57* (0.051)**
	F	2.86	\	0.75	0.47	**0.30* (0.011)**
GDB 7-9	E	1.04	\	0.32	0.21	**0.20 (0.012)**
	F	2.43	\	0.53	0.34	**0.22 (0.010)**
GDB 10–13	E	2.3	\	0.83	0.53	**0.43 (0.016)**
	F	2.67	\	1.52	0.62	**0.41 (0.016)**
S66x8	E	2.06	\	0.62	0.39	**0.34 (0.028)**
	F	1.6	\	0.33	0.22	**0.16 (0.009)**
Tripeptides	E	2.92	\	0.92	0.79	**0.62 (0.022)**
	F	2.49	\	0.62	0.44	**0.31 (0.011)**
COMP6 total	E	1.93	1.04	\	0.48	**0.39 (0.010)**
	F	2.09	1.41	\	0.52	**0.34 (0.010)**

View Article Online
DOI: 10.1039/D5SC07248D

* Note: ~150 DrugBank test points with errors >100 kcal/mol were excluded from test error calculation.

Table IV: MAEs for Energy per atom (meV/atom) and Force (meV/Å) of several MLIPs on the HME21 dataset. The boldfaced number indicates the model with the lowest error. The numbers in parentheses represent the error bars obtained from four independent training runs with different initial parameters.

	TeaNet	NequIP	MACE	NEMP
Energy	19.6	47.8	**16.5**	**16.5 (0.20)**
Forces	175	199	**140.2**	144.3 (0.18)

View Article Online
DOI: 10.1039/D5SC07248D

Data availability:

The NEMP Package is available from https://github.com/zhangylch/NEMP. All datasets used for model training and validation in this study are publicly accessible from previously published sources.