

# Global Neural Network Potential with Explicit Many-Body Functions for Improved Descriptions of Complex Potential Energy Surface

Pei-Lin Kang, Zheng-Xin Yang, Cheng Shang,\* and Zhi-Pan Liu\*

Cite This: <https://doi.org/10.1021/acs.jctc.3c00873>

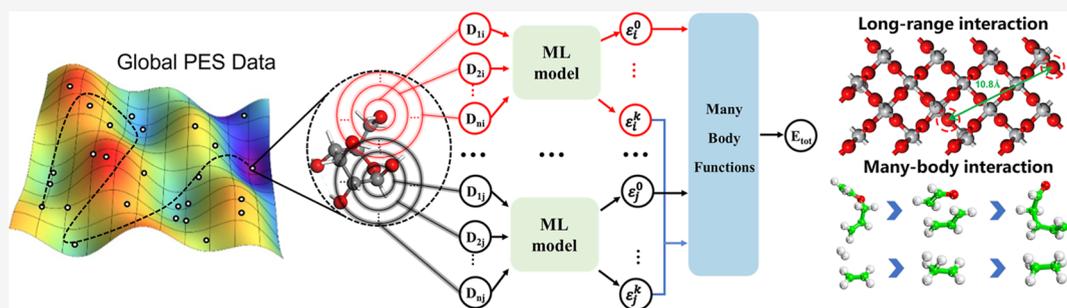
Read Online

ACCESS |

Metrics & More

Article Recommendations

Supporting Information



**ABSTRACT:** The high dimensional machine learning potential (MLP) that has developed rapidly in the past decade represents a giant step forward in large-scale atomic simulation for complex systems. The long-range interaction and the poor description of chemical reactions are typical problems of high dimensional MLP, which are mainly caused by the poor structure discrimination of the atom-centered ML model. Herein, we propose a low-cost neural-network-based MLP architecture for fitting global potential energy surface data, namely, G-MBNN, that can offer improved energy and force resolution on a complex potential energy surface. In G-MBNN, a set of many-body energy terms based on the local atomic environment are explicitly included in computing the total energy—the total energy of the system is written as the sum of atomic energy and many-body energy contributions. These extra many-body energy terms are computationally low-cost and, importantly, can provide easy access to delicate energy terms in complex systems such as very short repulsion, long-range attractions, and sensitive angular-dependent covalent interactions. We implement G-MBNN in the LASP code and demonstrate the improved accuracy of the new framework in representative systems, including ternary-element energy materials  $\text{LiCoO}_x$ ,  $\text{TiO}_2$  with defects, and a series of organic reactions.

## 1. INTRODUCTION

The past few years have seen tremendous progress in the use of machine learning (ML) potentials for accelerating atomic simulations for systems with complex potential energy surface (PES).<sup>1–3</sup> The atom-centered ML potential (AC-MLP) is a common approach, as introduced in the high-dimensional neural network (HDNN) by Behler and Parrinello,<sup>4</sup> which decomposes the total energy as a sum of local, environment-dependent atomic energies ( $E = \sum_{i=1}^n E_i$ ). Following this principle, a variety of ML potential recipes, such as ANI,<sup>5</sup> LASP,<sup>6,7</sup> GAP,<sup>8,9</sup> and message-passing networks,<sup>10–14</sup> are successfully developed that differ mainly in how the local atomic environments are represented. These HD MLPs utilize atom-centered descriptors and aim to describe the complex PES of hundreds to millions of atoms, being different from the earlier MLP used in studying full reaction trajectories of molecular systems (several to tens of atoms) that takes permutationally invariant polynomials of molecular coordinate as descriptors.<sup>15–17</sup> Comparatively, the less accurate description of long-range (electrostatic) interactions and high-energy structures (such as reaction transition state, TS) are two often referred problems in HD MLPs, although the addition of

global PES data can reduce the error.<sup>18</sup> How to improve the transferability of the AC-MLP, particularly with a small data set, is an intriguing but challenging task in developing the next generation of MLPs.

One obvious solution to improve the transferability of MLP is to design better input vectors, the so-called structure descriptors, to distinguish tiny structural changes between structures. While the descriptors used in HDNN are relatively simple, being the sums of two- or three-body functions composed by Gaussian-type functions of pairwise distances and trigonal functions,<sup>4,19</sup> the incorporation of more complex function forms, such as the spherical harmonics, dihedral angles, and power functions (SOAP,<sup>9</sup> ANI,<sup>5</sup> and PTSD<sup>7</sup>), does appear to improve the MLP performance. As the direct sums of

Received: August 9, 2023

two-body functions reduce the resolution on individual pairwise interactions, the Deep potential proposed to construct descriptors using the trainable transformation matrix acting on pairwise distances.<sup>14</sup> Similarly, graph convolution network methods were also applied to automatically extract structural information via trainable network parameters.<sup>10–12</sup> In general, these efforts on atomic descriptors are however limited by the cutoff range to the central atom, which is often below a few (e.g., 6–7) angstroms due to the rapid increase of computational cost in computing the long-range many-body descriptors.

