nature communications

Article

https://doi.org/10.1038/s41467-024-51886-6

E(n)-Equivariant cartesian tensor message passing interatomic potential

Received: 16 March 2024

Accepted: 16 August 2024

Published online: 01 September 2024

Check for updates

Junjie Wang[1,3], Yong Wang[1,2,3], Haoting Zhang[1], Ziyang Yang[1], Zhixin Liang[1], Jiuyang Shi[1], Hui-Tian Wang[1], Dingyu Xing[1] & Jian Sun[1]

Machine learning potential (MLP) has been a popular topic in recent years for its capability to replace expensive first-principles calculations in some large systems. Meanwhile, message passing networks have gained significant attention due to their remarkable accuracy, and a wave of message passing networks based on Cartesian coordinates has emerged. However, the information of the node in these models is usually limited to scalars, and vectors. In this work, we propose High-order Tensor message Passing interatomic Potential (HotPP), an E(n) equivariant message passing neural network that extends the node embedding and message to an arbitrary order tensor. By performing some basic equivariant operations, high order tensors can be coupled very simply and thus the model can make direct predictions of high-order tensors such as dipole moments and polarizabilities without any modifications. The tests in several datasets show that HotPP not only achieves high accuracy in predicting target properties, but also successfully performs tasks such as calculating phonon spectra, infrared spectra, and Raman spectra, demonstrating its potential as a tool for future research.

Molecular dynamics (MD) is a powerful computational technique allowing for the exploration of various physical and chemical phenomena at the atomic level and the study of the behavior of molecules and materials over time. It bridges the gap between theoretical predictions and experimental observations, enabling researchers to gain a comprehensive understanding of the behavior, properties, and interactions of molecules and materials. With sufficient computational resources, first principles calculations based on Density Functional Theory (DFT)[1] can simulate systems with hundreds or even thousands of atoms. However, it struggles when it comes to larger systems. Another approach to computing atomic interactions is empirical force fields, providing much quicker calculations and the ability to handle significantly larger systems. Nevertheless, many of these force fields rely on empirical observations, limiting their applicability to specific ranges and lacking universality and transferability. The machine learning potential (MLP)[2–8], which aim to accurately describe the potential energy surface of atomic configurations, combines the advantages of both DFT and empirical force fields. A well-trained machine learning force field can achieve accuracy close to DFT and even beyond DFT[9,10], and perform very large-scale, long-time simulations, offering a glimpse into the future of research in studying complex dynamical problems.

Most existing machine learning potentials are based on the framework proposed by Behler[2], which fits the total energy as a sum of atomic energies $E = \sum E_i$, and the atomic energies are determined by the atomic environment within a certain cutoff radius. This format ensures the scalability of the potential, allowing the network to be trained on small systems and extrapolated to larger systems. The quality of such a model is highly dependent on the choice of descriptors that describe the atomic environments[11]. A reasonable descriptor should first possess invariance to rotations, translations, and atom permutation; thus, the same atomic environment yields the same atomic energy. A common approach is to construct a series of symmetric functions based on interatomic distances and angles

[1]National Laboratory of Solid State Microstructures, School of Physics and Collaborative Innovation Center of Advanced Microstructures, Nanjing University, Nanjing 210093, China. [2]Department of Chemistry, Princeton University, Princeton, NJ 08544, USA. [3]These authors contributed equally: Junjie Wang, Yong Wang. e-mail: jiansun@nju.edu.cn

between atoms since these two quantities are naturally invariant under rotations and translations. Depending on the number of atoms involved, a series of so-called two-body and three-body descriptors can be obtained, such as atom-centered symmetry functions (ACSF)[2,12,13], the NEP descriptor[7,14–16], the smooth overlap of atomic positions (SOAP)[17], the DeePMD descriptor[5,18]. However, these descriptors are not complete[19], as different atomic environments can yield the same descriptor. Atomic cluster expansion (ACE)[6,20] and Moment Tensor Potential (MTP)[4,21] have proposed complete descriptors that can account for interactions of arbitrary order, but the number of descriptors can easily grow to tens of thousands as the order increases. Another issue is that such descriptors are only dependent on the coordinate information within the cutoff radius. When dealing with long-range interactions, simply increasing the cutoff radius would significantly raise computational complexity since the number of atoms is proportional to the cube of cutoff radius.

Message passing network (MPN)[22] can help address both of these issues. In the context of MLP, MPN is used to represent molecule or crystal structures as graphs, where atoms are nodes and bonds are edges. The key idea behind MPN is the iterative passing of messages between nodes, allowing information to be exchanged and aggregated. Such message passing processes can, on one hand, lead to the emergence of multiple atoms in the final descriptor (thus, resulting in n-body symmetric functions). On the other hand, it allows information from atoms beyond the cutoff radius to be transmitted to the current atom. As a result, many machine learning potentials based on message passing networks have achieved high levels of accuracy[8,23–31]. It is worth noting that as long as the energy (or other properties) obtained at the end satisfies the symmetry requirements, the messages used in the network do not necessarily have to be scalars[24,32]. For example, NequIP[8], BotNet[30], and MACE[31] utilize high-order tensors based on spherical harmonics in the message passing, coupling them through Clebsch-Gordon (CG) coefficients to construct equivariant networks. These methods have shown significant improvements in accuracy compared to approaches that only use scalar messages. Another category of methods, including PaiNN[26] and torchMD-Net[28], directly utilize vectors in Cartesian space as messages and obtain equivariant results through a series of designed layers. This approach does not require coupling through CG coefficients, but the vectors are only equivalent to $l=1$ tensors in the spherical harmonics method. TeaNet[33] can pass matrices information equivalent to $l=2$ tensors, but this introduces a multitude of artificially designed duplications, resulting in a highly intricate network structure that becomes challenging to extend to higher order tensors. Machine learning potential that use tensors of arbitrary orders as messages based on Cartesian coordinates have not been proposed.

In this work, we propose High-order Tensor Passing Potential (HotPP), which can utilize arbitrary order Cartesian tensors as messages. By combining some basic equivariant operations between tensors, all the high-order tensors used in the network are E(n)-equivariant, thus the output is consistent with the rotation of coordinates. In other words, if the output is a scalar, it remains invariant under rotations, while the vector output will rotate in accordance with the rotation of the coordinates, and the matrix output will transform as $M'=RMR^T$. Therefore, the method can directly predict high-order tensors such as dipole moments and polarizability tensors without any modifications. We validate HotPP in three prediction tasks: the energies and forces of molecular dynamics trajectory of small molecule; the energies, forces, and stresses of carbon with periodic boundary conditions; and the dipole moments and polarizability tensors of small molecules with coupled cluster singles and doubles (CCSD) accuracy. In these tests, our model achieves good performance with fewer parameters comparable to other high-order models, which provide a frame of equivariant network based on Cartesian coordinates.

Results

Equivariant functions of Cartesian tensors

Cartesian tensors are the tensors that transform under rotations in Euclidean space in a simple way. In other words, a Cartesian tensor is a tensor whose components transform as a product of vectors and covectors under rotations, without any additional factors that depend on the rotation matrix itself. Specifically, a n-th rank tensor transforms under rotation as:

$$\boldsymbol{T}_{i_1 i_2 \cdots i_n} \xrightarrow{\boldsymbol{R}} \boldsymbol{T}'_{i_1 i_2 \cdots i_n} = \left(\boldsymbol{R}_{i_1 j_1}\right)\left(\boldsymbol{R}_{i_2 j_2}\right)\cdots\left(\boldsymbol{R}_{i_n j_n}\right)\boldsymbol{T}_{j_1 j_2 \cdots j_n} \tag{1}$$

where R is an orthogonal matrix. Under this definition, it is easy to find that since $\mathbf{v}_i \to \mathbf{v}'_i = \boldsymbol{R}_{ij}\mathbf{v}_j$, the vectors are first-order tensors, and the dyadic product of two vectors is a second-order tensor since $(\mathbf{uv})_{i_1 i_2} \xrightarrow{\boldsymbol{R}} (\mathbf{uv})'_{i_1 i_2} = (\boldsymbol{R}_{i_1 j_1})(\boldsymbol{R}_{i_2 j_2})(\mathbf{uv})_{j_1 j_2}$.