Another important direction to construct MLP is to introduce new energy decomposition schemes with physically meaningful terms. In particular, considering that the long-range interaction should be of an electrostatic nature, recent years have seen great interests to retrieve the explicit electrostatic interaction<sup>20–24</sup> by exploiting the atomic charge, which can be computed from DFT (e.g., the Hirshfeld charges via the charge-localization scheme).<sup>22</sup> The atomic charge, strictly speaking, is however a nonlocal function due to the possibility of long-range charge transfer, and thus its prediction via ML, even with the reference atomic charges in the training set, is challenging with only local position quantities. For example, Goedecker et al. proposed a charge equilibration NN technique,<sup>21</sup> which utilized the empirical electronegativities based on the local environment to learn the atomic charge from DFT calculations. One step further, the Becke Population Neural Network (BpopNN) method<sup>23</sup> incorporates electronic populations obtained via the constrained DFT method in learning MLP and the MLP input vectors explicitly contain population variables. Since the training data in BpopNN are not from standard DFT calculations, the applications applied to complex material systems are limited. Overall, these atomic charge prediction methods are generally computationally expensive and not ideal for complex material systems with large PES data sets for learning.

Following our previous contributions on global NN (G-NN) potential<sup>7</sup> and LASP program (large-scale atomic simulation with a neural network potential, LASP 3.0),<sup>25,26</sup> here, we propose a low-cost NN-based MLP architecture that can offer much-improved resolution on complex PES, where many-body energy terms based on the local atomic environment are explicitly included in computing total energy. These extra many-body energy terms are computationally low-cost and, importantly, can provide easy access to delicate short- and long-range energy terms in complex systems, such as the very-short repulsion, long-range attractions, and sensitive angular-dependent covalent bondings. We implement the method in our LASP code and demonstrate the improved accuracy of the new framework in a few representative systems, including ternary energy materials, oxides with defects, and organic reactions.

## 2. METHODS

**2.1. Physically Meaningful Terms in Force Fields and G-NN Potential.** The purpose of MLP is very much similar to the classical force field developed since 1950s in the past century, aiming for a fast evaluation of PES. The energy decomposition scheme in AC-MLP does follow the idea of the embedded atom method, a general force field method designed to describe alloys,<sup>27</sup> which considers the total energy as the sum of atomic energies and the atomic energy is a function of the local electron density of the embedded atom in the

environment. The force field methods, while being physically meaningful, are difficult to generalize to treat all material systems due to the diverse nature of atom–atom interactions and the limited parameter space in fitting the PES. On the other hand, the AC-MLP connection between the atomic coordinate and the atomic energy lacks clear physically meaningful terms and is basically numerical by learning the training data set from quantum mechanics calculations using a huge parameter space. This could lead to poor transferability in MLP, for example, to predict the structures in the PES regions with low training data density.

On the other hand, apart from the atomic energy scheme, another common approach in the classical force field is to decompose the total energy into many-body terms, such as two-body (bond), three-body (angle), and four-body (dihedral angle) terms. For example, a general two-body term can be expressed as

$$E = f(r_{ij}; a, r_0, n) \quad (1)$$

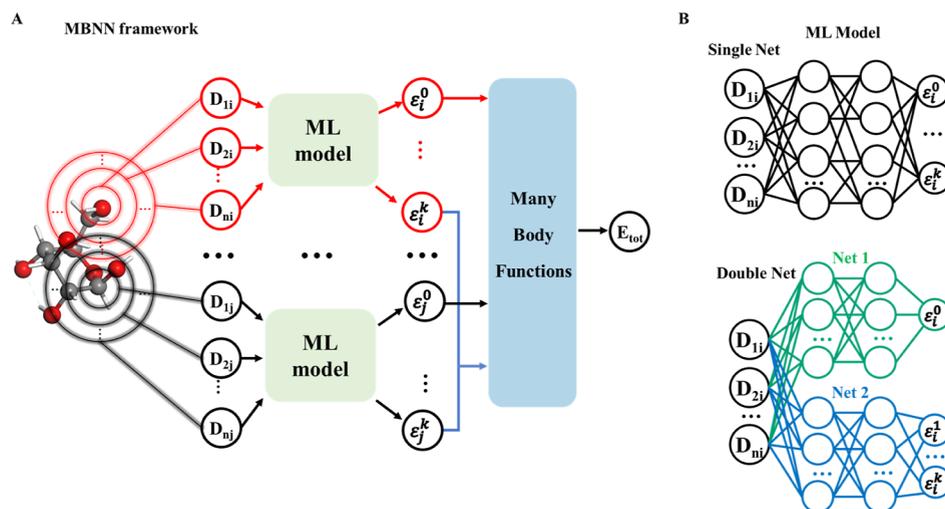
where  $r_{ij}$  is the atom pair distance;  $a$  and  $r_0$  are optional parameters related to the pair for distinguishing different elemental types, bond order, and the equilibrium distance;  $n$  is the parameter describing the decay behavior between the atom pair, which is physically interpretable, for example, in electrostatic interaction  $n = 1$ , in dipole–dipole interaction  $n = 3$  and in van der Waals interaction  $n = 6$ . The reason that AC-MLP does not utilize the many-body energy decomposition scheme lies in the fact that the number of many-body terms increases rapidly with the increase of system size—it is practically impossible for MLP to connect position coordinates with each many-body term in complex systems.