And equivariance is a property of functions or transformations between two spaces, where the transformation preserves the relationships between the elements of those spaces. More formally, a function $\phi : \mathbf{X} \to \mathbf{Y}$ is said to be equivariant with respect to a group G acting on two sets $\mathbf{X}$ and $\mathbf{Y}$ if for all $g \in \mathbf{G}$ and $x \in \mathbf{X}$, we have:

$$\phi\left(g \circ x\right) = g \circ \phi(x) \tag{2}$$

This means that applying a function ϕ to an object x and then applying a group element g to the resulting object should give the same result as first applying the group element to the object and then applying the function. And the composition of equivariant maps is also equivariant:

$$\psi\left(\phi\left(g \cdot x\right)\right) = \psi\left(g \cdot \phi(x)\right) = g \cdot \psi(\phi(x)) \tag{3}$$

Therefore, by providing some basic equivariant functions between Cartesian tensors and combining them, we can obtain an equivariant neural network. Here, we use the following three equivariant operations, whose equivariances are proven in the Supplementary Note 1:

1. Linear combinations of tensors with the same order: $f(T_1, T_2, \cdots, T_m) = \sum c_i T_i$.
2. Contraction of two tensors.

The contraction of tensors is a mathematical operation that reduces the rank of tensors by summing over one or more pairs of indices. For example, consider a 3-order tensor $\boldsymbol{A}$ and a 2-order tensor $\boldsymbol{B}$, the contraction of them can be $\boldsymbol{C}_{ijk} = \boldsymbol{A}_{ijl}\boldsymbol{B}_{kl}$, this will reduce the rank of tensors by 2. More generally, if we sum over more than one pair of indices between an x-order tensor $\boldsymbol{T}_1$ and a y-order tensor $\boldsymbol{T}_2$ such as:

$$\Phi_z\left(\boldsymbol{T}_1, \boldsymbol{T}_2\right)_{a_1 \cdots a_{x-z} b_1 \cdots b_{y-z}} = \boldsymbol{T}_{1 a_1 \cdots a_{x-z} c_1 \cdots c_z} \cdot \boldsymbol{T}_{2 b_1 \cdots b_{y-z} c_1 \cdots c_z} \tag{4}$$

We can get a new tensor with $x+y-2z$ order, where $0 \le z \le \min(x,y)$. When $z=0$, none of the indices are contracted and the Eq. (4) becomes tensor product $\boldsymbol{T}_{a_1 \cdots a_x b_1 \cdots b_y} = \boldsymbol{T}_{1 a_1 \cdots a_x} \cdot \boldsymbol{T}_{2 b_1 \cdots b_y}$.

3. Partial derivative with respect to another Cartesian tensor: $\frac{\partial}{\partial \boldsymbol{T}_{j_1 j_2 \cdots j_n}}$.

By the combination of these operations, we can get many equivariant functions. Many common operations frequently used in other equivariant neural networks that operate on vectors, such as scaling of vectors: $s \cdot \vec{v}$, scalar products $\langle \vec{v}_1, \vec{v}_2 \rangle$, vector products $\vec{v}_1 \times \vec{v}_2$ (the upper triangle part of $\vec{v}_1 \otimes \vec{v}_2 - \vec{v}_2 \otimes \vec{v}_1$) can all be viewed as special cases of these three operations. Some other more complex descriptors, such as MTP descriptors[4,21], can also be obtained through combinations of these operations as shown in the Supplementary Note 3.

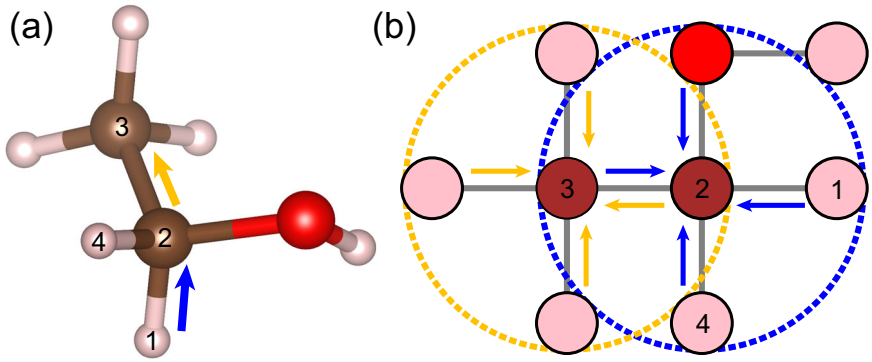

Fig. 1 | The schematic diagram of message-passing networks. a the origin structure; **b** the graph corresponding (**a**). The blue arrows represent the first time of message passing of the atom 2, while the yellow arrows denote the second time of the atom 3.

Equivariant message passing neural network

To obtain an end-to-end machine learning model for predicting material properties, the input should be the positions $\{\mathbf{r}_i\}$ and chemical elements $\{Z_i\}$ of all atoms, and for periodic crystals, the lattice parameters should also be considered. To apply graph neural networks, we first transform the crystal structure into a graph $\{n_i, e_{ij}\}$, where each node n_i corresponds to an atom i in the unit cell, and all atoms j within a given cutoff distance r_{cut} are considered connected to n_i labeled with their relative positions $\mathbf{r}_{ij}$. For periodic structures, since atom j and its equivalent atom j' may both lie within the cutoff distance of atom i, n_i may have more than one edge connected to n_j. To extract the information of the nodes, we use the scheme of MPN. A normal MPN can be described as:

$$m_i^{t+1} = \bigoplus_{j \in N(i)} \mathrm{M}_t\left(h_i^t, h_j^t, e_{ij}\right) \tag{5}$$

$$h_i^{t+1} = \mathrm{U}_t\left(h_i^t, m_i^{t+1}\right) \tag{6}$$

where h_i^t is the hidden feature of n_i at layer t that captures its local information, messages are then passed between nodes along edges, with the message at each edge e_{ij} being a function M_t of the features of the nodes connected by that edge. The $\bigoplus$ is a differentiable and permutation invariant function such as sum, mean or max to aggregate the message at each node together to produce an updated message m_i^{t+1} for that node, which in turn is used to update the hidden feature with the function U_t for the next iteration.

A concrete example illustrating the principles of MPN is presented in Fig. 1. We first determine the connectivity of a structure based on a given cutoff radius and convert it into a graph (quotient graph for periodic structures). Then the messages on the nodes can be passed through the edges by a two-body interaction. As the process of message passing, the information from atoms beyond the cutoff radius can also be conveyed to the central atom. As illustrated in Fig. 1b, the blue arrows represent the first time of message passing, while the yellow arrows denote the second. To be noticed, each layer of message passing is performed simultaneously on all atoms; here, we focus on a specific atom in each layer for ease of analysis. During the initial time of message passing, information of atom *1* is encoded into the hidden information of atom *2*. Subsequently, in the second time of message passing, the information of atom *2* including some information of atom *1* is collectively transmitted to atom *3*, thereby achieving nonlocal effects from atom *1* to atom *3*. On the other hand, due to the interaction between atom *4* and atom *2* in the second time of message passing also containing the information from atom *1*, the effective interaction is elevated from a two-body interaction to a three-body interaction. This indeed encapsulates the two advantages of the message-passing architecture.