Instead, the MLP envelops all structure information in the structure descriptors, which are, strictly speaking, pure numerical functions satisfying the translation, rotation, and permutation invariance of structure as required for describing PES. Taking the G-NN potential as an example, the power-type structure descriptors (PTSD) are utilized as the input, which is a set of highly sophisticated functions, and designed to be compatible with the SSW global optimization data set. In PTSD, not only the traditional two-body (S1–S2) and three-body (S3–S5) terms are available but also the four-body terms (S6) are implemented and the spherical functions are introduced to enhance the structure discrimination. The formulas of S2 and S5 are given below, where the combination of the power function and spherical harmonic function, similar to atomic wave functions, provides a convenient way to couple the radial and the angular information on atom.

$$f_c(r_{ij}) = \begin{cases} 0.5 \times \tanh^3 \left[ 1 - \frac{r_{ij}}{r_c} \right], & \text{for } r_{ij} \leq r_c \\ 0 & \text{for } r_{ij} > r_c \end{cases} \quad (2)$$

$$R^n(r_{ij}) = r_{ij}^n \cdot f_c(r_{ij}) \quad (3)$$

$$S_i^2 = \left[ \sum_{m=-L}^L \left| \sum_{j \neq i} R^n(r_{ij}) Y_{Lm}(r_{ij}) \right|^2 \right]^{1/2} \quad (4)$$



**Figure 1.** Illustration of the architecture of MBNN. (A) Each atomic ML model (NN utilized in this work) outputs a series of coefficients  $\epsilon_i^k$ , which act as the parameters for different many-body functions. The total energy of the system is written as the sum of many-body functions. (B) Two different implementations of the ML model based on feed-forward NN. Single net: one single NN outputs all coefficients; double net: two separate NNs output the  $\epsilon_i^0$  and  $\epsilon_i^k$  ( $k > 0$ ) coefficients.

$$S_i^S = \left[ \sum_{m=-L}^L \left| \sum_{j,k \neq i} R^n(r_{ij}) \cdot R^m(r_{ik}) \cdot R^p(r_{jk}) \cdot (Y_{Lm}(r_{ij}) + Y_{Lm}(r_{ik})) \right|^2 \right]^{1/2} \quad (5)$$

In the equations,  $r_{ij}$  is the internuclear distance between atom  $i$  and  $j$ , and  $\theta_{ijk}$  is the angle centered at  $i$  atom with  $j$  and  $k$  being neighbors ( $i$ ,  $j$ , and  $k$  are atom indices). The key ingredients in PTSD are the cutoff function  $f_c$  that decays to zero beyond the  $r_c$  (eq 2), power-type radial function, trigonometric angular functions, and spherical harmonic function. It should be noted that the terms of the PTSD will increase rapidly as cutoff  $r_c$  becomes larger.

**2.2. MBNN Architecture.** To overcome the drawbacks in AC-MLP, in this work, we propose a new MLP architecture, namely, the many-body-function corrected global neural network (G-MBNN) as a replacement to the G-NN potential utilized for fast global PES exploration.<sup>28</sup> In G-MBNN, the total energy is expressed as the sum of both the atomic energies and many-body functions. G-MBNN is also targeted to learn global PES data sets generated by stochastic surface walking (SSW) global optimization, which involves highly complex atomic environments and multiple chemical compositions and element combinations. The parameters in these many-body functions are learned from local atomic environments via machine learning. The G-MBNN is implemented in our large-scale atomic simulation with a neural network potential (LASP) program by using the power-type-structure descriptors (PTSD) to distinguish structures and the SSW global PES data set for training. More than 100 MBNN potentials are now available for describing complex PES systems covering more than 50 different elements (<http://www.lasphub.com/#/lasp/nnLibrary>). In the following, we elaborate on the MBNN architecture and provide benchmarks for treating different PES problems.

Figure 1 illustrates how the total energy is derived from a molecular structure by using the G-MBNN framework. Two different implementations of G-MBNN, a single feed-forward neural network (FFNN) version and a double FFNNs version,

are illustrated in Figure 1B, which differ in how the parameters in MB functions are yielded from NNs.

As shown in Figure 1A, for each atom in a molecule, a set of PTSD functions  $\{D_i\}$  are calculated from the Cartesian coordinates  $\{R_i\}$ . These functions describe the local chemical environment of each atom and are used as input for atomic neural networks. Unlike standard G-NN, which output only a single value as the atomic energy, the G-MBNN outputs a vector, including the atomic energy contribution  $\epsilon_i^0$ , a one-body function ( $f_s$ ), and a series of coefficients  $\epsilon_i^k$  ( $k = 1, 2, \dots$ ) for building different many-body functions, namely,  $f_d$ ,  $f_v$ , and  $f_q$  for the two-body, three-body, and four-body functions, respectively. As shown in eq 6, the total energy  $E_{\text{tot}}$  of the system can thus be expressed as a sum of the atomic energy contributions ( $f_s$ ) and the many-body functions ( $f_d, f_v, \dots$ ). The double FFNN implementation, different from the single FFNN with all  $\epsilon_i^k$  come from the same network (Figure 1 upper), yields  $\epsilon_i^0$  and  $\epsilon_i^k$  in separated networks (Figure 1B lower).