However, the scalar hidden feature, message, and edge information (always relatively distance between two atoms) here will limit the expressive capacity and may cause the incompleteness of atomic structure representations. As shown in Fig. 2a and d, if we only use scalar information h_i, h_j, and d_{ij} to pass the message in Eq. (5) and update the feature in Eq. (6), all nodes will always produce the same embedding information. As a result, the network will be unable to distinguish between these two structures and give the same total energy. Even if the 3-body message includes angles α_{ijk} are taken into consideration, some structures with only 4 atoms cannot be distinguished[26], as shown in Fig. 2b and e. Due to the identical atomic environments within the truncation radius, no matter how many message passing iterations are performed, these two different structures will only yield the same result. To alleviate this problem, a series of models that use high-order geometric tensors during the message passing have been proposed. For example, allowing vectors in the message passing process can differentiate Fig. 2b and e, but in the case of Fig. 2c and f which have different α, the summation in Eq. (5) would cause the network to confuse these two structures (a more detailed explanation can be seen in the Supplementary Note 5). It can be anticipated that increasing the order of tensors in message passing would enhance the expressive power of the network. Previously, the order of high-order tensor networks based on Cartesian space was typically limited to 2[33,34], while our method can work with any order Cartesian tensors. In the following, we use ${}^{l}\boldsymbol{h}_i^t$ to represent the l-order Cartesian tensor features of node i in the t-th layer, and $\mathbf{r}^{\otimes n}$ to represent tensor product of a vector $\mathbf{r}$ for n times: $\mathbf{r}\otimes\mathbf{r}\otimes\cdots\otimes\mathbf{r}$. In particular, for n = 0 we define this to a learnable function of $||\mathbf{r}||$: $\mathbf{r}^{\otimes 0} = \mathrm{f}(||\mathbf{r}||)$.

Initialize of node features. The scalar features in the first layer ${}^{0}h_i^0$ should be invariant to rotation, translation, and permutation of the atoms with the same chemical species. This is also the requirement for most descriptors used in machine learning potentials, so these descriptors such as ACSF, SOAP, ACE, MTP, etc., can be used directly to expedite the process of feature extraction. Here, we used the trainable chemical embedding similar to SchNet[23] to minimize human-designed elements as much as possible. Specifically, the atomic numbers are first encoded by one-hot and then multiplied by a learnable weight matrix, resulting in a learnable embedding for each element Z_i. For high-order features ${}^{l}\boldsymbol{h}_i^0$ with $l > 0$, we set them all to 0 at the beginning.

Message and aggregate. To combine the information of neighboring nodes, we need to design a message passing function M_t in Eq. (5). Considering that the hidden feature ${}^{l_i}\boldsymbol{h}_i^t$, ${}^{l_j}\boldsymbol{h}_j^t$, the bond info e_{ij}, and the target message ${}^{l_{\text{out}}}\boldsymbol{m}_{ij}^{t+1}$ can be tensors of arbitrary order. Therefore, we need to find an equivariant way to compose the two tensors to a new

https://doi.org/10.1038/s41467-024-51886-6

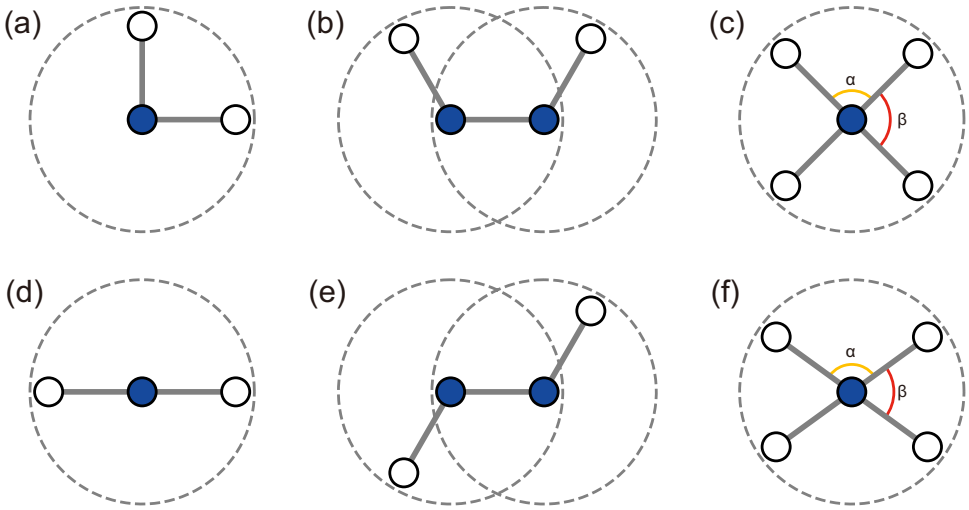

Fig. 2 | Some structures that cannot be distinguished by message passing networks that do not utilize high-order tensor information. a, d cannot be distinguished by using two-body scalar information; **b**, **e** cannot be distinguished by using three-body scalar information[26]; and **c**, **f** cannot be distinguished by using two-body vector information. **c**, **f** have different α and β, so they have the same center of mass but different moment of inertia.

tensor with different order, and Eq. (4) is such an operation. In our model, we write M_t in Eq. (5) as:

$$ {}^{l_o}_{l_i,l_r}\boldsymbol{m}_{ij}^t = \mathrm{M}_t\left(\boldsymbol{h}_i^t,\boldsymbol{h}_j^t,\boldsymbol{e}_{ij}\right) = \mathrm{f}_{l_r}^t\left(d_{ij}\right)\cdot {}^{l_i}\boldsymbol{h}^t_{j_{a_1\cdots a_{l_i-l_c}c_1\cdots c_{l_c}}}\cdot\left(\mathbf{u}_{ij}^{\otimes l_r}\right)_{c_1\cdots c_{l_c}b_1\cdots b_{l_r-l_c}} \tag{7} $$

Where $d_{ij}=||r_{ij}||$ is the relative distance between atom i and atom j, $\mathbf{u}_{ij}=\frac{\mathbf{r}_{ij}}{d_{ij}}$ is the direction vector, $0\le l_c\le \min(l_i,l_r)$ is the number of the indices summing up during the contraction. $\mathrm{f}_{l_r}^t\left(d_{ij}\right)$ is the radial function, which is a learnable multi-layer perceptron of radial basis functions such as Bessel basis and Chebyshev basis. The result is a Cartesian tensor with order $l_o=|l_i+l_r-2l_c|$, which is between $|l_i-l_r|$ and l_i+l_r. Since l_r can be chosen arbitrarily, we can obtain an equivariant tensor of order from 0 to any arbitrary order.