$$E = \sum_{s=1}^{ns} \sum_{i=1}^{na} f_s(\epsilon_i^s) + \sum_{d=1}^{nd} \sum_{i=1}^{na} \sum_{j>i}^{na} f_d(\epsilon_i^d, \epsilon_j^d, r_{ij}) + \sum_{t=1}^{nt} \sum_{i=1}^{na} \sum_{j>i}^{na} \sum_{k>j}^{na} \sum_{\text{ang}=1}^3 f_t(\epsilon_i^t, \epsilon_j^t, \epsilon_k^t, r_{ij}, r_{ik}, r_{jk}, \theta_{\text{ang}}) + \quad (6)$$

These many-body functions can be written into general function forms shown in eqs 7–13, as implemented in LASP program, where  $\epsilon_i^k$  from NN acts as the parameters. In these functions,  $r_{ij}$  is the distance between atoms  $i$  and  $j$  and  $\theta_{ijk}$  is the angle centered at the  $i$  atom with neighbors  $j$  and  $k$  ( $i$ ,  $j$ , and  $k$  are atom indices);  $n_a$  is the total atom number of the system;  $n_s$ ,  $n_d$ , and  $n_c$  are the number of one-body, two-body, three-body function in MBNN architecture, respectively ( $s$ ,  $d$ , and  $t$  are function indices);  $\lambda$ ,  $m_d$ ,  $m_v$ ,  $l_v$ , and  $l_q$  is adjustable parameters of these functions. The  $f_c$  in all of the equations is the cutoff function with the cutoff radius  $r_{\text{cut}}$  and the adjustable parameter  $\alpha$  and  $\beta$ , as defined in eq 14 (the cutoff function in MBNN functions can be different from that in PTSD as shown in eq 2). These functional forms with adjustable parameters

allow NN to quickly identify physical interactions that can satisfy the designed function relation. In principle, a shorter cutoff and larger  $m$  in eq 8 describe the (very) short interactions (e.g., Pauli repulsion), while a longer cutoff and smaller  $m$  functions are designed to improve the long interaction (e.g., electrostatic origin).

$$f_s = \lambda \varepsilon_i^s \quad (7)$$

$$f_{d1} = \lambda \frac{\varepsilon_i^d \varepsilon_j^d}{r_{ij}^{m_d}} f_c(r_{ij}) \quad (8)$$

$$f_{d2} = \lambda \varepsilon_i^d \varepsilon_j^d e^{-r_{ij}} f_c(r_{ij}) \quad (9)$$

$$f_{t1} = \lambda \frac{\varepsilon_i^t (\varepsilon_j^t + \varepsilon_k^t) (2 + \cos(\theta_{ijk}))^t}{r_{ij}^{m_t} r_{ik}^{m_t} r_{jk}^{m_t}} f_c(r_{ij}) f_c(r_{ik}) f_c(r_{jk}) \quad (10)$$

$$f_{t2} = \lambda \frac{\varepsilon_i^t (\varepsilon_j^t + \varepsilon_k^t) (2 - \cos(\theta_{ijk}))^t}{r_{ij}^{m_t} r_{ik}^{m_t}} f_c(r_{ij}) f_c(r_{ik}) \quad (11)$$

$$f_{t3} = \lambda \frac{\varepsilon_i^t \varepsilon_j^t \varepsilon_k^t}{r_{ij}^{m_t} r_{ik}^{m_t} r_{jk}^{m_t}} f_c(r_{ij}) f_c(r_{ik}) f_c(r_{jk}) \quad (12)$$

$$f_q = \lambda (2 + \cos(\delta_{ijkl}))^q f_c(r_{ij}) f_c(r_{ik}) f_c(r_{il}) \quad (13)$$

$$f_c(r_{ij}) = 0.5 \times (1 - \tanh(\alpha(r_{ij} - (r_{\text{cut}} - \beta)))) \quad (14)$$

In the MBNN architecture, the atomic force is analytically derived based on eq 6, where the force component  $F_u$  acting on the atom  $u$  is the derivative of the total energy with respect to its coordinate vector  $\mathbf{r}_u$  that is further related to the derivatives to the atomic coefficients  $\varepsilon_i^k$  and the pairwise distance  $r_{iu}$ . For illustration, we explain how the force  $F_u$  can be calculated for the one-body (eq 7) and two-body (eq 8) functions in the following, as shown in eq 15.

$$F_u = - \frac{\partial (\sum_{s=1}^{ns} \sum_{i=1}^{na} f_s(\varepsilon_i^s) + \sum_{d=1}^{nd} \sum_{i=1}^{na} \sum_{j>i}^{na} f_d(\varepsilon_i^d, \varepsilon_j^d, r_{ij}))}{\partial \mathbf{r}_u} \quad (15)$$

We define the following term in eq 16 for simplicity.