We use a summation operation as the aggregation function for the messages in Eq. (5), that is, directly adding all the messages obtained from neighboring nodes. For tensors of the same order obtained from different (l_i,l_r), we add them together with different coefficients. Due to the arbitrariness of l_o and l_r, we need to specify their maximum values. With given l^t_{omax}, l^t_{rmax}, we sum up all possible (l_i,l_r):

$$ {}^{l_o}\boldsymbol{m}_i^{t+1} = \sum_{l_r\le l^t_{\mathrm{rmax}}}\sum_{l_i} c_{l_o,l_r,l_i}\sum_{j\in N(i)}\mathrm{f}_{l_r}\left(d_{ij}\right)\cdot {}^{l_i}\boldsymbol{h}^t_{j_{a_1\cdots a_{l_i-l_c}c_1\cdots c_{l_c}}}\cdot\left(\mathbf{u}_{ij}^{\otimes l_r}\right)_{c_1\cdots c_{l_c}b_1\cdots b_{l_r-l_c}} \tag{8} $$

Update. For scalar message ${}^0m_i^{t+1}$, we feed it to a fully connected layer followed by a non-linear activation function to extract the information, and update the hidden feature with residual neural networks:

$$ {}^0h_i^{t+1} = {}^0h_i^t + \sigma\left({}^0\mathrm{W}^t({}^0m_i^{t+1})\right) \tag{9} $$

Where σ is the nonlinear activation function, ${}^0\mathrm{W}^t$ is a linear layer with bias in the t layer for the scalar message. However, for the tensors above 0 order, both the bias and the activation function will break the equivariance (Supplementary Note 2). Therefore, we only apply bias when $l=0$.

For the high-order activation function, as shown in Eq. (3), tensor multiplication by a scalar is equivariant. Hence, we need to find a mapping from an n-order tensor to a scalar. One simple idea is to use the squared norm of the tensor $||\boldsymbol{T}||^2=\sum_{i_1\cdots i_l}\boldsymbol{T}^2_{i_1\cdots i_l}$ since it is invariant by definition. Therefore, for $l>0$, we write the element-wise non-linear function as:

$$ {}^l\sigma(\boldsymbol{T})_{i_1\cdots i_l} = \sigma'\left({}^l\mathrm{W}(||\boldsymbol{T}||^2)\right)\cdot\boldsymbol{T}_{i_1\cdots i_l} \tag{10} $$

It should be noted that different notations σ and σ' were used for the activation function in Eqs. (9) and (10), as the choice of activation function may vary for scalar and high-order tensors. Take the SiLU function $\sigma(x)=x\cdot\mathrm{sigmoid}(x)$ for example and suppose ${}^l\mathrm{W}$ is the identity function. For scalar, SiLU maps x to x itself when $x\gg 0$. However, for higher-order tensors, Eq. (10) will map $\boldsymbol{T}_{i_1\cdots i_l}$ to $||\boldsymbol{T}||^2\cdot\boldsymbol{T}_{i_1\cdots i_l}$ instead of $\boldsymbol{T}_{i_1\cdots i_l}$ when $\boldsymbol{T}_{i_1\cdots i_l}\gg 0$. This is because if we apply the formula for higher-order tensors to a scalar, which is $y_i=\sigma(\mathrm{W}(x_i))\cdot x_i$, an extra x_i is multiplied. Therefore, if we use SiLU function for σ, we should use Sigmoid function for σ'. Other activation functions can be handled using a similar approach, and hence the update function for high-order tensors is:

$$ {}^l\boldsymbol{h}_i^{t+1} = {}^l\boldsymbol{h}_i^t + {}^l\sigma\left({}^l\mathrm{W}^t({}^l\boldsymbol{m}_i^{t+1})\right) \tag{11} $$

Readout. For a target n-order property, we utilize a two-layer nonlinear MLP to operate on the n-order tensor at the last hidden layer. For the same reason, bias and element-wise nonlinear functions cannot be used in the high-order tensor.

$$ {}^l\boldsymbol{o}_i = {}^l\mathrm{W}_2\left({}^l\sigma\left({}^l\mathrm{W}_1({}^l\boldsymbol{h}_i^t)\right)\right) \tag{12} $$

Performance of HotPP

We validate the accuracy of our method on a diverse range of systems, including small organic molecules, periodic structures, and predictions of dipole moments and polarizability tensors. For each system, we trained the HotPP model on the commonly used dataset and compared the results with other models. To demonstrate the robustness of our model, most models were trained using the same network architecture shown in Fig. 3 and similar hyperparameters. More training details can be seen in Methods section.

Small organic molecule. We first test our model on molecular dynamics trajectories of small organic molecules. The ANI-1x dataset[35,36] contains DFT calculations for approximately five million diverse molecular conformations obtained through an active learning algorithm. To evaluate the extrapolation capability of HotPP, we train our model with 70% data from the ANI-1x dataset and test on the

https://doi.org/10.1038/s41467-024-51886-6

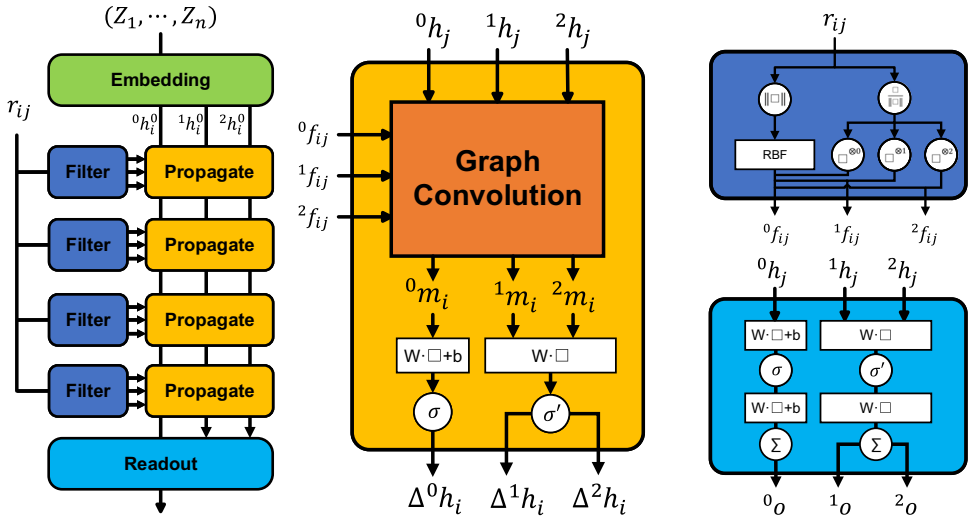

Fig. 3 | The architecture of HotPP. After embedding atomic information into scalars, vectors, and tensors, 4 propagation layers are used to further extract information and outputs the target properties. For higher-order tensors or additional propagation layers, the frame remains similar. The inputs to the network are the atomic numbers Z_i and relative displacements between atoms $\mathbf{r}_{ij}$. After Embedding layer, the scalar node embedding $^0h^0$ are initialized based on Z_i and the higher order embeddings are set to zero. In each Propagate layer, the inputs are the node embeddings $^{l_{in}}h$ and the filter tensors $^{l_f}f_{ij}$ obtained by the Filter layer as $f_{l_r}(d_{ij}) \cdot \mathbf{u}_{ij}^{\otimes l_r}$, where d_{ij} is the relative distance and $\mathbf{u}_{ij}$ is the unit vector of $\mathbf{r}_{ij}$. The output messages $^{l_{out}}m_i$ are transformed through a linear layer and activated by the nonlinear function σ or $\sigma\prime$, dependent on the order of the message tensors. In the Readout layer, the node embeddings are transformed and summed up to get the target properties ^lo, such as ^0o for energy, and ^1o for dipole.

COmprehensive Machine-learning Potential (COMP6) benchmark[35], which samples the chemical space of molecules larger than those included in the training set. The results are shown in Table 1. Compared to ani-1x, our model has demonstrated superior performance across the majority of prediction tasks.

Periodic systems. After testing HotPP on small molecule datasets without periodicity, we evaluated its performance on periodic systems. We selected the carbon system with various phases as the example[37]. It is a complicated dataset with a wide range of structures containing structural snapshots from ab initio MD and iteratively extended from GAP-driven simulations, and randomly distorted unit cells of the crystalline allotropes, diamond, and graphite. We show the results in Table 2. Clearly, our model with about 150k parameters demonstrates advantages in predicting forces and virials compared to most models and can achieve accuracy close to $l = 3$ NequIP model with around 2 million parameters. And when the parameters are expanded to 600k, HotPP achieves the best results on this dataset.

Next, we verify the accuracy of our potential in calculating phonon dispersions of diamond, which was not well-predicted in some of the previous models for carbon[38]. We can obtain the force constant matrix of the structure directly through automatic differentiation $\frac{\partial^2 E}{\partial r_{i\alpha} \partial r_{j\beta}}$. We used Phonopy Python package[39,40] to calculate the phonon spectrum of diamond and compared it with the results from DFT in Fig. 4. The results show that HotPP can describe the vibrational behavior well. Although there are relatively large errors in the high-frequency part at the gamma point, this could be attributed to the inaccuracy of the DFT calculations within the training dataset. We retrained the model using a more accurate dataset[38], and the newly calculated phonon spectrum almost perfectly matches the results from DFT, which demonstrates the reliability of our model.

Table 1 | Test results on COMP6 dataset

		ANI-MD	DrugBank	GDB7-9	GDB10-13	s66x8	Tripeptides
HotPP	E	1155	534	209	332	231	311
	F	61	62	61	92	32	51
Ani-1x[35]	E	2249	597	56	134	568	563
	F	249	174	108	170	136	149

The upper and lower portions of each row in the table correspond to energy (with the legend E) and forces (with the legend F) mean absolute errors in meV and meV/Å, and each column represents to a benchmark set of the COmprehensive Machine-learning Potential (COMP6) benchmarks[35].

Dipole moment and polarizability tensor. Since our model can directly output vectors and matrices, we attempted to directly predict the dipole moments and polarizability tensors of structures in the final section. We consider the water systems including water monomer, water dimer, Zundel cation, and liquid water bulk[41]. The dipole and polarizability of the aperiodic systems were calculated by CCSD theory and those of liquid water were calculated by DFT. Each system contains 1000 structures and we use 70% of them as a training set and the rest as a testing set. We calculate the RMSEs relative to the standard deviation of the testing samples to compare with previous results obtained by other models[41–43] as shown in Table 3 In most cases, HotPP gets the best results except for the polarizability tensor of the water monomer. Compared to T-EANN[42] and REANN[43], HotPP performs particularly well in the case of the dipole moment of liquid water. This may be because they fit the dipole moment by learning q_i and calculating $\boldsymbol{\mu} = \sum_{i=1}^{N} q_i \mathbf{r}_i$, which is inappropriate for periodic systems. In contrast, the output results of our model are all obtained through relative coordinates and thus we can get rid of the selecting of reference point.

Since now we can obtain the dipole moment and polarizability by HotPP, we can calculate the infrared (IR) absorption spectrum and Raman spectrum for liquid water. We separately trained a machine learning potential to learn the energy, forces, and stresses of liquid water to assist us in conducting dynamic simulations. With this potential, we perform a classical MD simulation under ambient conditions (300 K, 1 bar) for 100 ps and calculate the dipole moment and the polarizability tensor every 1 fs. Then we compute the IR and Raman spectra by Fourier transforming the autocorrelation function (ACF) of them and the results are compared to the experiment data[44,45] as shown in Fig. 5. We can observe that both HotPP model and DeePMD

Table 2 | Test results on Carbon

	GAP[7]	DP[7]	MTP[7]	REANN[7]	NEP[7]	NequIP (l = 1)	NequIP (l = 2)	NequIP (l = 3)	HotPP (153k para)	HotPP (602k para)
E	46	44	35	31	42	67	23	17	22	**16**
F	1100	800	630	640	690	746	507	431	439	**395**
W	-	170	200	-	160	-	-	-	64	**56**

E, F, and W are RMSE of energy, forces, and virial in meV/atom, meV/Å, and meV/atom respectively. For MTP, REANN, and NEP, we only show the models that can deliver the best results (se2+se3 for DP, 5 Å cutoff for MTP, 4 Å cutoff for MTP, and 4.2 Å cutoff for NEP). More details of training parameters can be found in Fan's previous work[7].

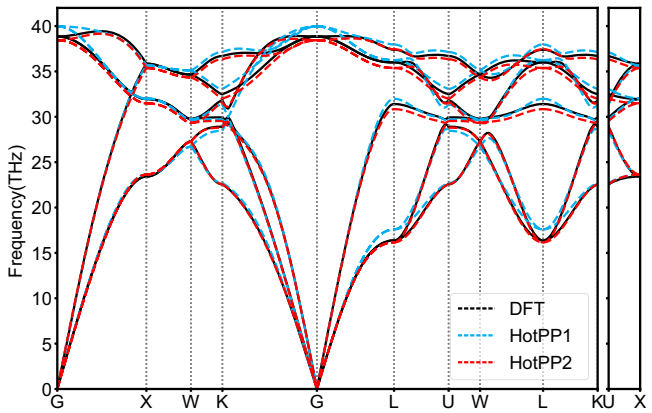

Fig. 4 | Phonon spectrum of diamond. Phonon dispersion relation for diamond as predicted by HotPP1 (blue dotted) and HotPP2 (red dotted) with comparison to DFT reference data (black solid). The HotPP1 was trained with the dataset GAP2017[37], and the HotPP2 was trained with GAP2020[38]. Source data are provided as a Source Data file.

Table 3 | Comparison the results obtained by SA-GPR, T-EANN, REANN and HotPP in different water systems

	SA-GPR[41]	T-EANN[42]	REANN[43]	HotPP
H_2O	~0.11	0.02	0.05	0.02
	~0.02/0.12	0.02	0.06	0.05
$(H_2O)_2$	~5.3	6.6	3.0	2.36
	~6.4/7.8	4.2	1.6	0.99
$H_5O_2^+$	~2.4	1.3	0.4	0.15
	~3.8/0.97	0.3	0.1	0.09
Liquid water	-	16	15	0.70
	~5.8/19	2.2	2.1	0.48

In each system, the first line represents the relative root mean square error (RMSE) of the dipole moment, while the second line indicates the relative RMSE of the polarizability tensor. For SA-GPR, isotropic and anisotropic terms are learned separately, and the values of liquid water are used for the qualitative comparison only as mentioned in the work of REANN[43].

model[46,47] can closely approximate the experimental IR spectra. Our results accurately fit the first three peaks, corresponding to the hindered translation, libration, and H-O-H bending respectively, but there is a long tail in comparison to the experimental data for the O-H stretching mode. This discrepancy may arise from not accounting for quantum effects in our classical molecule dynamic simulation. And for Raman spectra, our model also gives the result in agreement with experimental data.

Discussion

In this work, we introduce HotPP, an E(n) equivariant high-order tensor message passing network based directly on Cartesian space. Compared to other Cartesian-space based high-order tensor networks, HotPP can utilize tensors of arbitrary order, providing enhanced expressive power. In contrast to high-order tensor networks based on spherical harmonics and coupled with CG coefficients, HotPP employs simple tensor contraction operations, resulting in a reduced number of parameters and better computational efficiency for small l (Supplementary Note 6). Moreover, the network's output can be any-order tensor, enabling convenient prediction of vector or tensor properties. With its ability to achieve high accuracy while saving substantial computational time compared to first-principles calculations, HotPP holds great promise in scenarios such as molecular dynamics simulations and structure optimizations, where exploration of potential energy landscapes is essential. In future work, we would investigate approaches to eliminate redundancies in high-order Cartesian tensors to further enhance accuracy. Additionally, due to its E(n) equivariance (rather than E(3)), HotPP can be explored for high-dimensional structure optimization[48] to expedite potential energy surface exploration. It can also serve as a foundation for generating models to directly generate structures or predict wave functions. Overall, HotPP is a promising neural network that we believe can facilitate further explorations in physical chemistry, biology, and related fields.