$$R(r_{iu}) = \frac{f_c(r_{iu})}{r_{iu}^{m_d}} \quad (16)$$

Substituting eqs 8 and 9 with eq 15, we obtain eq 17

$$F_u = - \sum_{s=1}^{ns} \sum_{i=1}^{na} \frac{\partial \varepsilon_i^s}{\partial \mathbf{r}_u} - \sum_{d=1}^{nd} \sum_{i=1}^{na} \left( \frac{\partial \varepsilon_i^d}{\partial \mathbf{r}_u} \sum_{j \neq i}^n \varepsilon_j^d R(r_{ij}) \right) - \sum_{d=1}^{nd} \sum_{i=1}^{na} \sum_{j>i}^{na} \varepsilon_i^d \varepsilon_u^d \frac{\partial R(r_{ij})}{\partial \mathbf{r}_u} \quad (17)$$

$$F_u = - \sum_{i=1}^{na} \frac{\partial \varepsilon_i^s}{\partial D_i} \frac{\partial D_i}{\partial \mathbf{r}_u} - \sum_{d=1}^{nd} \sum_{i=1}^{na} \left( \frac{\partial \varepsilon_i^d}{\partial D_i} \frac{\partial D_i}{\partial \mathbf{r}_u} \sum_{j \neq i}^n \varepsilon_j^d R(r_{ij}) \right) - \sum_{d=1}^{nd} \sum_{i=1}^{na} \varepsilon_i^d \varepsilon_u^d \frac{\partial R(r_{iu})}{\partial \mathbf{r}_u} \quad (18)$$

In eq 17, the derivatives of  $\varepsilon_i^s$  and  $\varepsilon_i^d$  with respect to the atomic positions in the first two terms are determined by the NN back-propagation, and the third term, the direct force of the two-body function, is nonzero only if  $r_{ij}$  is related to  $u$ . Finally, eq 17 can be further derived by inserting the input PTSD functions  $\{D_i\}$  as eq 18. In the MBNN framework, all forces and stresses can be obtained analytically and together with the total energy, all the quantities are trained simultaneously in LASP. It is important to note that as derived in eq 18, the computational cost for the derivative calculation of the descriptor to the atom  $(\frac{\partial \varepsilon_i^d}{\partial D_i} \frac{\partial D_i}{\partial \mathbf{r}_u})$  does not increase with

respect to the increase of the cutoff of MB functions. This has a great benefit in computation speed since the derivative-related computation becomes a dominant expense with the increase of the PTSD cutoff radius and the number of PTSDs encountered in complex material systems. As a result, the extra cost in G-MBNN computation is generally low (e.g., 10–20%), which is governed by the MB function cutoff and the number of MB functions. In practice, we can reduce the cutoff of relatively expensive PTSD functions and enlarge the cutoff of MB functions to balance accuracy and efficiency.

Apart from the low computational costs, we may emphasize two key advantages of G-MBNN. First, the MB functions provide explicit functional forms to fit complex PES, which can help to describe sharp energy variation due to small local structural changes. The large Pauli repulsion with a very short atom–atom distance is a case. More importantly, for some molecular chemical reactions, it also happens that the local bond breaking/making causes a violent change in energy, leading to a sharp transition state. While the specific MLP using permutationally invariant polynomials of molecular coordinate as descriptors can well describe such difficult reaction profiles,<sup>15–17</sup> in our applications of HD AC-MLP we found that the structural descriptors in the summation forms (e.g., PTSD), although satisfy the translation, rotation, and exchange invariance requirements of descriptors, sacrifice the high resolution to tiny structural changes and is often less accurate at the transition state regions than the initial state region, particular when the global PES data set is utilized for fitting the MLP. Instead of the direct approach to augment the training data set with a set of closely related configurations to capture the specific PES characteristic, the MB functions incorporating the individual pairwise distance, the energy and gradient contributions to the target atom pairs can be learned readily to produce large energy fluctuations due to a local structural change.

Second, the MB functions enhance the long-range coupling between atoms, and thus, the long-range interactions exceeding the cutoff radius of atom-centered descriptors can be better described. The ionic solids with low-concentration defects represent such an example, where the distance between defects can easily exceed 10 Å and thus are not directly visible to structure descriptors. To allow the defects to see each other, one may have to increase the cutoff of descriptors, which, however, will heavily hamper the MLP speed. Instead, MB functions can be designed to allow such interactions to contribute explicitly to the total energy, which should be reflected by the change in coefficient  $\varepsilon$  of atoms near the defect site.

It might be mentioned that our G-MBNN potential is different from the linear atomic cluster expansion (ACE) force field recently proposed by Kovács et al.<sup>29</sup> both in the descriptor