Methods

Software

All experiments were run with the HotPP software available at https://gitlab.com/bigd4/hotpp, git commit be36dae6c2b35148ba214d5626f9960a8eaf5a07. In addition, the Pytorch version was 2.0.1+cu117, PyTorch Lightning under version 2.0.7, and Python under version 3.9.17.

Datasets

Ani-1x: The ANI-1x dataset[35,36] contains approximately five million diverse molecular conformations obtained through an active learning algorithm. We use 70% data of ani1x.h5 dataset downloaded from https://doi.org/10.6084/m9.figshare.c.4712477.v1 to train the model, and then test on the COmprehensive Machine-learning Potential (COMP6) benchmark[34]. The reference energies are extracted from "wb97x_dz.energy" label, and forces from "wb97x_dz.forces" label.

Carbon: The GAP-17 dataset[37] consists of a training set comprising 4080 structures and a test set comprising 450 structures. The reference data are got by single-point DFT-LDA computations with dense reciprocal-space meshes. The GAP-20 dataset[38] contains 6088 structures and are calculated with the optB88-vdW dispersion inclusive exchange–correlation functional. We use 70% of the data as the training set.

Water: This dataset includes water monomer, water dimer, Zundel cation, and liquid water bulk[41], each systems contains 1000 configurations, 70% of which are used for training. The water monomer, water dimer, and Zundel cation are calculated at (CCSD)/d-aug-cc-pVTZ level, and the liquid water bulk at the DFT/PBE-USPP level.

Training details

The tensor of the out layer l^t_{omax} and the tensor product of the relative coordinates l^t_{rmax} in the models are both set to 2 and a discussion about the effect of these values can be seen in the Supplementary Note 4.

https://doi.org/10.1038/s41467-024-51886-6

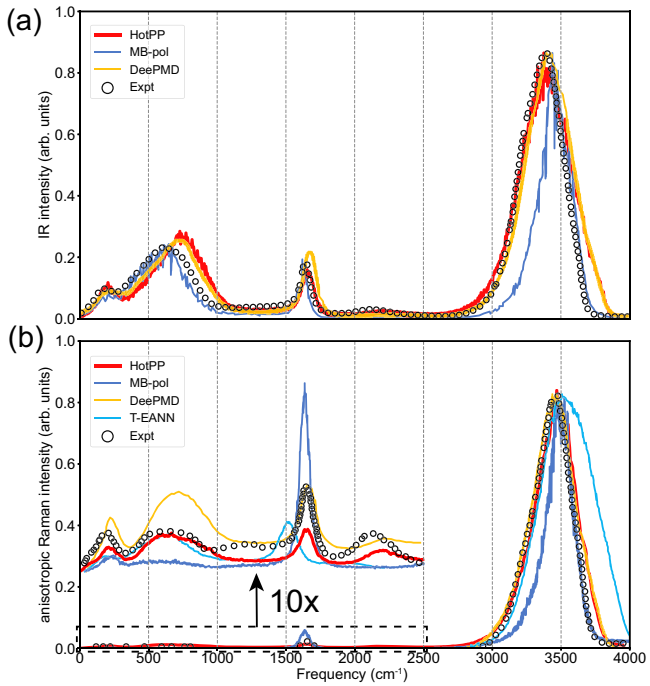

Fig. 5 | Experimental and simulated infrared absorption spectrum and reduced anisotorpic Raman spectra of liquid water under ambient condition. a The comparasion of the infrared abosption spectrum **b** The comparasion of the reduced anisotropic Raman spectra. The black circles are the experimental data[44,45], and the lines are calculated with molecular dynamics trajectories performed by HotPP, MB-pol[56], DeePMD[46,47], and T-EANN[42]. Source data are provided as a Source Data file.

The number of chemical embedding channels and features of node messages is 64. The radial function is a 3 layer MLP of dimensions [64, 64, 64] with SiLU nonlinearity and the basis function is 8 trainable Bessel functions similar to NequIP[8]. The readout layer is a 2 layer MLP of dimensions [64, 64] and SiLU nonlinearity. The models were trained with the Adam optimizer[49] in PyTorch with default parameters. We used a learning rate of 0.01 and the learning rate was reduced using an on-plateau scheduler based on the validation loss with a decay-factor of 0.8.

Ani-1x: We used 5 propagation layers, a radial cutoff of 4.5 Å, a batchsize of 128, and the following loss function with $\lambda_E = 0.1$ and $\lambda_f = 1$:

$$L = \frac{\lambda_E}{N}\left|\left|\hat{E} - E\right|\right|^2 + \frac{\lambda_f}{3N}\sum_{i=1}^{N}\sum_{\alpha=1}^{3}\left|\left|-\frac{\partial \hat{E}}{\partial \mathbf{r}_{i\alpha}} - \mathbf{F}_{i\alpha}\right|\right|^2$$

Where N, E, $\hat{E}$, $\mathbf{F}_{i\alpha}$ denote the number of atoms, target energies, predicted energies, and the force of atom i on direction α.

Carbon: The cell is multiplied by a unit matrix $\boldsymbol{S}$ before calculations. We used 5 propagation layers, the difference between the models with different parameters is that, for the model with more parameters, we updated both node information and edge information in the propagation layers. We used a radial cutoff of 3 Å, a batchsize of 16, and the following loss function with $\lambda_E = 0.5$, $\lambda_f = 1$, and $\lambda_v = 0.2$:

$$L = \frac{\lambda_E}{N}\left|\left|\hat{E} - E\right|\right|^2 + \frac{\lambda_f}{3N}\sum_{i=1}^{N}\sum_{\alpha=1}^{3}\left|\left|\frac{\partial \hat{E}}{\partial r_{i\alpha}} - \mathbf{F}_{i\alpha}\right|\right|^2 + \frac{\lambda_v}{9}\sum_{\alpha,\beta}\left|\left|-\frac{\partial \hat{E}}{\partial \boldsymbol{s}_{\alpha\beta}} - \mathbf{V}_{\alpha\beta}\right|\right|^2$$

Where $\boldsymbol{V}_{\alpha\beta}$, $\boldsymbol{s}_{\alpha\beta}$ denotes the α,β component of the virial and the unit matrix $\boldsymbol{S}$.

The NequIP model was trained on the same dataset with NVIDIA GeForce RTX 4090 24GB. We used 4 layers with 64 channels for even and odd parity for both $l=1$, $l=2$, and $l=3$. Radial features are generated using 8 trainable Bessel basis functions and a polynomial envelope for the cutoff with p = 6, the numbers of invariant layers and neurons were set to 2 and 64. Such hyperparameter settings result in model parameter quantities of 389k, 971k, and 1,970k respectively. Models were trained with Adam optimizer with default parameters. We initialized the learning rate to 0.01 and used an on-plateau scheduler based on the validation loss with a patience of 100 and a decay factor of 0.8. We used an exponential moving average with weight 0.99.