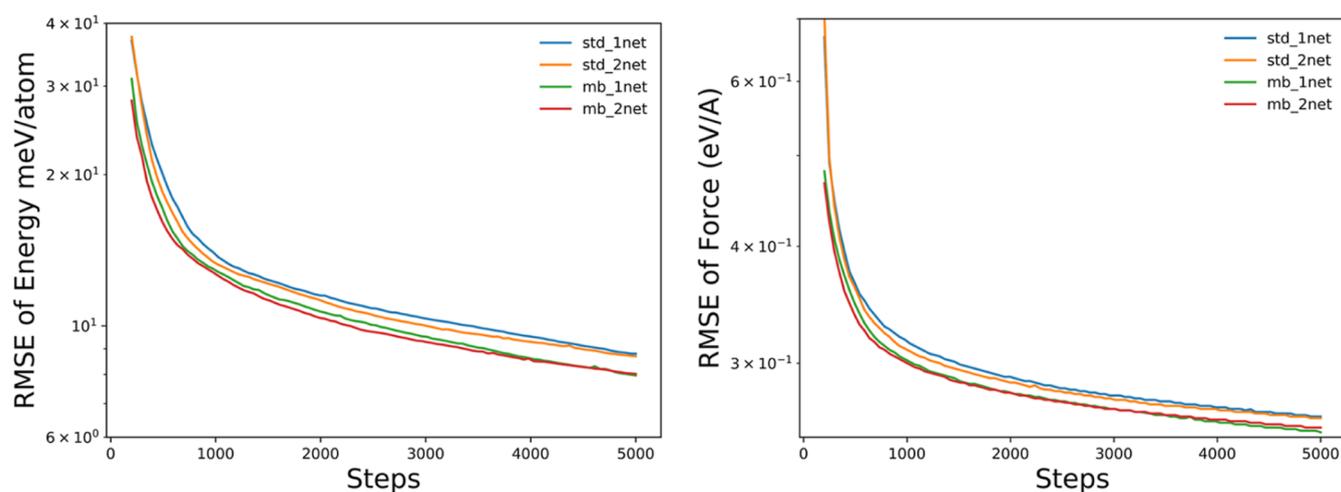


Figure 2. Learning curve for four different MLPs in training LiCoO global PES data set.

and the nonlinearity of the function forms, although both approaches construct the MB functions to describe the total energy. The ACE force field is proposed to fit the MD data set of small organic molecules (without periodicity). It generalizes the power spectrum (also known as the SOAP) descriptor, a two-body density function (one-particle basis), to an arbitrary body order for describing the atomic chemical environment. In addition, the ACE force field is linear with respect to the fitting parameters which are the coefficients of many-body terms. As illustrated below, G-MBNN will be applied to the global PES problems of complex material systems with and without periodicity. In G-MBNN, more complex PTSDs are utilized as the atomic descriptor, which by themselves contain up to four-body terms (e.g., S6) to account for the complex atomic environment. Furthermore, the nonlinear NN is utilized to generate each atomic parameter in MB expansions (eqs 7–14), which is numerically more complex and flexible and thus offers a power solution to complex PES problems in general.

### 3. RESULTS AND DISCUSSION

Below we present the performance of the G-MBNN potential on a variety of systems including metal oxides and covalent organic molecules. By comparing the results from standard G-NN potential and DFT calculation results, we show that the G-MBNN method generally offers a better description for complex material and molecular PES, featuring a shortened training time and improved accuracy on long-range interaction and chemical reactions.

All data set for fitting NN potential in this work is generated by first-principles periodic DFT calculations with plane wave basis set as implemented in VASP package.<sup>30,31</sup> The projector augmented wave (PAW) pseudopotential is utilized to describe the ionic core electrons. The electron exchange and correlation effects are described by the GGA-PBE functional. The kinetic energy cutoff for plane wave basis is 450 eV and the fully automatic Monkhorst–Pack  $k$ -mesh grid is generated with 25 times the reciprocal lattice vectors.

**3.1. Training Accuracy.** Taking the ternary-element energy material system LiCoO as an example, we illustrate the efficiency of G-MBNN by comparing four different NN architectures at comparable parameter sizes. In constructing the NN potential, we utilize the same size of input layers, i.e., 338 PTSDs for each element, and only consider two-body

radial functions  $f_{d1}$  in MBNN for clarity, as described in eq 8, where the  $r_{\text{cut}}$  is set as 16.0 Å and  $m_d = 3$ . These NN frameworks are detailed as follows.

**3.1.1. mb\_2net.** Two separate NNs with three hidden layers, i.e., 338-144-80-80-1 nodes and 338-80-80-80-1 nodes, respectively. The numbers 144, 80, and 80 are hidden layer neuron numbers and the two outputs from two network are the atomic energy contribution  $\epsilon_i^1$  and two-body function parameter  $\epsilon_i^2$ , respectively.

**3.1.2. mb\_1net.** One net with three hidden layers (338-178-144-144-2) with the output being atomic energy contribution  $\epsilon_i^1$  together with two-body function parameter  $\epsilon_i^2$ .

**3.1.3. std\_2net.** In analogy to **mb\_2net**, two standard networks (338-144-80-80-1 nodes and 338-80-80-80-1 nodes, respectively) are utilized to obtain the atomic energy  $E_i$  by summing up two outputs ( $\epsilon_i^1$ ,  $\epsilon_i^2$ ) of the networks, i.e.,  $\epsilon_i = \epsilon_i^1 + \epsilon_i^2$ .

**3.1.4. std\_1net.** In analogy to **mb\_1net**, one standard network (338-178-144-144-1) is utilized to generate the atomic energy.

By using SSW global optimization<sup>32</sup> and the standard iterative self-learning procedure,<sup>6,7</sup> we first generate a global data set of LiCoO system, which contains 11,008 distinct structures that vary in morphology (clusters, layer, and bulk structures) and chemical composition ( $\text{Li}$ ,  $\text{CoO}_x$ ,  $\text{LiCoO}_x$ ), as detailed in Table S1. This data set is recently utilized for training a G-MBNN potential for describing the Li ion dynamics in  $\text{CoO}_2$  material.<sup>33</sup>

The four NN frameworks are then utilized to train the global LiCoO data set. The learning curves using BFGS optimizer<sup>34,35</sup> are plotted in Figure 2. It can be seen that although the architectures are different, the NN with MB function corrections (**mb\_2net** and **mb\_1net**) generally improve the training speed and accuracy compared to the architectures without MB corrections (**std\_2net** and **std\_1net**). Specifically, in 5000 epochs, the root-mean-square error (RMSE) of energy and force of **mb\_1net** could reach 7.9 meV/atom and 0.253 eV/Å, while the RMSE of energy and force of **std\_1net** are higher, being 8.8 meV/atom and 0.263 eV/Å. As shown in Figure 2, **mb\_1net** reaches the same accuracy level with 30% fewer epochs (3700 vs 5000) compared to **std\_1net** and the training cost for each epoch increases by  $\sim 10\%$ . As a result, MBNN improves the training speed greatly and achieves higher accuracy.