Water: We used 4 propagation layers, a radial cutoff of 4 Å, a batchsize of 4, and the following loss functions:

$$L = \left|\left|\hat{\mathbf{P}} - \mathbf{P}\right|\right|^2$$

$$L = \left|\left|\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}\right|\right|^2$$

Where $\mathbf{P}, \hat{\mathbf{P}}, \boldsymbol{\alpha}, \hat{\boldsymbol{\alpha}}$ are target dipoles, predicted dipoles, target polarizations, predicted polarizations.

Molecluar dynamics simulations

We constructed the PES of liquid water using HotPP with the dataset provided by DeePMD[50], including 1888 structures computed with the strongly constrained and appropriately normed (SCAN) functional[51]. Our model gives the RMSE of 2 meV/atom for the per atom energy, 49 meV/ Å for forces, and 11 meV/atom for per atom virials.

Next we used LAMMPS[52] to perform the MLMD with the system of 512 water molecules in a cubic≈24.8 Å supercell to make sure the density close to 1 g cm^{-3}. The system was equilibrated at ambient conditions using Nosé-Hoover chain thermostat[53,54] for 200 ps. 10 statistically independent initial conditions were then sampled from the last 100 ps simulation to initialize NVE trajectories of 200 ps length using a time step of 0.5 fs. The initial and final configurations of these 10 simulations can be found in Supplementary Data 1. The dipole moment and polarizability tensor were calculated every 1 fs.

Calculation of infrared absorption spectrum and Raman spectrum

We can get different types of vibrational spectra by Fourier transforming the time autocorrelation functions (ACF) of different physical properties with a trajectory sampled by molecule dynamics. For IR absorption, it can be computed by the ACF of the dipole moment as:

$$I(\omega) \propto \omega^2 \int_{-\infty}^{+\infty} e^{-i\omega t}\langle \mu(0)\cdot\mu(t)\rangle dt$$

Where μ is dipole moment, ω is the vibrational frequency, and the bracket is the average over the time origin.

And for Raman spectrum, the isotropic components can be calculated by:

$$R_{iso}(\omega) \propto \omega \tanh(\frac{\hbar\omega}{2kT})\int_{-\infty}^{+\infty} e^{-i\omega t}\langle \alpha_{iso}(0)\cdot\alpha_{iso}(t)\rangle dt$$

Where $\alpha_{iso} \equiv Tr(\alpha)/3$ is the isotropic components of the polarizability tensor. And,

$$R_{aniso}(\omega) \propto \omega \tanh(\frac{\hbar\omega}{2kT})\int_{-\infty}^{+\infty} e^{-i\omega t} Tr\langle \beta(0)\cdot\beta(t)\rangle dt$$

Where β is the anisotropic traceless tensor $\beta = \alpha - \alpha_{iso}I$.

The IR spectrum and Raman spectrum are scaled to make the maximum values consistent with the experimental data.

Data availability

The datasets used in this paper (ANI-1x, carbon, and water) are publicly available (see "Method"). Source data are provided with this paper.

Code availability

The HotPP code used in the current study is available at GitLab (https://gitlab.com/bigd4/hotpp) and Zenodo (https://doi.org/10.5281/zenodo.12952612), ref. 55.

References

1. Hohenberg, P. & Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **136**, B864–B871 (1964).
2. Behler, J. & Parrinello, M. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.* **98**, 146401 (2007).
3. Bartók, A. P., Payne, M. C., Kondor, R. & Csányi, G. Gaussian approximation potentials: the accuracy of quantum mechanics, without the electrons. *Phys. Rev. Lett.* **104**, 136403 (2010).
4. Shapeev, A. V. Moment tensor potentials: a class of systematically improvable interatomic potentials. *Multiscale Model. Simul.* **14**, 1153–1173 (2016).
5. Zhang, L., Han, J., Wang, H., Saidi, W. A. & Car, R. End-to-end Symmetry Preserving Inter-atomic Potential Energy Model for Finite and Extended Systems. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*. 4441–4451 (Curran Associates Inc., Red Hook, NY, USA, 2018).
6. Drautz, R. Atomic cluster expansion for accurate and transferable interatomic potentials. *Phys. Rev. B* **99**, 014104 (2019).
7. Fan, Z. et al. GPUMD: a package for constructing accurate machine-learned potentials and performing highly efficient atomistic simulations. *J. Chem. Phys.* **157**, 114801 (2022).
8. Batzner, S. et al. E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nat. Commun.* **13**, 2453 (2022).
9. Smith, J. S. et al. Approaching coupled cluster accuracy with a general-purpose neural network potential through transfer learning. *Nat. Commun.* **10**, 2903 (2019).
10. Daru, J., Forbert, H., Behler, J. & Marx, D. Coupled cluster molecular dynamics of condensed phase systems enabled by machine learning potentials: liquid water benchmark. *Phys. Rev. Lett.* **129**, 226001 (2022).
11. Musil, F. et al. Physics-inspired structural representations for molecules and materials. *Chem. Rev.* **121**, 9759–9815 (2021).
12. Behler, J. Atom-centered symmetry functions for constructing high-dimensional neural network potentials. *J. Chem. Phys.* **134**, 074106 (2011).
13. Gastegger, M., Schwiedrzik, L., Bittermann, M., Berzsenyi, F. & Marquetand, P. WACSF - weighted atom-centered symmetry functions as descriptors in machine learning potentials. *J. Chem. Phys.* **148**, 241709 (2018).
14. Fan, Z. et al. Neuroevolution machine learning potentials: combining high accuracy and low cost in atomistic simulations and application to heat transport. *Phys. Rev. B* **104**, 104309 (2021).
15. Fan, Z. Improving the accuracy of the neuroevolution machine learning potential for multi-component systems. *J. Phys.: Condens. Matter* **34**, 125902 (2022).
16. Xu, N. et al. Tensorial properties via the neuroevolution potential framework: fast simulation of infrared and raman spectra. *J. Chem. Theory Comput.* https://doi.org/10.1021/acs.jctc.3c01343 (2024).
17. Bartók, A. P., Kondor, R. & Csányi, G. On representing chemical environments. *Phys. Rev. B* **87**, 184115 (2013).
18. Zhang, L. et al. Deep potential molecular dynamics: a scalable model with the accuracy of quantum mechanics. *Phys. Rev. Lett.* **120**, 143001 (2018).
19. Pozdnyakov, S. N. et al. Incompleteness of atomic structure representations. *Phys. Rev. Lett.* **125**, 166001 (2020).
20. Drautz, R. Atomic cluster expansion of scalar, vectorial, and tensorial properties including magnetism and charge transfer. *Phys. Rev. B* **102**, 024104 (2020).
21. Novikov, I. S., Gubaev, K., Podryabinkin, E. V. & Shapeev, A. V. The MLIP package: moment tensor potentials with MPI and active learning. *Mach. Learn.: Sci. Technol.* **2**, 025002 (2021).
22. Gilmer, J., Schoenholz, S. S., Riley, P. F., Vinyals, O. & Dahl, G. E. Neural message passing for Quantum chemistry. In *Proceedings of the 34th International Conference on Machine Learning*. Vol. 70, 1263–1272 (JMLR.org, Sydney, NSW, Australia, 2017).
23. Schütt, K. T. et al. SchNet: A continuous-filter convolutional neural network for modeling quantum interactions. *Adv. Neural Inf. Process. Syst.* **2017**, 992–1002 (2017).
24. Thomas, N. et al. *Tensor field networks: Rotation- and translation-equivariant neural networks for 3D point clouds*. Preprint at http://arxiv.org/abs/1802.08219 (2018).
25. Unke, O. T. & Meuwly, M. PhysNet: a neural network for predicting energies, forces, dipole moments, and partial charges. *J. Chem. Theory Comput.* **15**, 3678–3693 (2019).
26. Schütt, K., Unke, O. & Gastegger, M. Equivariant message passing for the prediction of tensorial properties and molecular spectra. in *Proceedings of the 38th International Conference on Machine Learning* (eds. Meila, M. & Zhang, T.) 139 9377–9388 (PMLR, 2021).
27. Gasteiger, J., Groß, J. & Günnemann, S. Directional Message Passing for Molecular Graphs. In *International Conference on Learning Representations (ICLR)* (2020).
28. Thölke, P. & Fabritiis, G. D. Equivariant Transformers for Neural Network based Molecular Potentials. In *International Conference on Learning Representations* (2022).
29. Haghighatlari, M. et al. NewtonNet: a Newtonian message passing network for deep learning of interatomic potentials and forces. *Digital Discov.* **1**, 333–343 (2022).
30. Batatia, I. et al. *The Design Space of E(3)-Equivariant Atom-Centered Interatomic Potentials*. Preprint at http://arxiv.org/abs/2205.06643 (2022).
31. Batatia, I., Kovacs, D. P., Simm, G. N. C., Ortner, C. & Csanyi, G. MACE: Higher Order Equivariant Message Passing Neural Networks for Fast and Accurate Force Fields. In *Advances in Neural Information Processing Systems* (eds. Oh, A. H., Agarwal, A., Belgrave, D. & Cho, K.) (2022).
32. Finkelshtein, B., Baskin, C., Maron, H. & Dym, N. A Simple and Universal Rotation Equivariant Point-cloud Network. In *TAG-ML* (2022).
33. Takamoto, S., Izumi, S., Li, J. & TeaNet Universal neural network interatomic potential inspired by iterative electronic relaxations. *Computational Mater. Sci.* **207**, 111280 (2022).
34. Simeon, G. & Fabritiis, G. D. TensorNet: Cartesian Tensor Representations for Efficient Learning of Molecular Potentials. In *Thirty-seventh Conference on Neural Information Processing Systems* (2023).
35. Smith, J. S., Nebgen, B., Lubbers, N., Isayev, O. & Roitberg, A. E. Less is more: sampling chemical space with active learning. *J. Chem. Phys.* **148**, 241733 (2018).
36. Smith, J. S. et al. The ANI-1ccx and ANI-1x data sets, coupled-cluster and density functional theory properties for molecules. *Sci. Data* **7**, 134 (2020).
37. Deringer, V. L. & Csányi, G. Machine learning based interatomic potential for amorphous carbon. *Phys. Rev. B* **95**, 094203 (2017).
38. Rowe, P., Deringer, V. L., Gasparotto, P., Csányi, G. & Michaelides, A. An accurate and transferable machine learning potential for carbon. *J. Chem. Phys.* **153**, 034702 (2020).
39. Togo, A. First-principles phonon calculations with phonopy and phono3py. *J. Phys. Soc. Jpn.* **92**, 012001 (2023).

40. Togo, A., Chaput, L., Tadano, T. & Tanaka, I. Implementation strategies in phonopy and phono3py. *J. Phys.: Condens. Matter* **35**, 353001 (2023).
41. Grisafi, A., Wilkins, D. M., Csányi, G. & Ceriotti, M. Symmetry-adapted machine learning for tensorial properties of atomistic systems. *Phys. Rev. Lett.* **120**, 036002 (2018).
42. Zhang, Y. et al. Efficient and accurate simulations of vibrational and electronic spectra with symmetry-preserving neural network models for tensorial properties. *J. Phys. Chem. B* **124**, 7284–7290 (2020).
43. Zhang, Y., Xia, J. & Jiang, B. REANN: a PyTorch-based end-to-end multi-functional deep neural network package for molecular, reactive, and periodic systems. *J. Chem. Phys.* **156**, 114801 (2022).
44. Bertie, J. E. & Lan, Z. Infrared intensities of liquids xx: the intensity of the oh stretching band of liquid water revisited, and the best current values of the optical constants of H_2O(l) at 25 °C between 15,000 and 1 cm^{-1}. *Appl. Spectrosc.,* **50**, 1047–1057 (1996).
45. Brooker, M. H., Hancock, G., Rice, B. C. & Shapter, J. Raman frequency and intensity studies of liquid H2O, H218O and D2O. *J. Raman Spectrosc.* **20**, 683–694 (1989).
46. Zhang, C. et al. Modeling liquid water by climbing up jacob's ladder in density functional theory facilitated by using deep neural network potentials. *J. Phys. Chem. B* **125**, 11444–11456 (2021).
47. Sommers, G. M., Calegari Andrade, M. F., Zhang, L., Wang, H. & Car, R. Raman spectrum and polarizability of liquid water from deep neural networks. *Phys. Chem. Chem. Phys.* **22**, 10592–10602 (2020).
48. Pickard, C. J. Hyperspatial optimization of structures. *Phys. Rev. B* **99**, 054102 (2019).
49. Kingma, D. P. & Ba, J. Adam: *A Method for Stochastic Optimization*. Preprint at https://doi.org/10.48550/arXiv.1412.6980 (2017).
50. Zhang, L. et al. Phase diagram of a deep potential water model. *Phys. Rev. Lett.* **126**, 236001 (2021).
51. Sun, J., Ruzsinszky, A. & Perdew, J. P. Strongly constrained and appropriately normed semilocal density functional. *Phys. Rev. Lett.* **115**, 036402 (2015).
52. Thompson, A. P. et al. LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Computer Phys. Commun.* **271**, 108171 (2022).
53. Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **52**, 255–268 (1984).
54. Hoover, W. G. Canonical dynamics: equilibrium phase-space distributions. *Phys. Rev. A* **31**, 1695–1697 (1985).
55. Wang, J. et al. E(n)-equivariant cartesian tensor passing potential: HotPP. *Zenodo* https://doi.org/10.5281/zenodo.12952612 (2024).
56. Medders, G. R. & Paesani, F. Infrared and raman spectroscopy of liquid water through "first-principles" many-body molecular dynamics. *J. Chem. Theory Comput.* **11**, 1145–1154 (2015).

Acknowledgements

This work was supported by the National Natural Science Foundation of China (grant number. 12125404), the Basic Research Program of Jiangsu (Grant BK20233001, BK20241253), the Jiangsu Funding Program for Excellent Postdoctoral Talent (2024ZB075), the Postdoctoral Fellowship Program of CPSF (Grant GZC20240695), and the Fundamental Research Funds for the Central Universities. Y. W. was partially supported by the Computational Chemical Sciences Center: Chemistry in Solution and at Interfaces (CSI) funded by DOE Award DE-SC0019394, and used resources of the National Energy Research Scientific Computing Center (NERSC) operated under Contract No.DE-AC02-05CH11231 using NERSC award ERCAP0021510. The calculations were carried out using supercomputers at the High-Performance Computing Center of Collaborative Innovation Center of Advanced Microstructures, the high-performance supercomputing center of Nanjing University.

Author contributions

J.Sun. conceived and led the project. J.W. and Y.W. deduced the formula and completed the coding. H.Z. helped to optimize the efficiency of the code. Z.Y., Z.L. and J.Shi. trained the model in different datasets. J.W., Y.W., H.-T.W., D.X. and J. Sun wrote the manuscript. All authors discussed the results and commented on the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-51886-6.

Correspondence and requests for materials should be addressed to Jian Sun.

Peer review information *Nature Communications* thanks Wilson Gregory and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2024