In addition, by comparing the performance of **mb\_2net** and **mb\_1net**, we found that single-network G-MBNN performs better. Although the energy error in **mb\_2net** drops faster initially, **mb\_1net** catches up after 4000 epochs. Furthermore, after 5000 epochs, **mb\_1net** has a better performance in the force fitting (RMSE 0.253 eV/Å), compared to that (0.256 eV/Å) in **mb\_2net**. Since two MBNN potentials have comparable parameter space, the better performance of **mb\_1net** could be attributed to the coupling of NN parameters in outputting multiple targets. A single fully connected network helps to slowly convolute the parameter space and transform the structural information into multiple target outputs.

**3.2. Long-Range Interaction.** Taking the ionic crystal TiO system as the example, below we will show that the explicit long-range two-body  $f_d$  can markedly improve the description of long-range interaction. The presence of defective sites (e.g., oxygen vacancies, dislocations) in the TiO<sub>2</sub> system is of great significance to the material properties and catalytic performance.<sup>6</sup> The interaction range between these low-concentration defected sites is generally far beyond the cutoff radius (6–7 Å) of structure descriptors, and the defected TiO<sub>2</sub> systems are the typical examples that are difficult to describe accurately using the standard AC-MLP framework.

Our TiO global PES database has 27279 structures, containing different chemical compositions (Ti, TiO<sub>x</sub>). The database is a subdatabase of Ti–O–H global PES data set utilized previously for identifying new TiO<sub>2</sub> crystal phases<sup>36</sup> and exploring the catalytic reactions. For comparison, four NN potentials with a similar parameter space are trained on the same TiO database. They are described as follows and the network parameters are listed in Table 1.

**Table 1.** MB Function Parameters in Different G-MBNN<sup>a</sup>

MLP	MB terms	$r_{\text{cut}}$ (Å)	$m_d/m_t$	$l_t$
<b>mb1</b>	$f_{d1}$	7	4	
<b>mb2</b>	$f_{d1}$	15	4	
<b>mb3</b>	$f_{d1}$	15	3	
	$f_{d1}$	3	6	
	$f_{t1}$	4.5	1	2
	$f_{t2}$	4.5	1	2

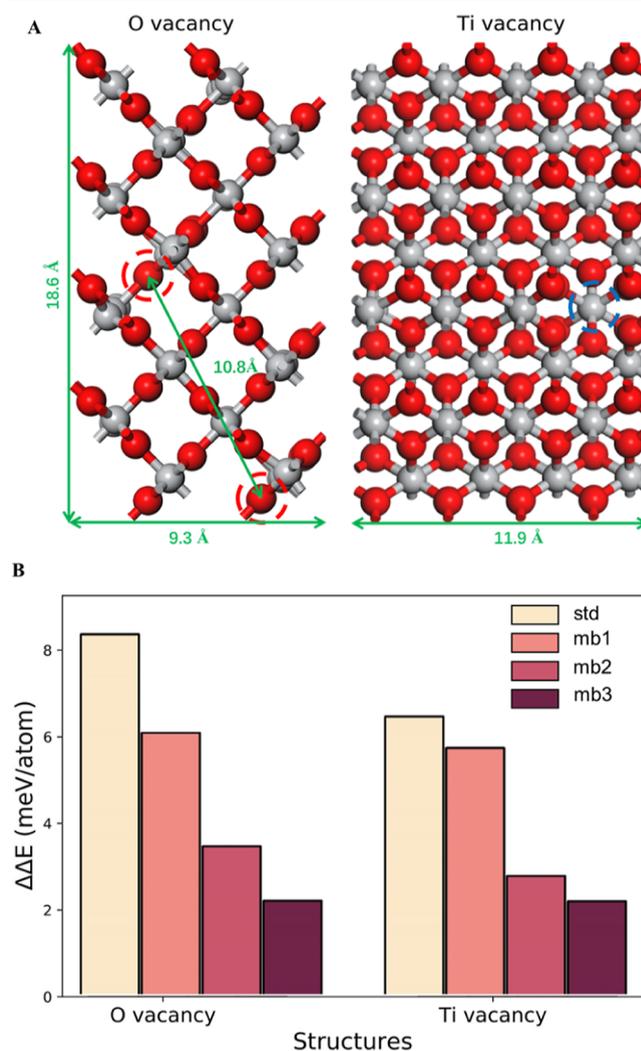
<sup>a</sup>The maximum PTSD cutoff is 7 Å for all NN potentials.

- standard G-NN potential without MB corrections (**std**) with the network architecture 173-144-80-80-1;
- 1-net MBNN potential (**mb1**) with one additional short-range MB function  $f_{d1}$  ( $m = 4$ ,  $r_{\text{cut}} = 7$  Å) and a network architecture of 173-144-80-80-2.
- 1-net MBNN potential (**mb2**) with one additional long-range MB function  $f_{d1}$  ( $m = 4$ ,  $r_{\text{cut}} = 15$  Å) and a network architecture of 173-144-80-80-2.
- 1-net MBNN potential (**mb3**) with four many-body functions (short-range  $f_{d1}$ , long-range  $f_{d1}$ , and two three-body MB functions  $f_{t1}$ ,  $f_{t2}$ ) and a network architecture of 173-144-80-80-5.

After 10000 epoch training, the RMSE of energy and force of **std**, **mb1**, **mb2**, and **mb3** are 5.97 meV/atom, 0.140 eV/Å; 5.73 meV/atom, 0.133 eV/Å; 5.80 meV/atom, 0.133 eV/Å; and 5.76 meV/atom, 0.132 eV/Å, respectively (the details of the learning curve are shown in Figure S1). Consistent with the previous results for LiCoO, three G-MBNN of TiO all achieve

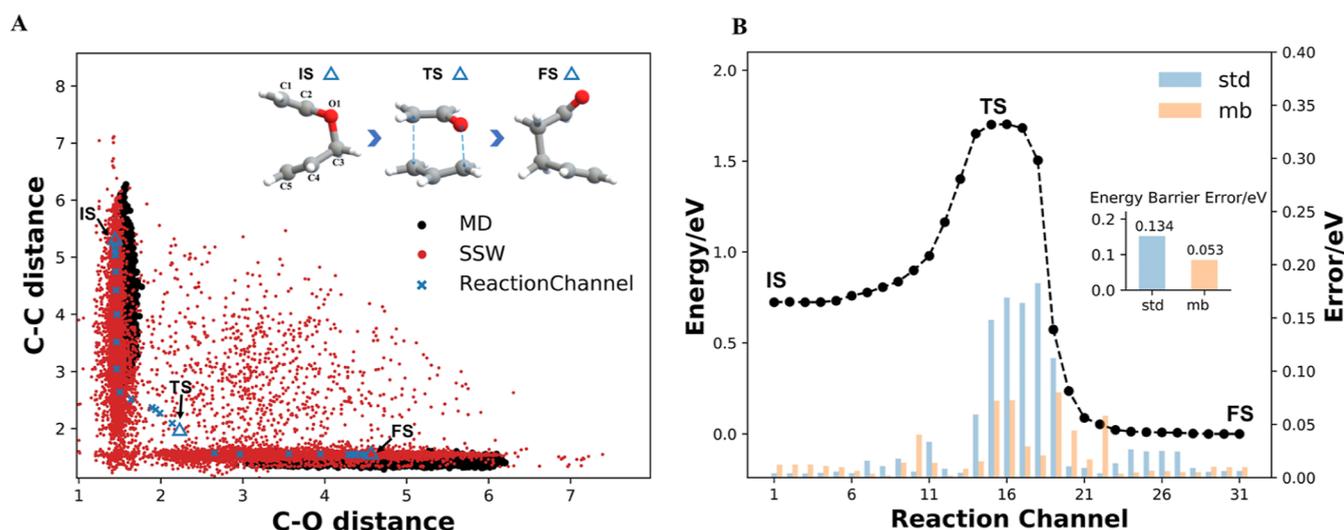
an accuracy higher than that of the standard G-NN potential. Among three G-MBNN, the **mb1** and **mb3** are better in the energy fitting, suggesting the presence of the short-range MB correction  $f_{d1}$  is critical for improving the overall accuracy of the PES.

To examine the transferability of the MLP in describing the long-range interactions, we design two large TiO<sub>2</sub> systems with low concentrations of defects, as shown in Figure 3A, both



**Figure 3.** MBNN performance for predicting defects formation. The benchmark of MLP with DFT results for two TiO<sub>2</sub>(rutile) structures with defects. The positions of the atomic defects are marked with a hollow circle. Red: O, blue: Ti.  $\Delta\Delta E$  is the comparison of NN calculated vacancy formation energy  $\Delta E$  with respect to DFT calculations.

originating from a rutile crystal bulk [192 atoms,  $p(2^*4^*4)$ ]. The first one deletes two O atoms (two O vacancies) and the second one deletes one Ti atom (Ti vacancy) per supercell. The defect–defect distances are indicated in the figure. In our comparison, the two structures and the corresponding pristine rutile structure are fully optimized using the four NN potentials and DFT calculations. The energy difference between the defect structures and the pristine structure, i.e., the energy cost to create the defects, are thus obtained from different potentials/methods, and the results are shown in Figure 3B.



**Figure 4.** MBNN performance for predicting Claisen reaction. (A) The data points are distributed from MD simulations and from SSW global search. The points along the reaction channel from DFT calculations are indicated for illustrating the initial (IS), transition (TS), and final state (FS) regions. (B) The benchmark was 31 structures along the reaction channel using MBNN (**mb**) and standard NN potentials (**std**). The black dots are the DFT relative energetics (left y-axis) of the structures along the reaction channel, and the bars are the corresponding errors (right y-axis) of two NN potentials to DFT calculations.

Figure 3B shows that the long-range MB correction effectively improves the prediction of structures with long-range charge transfer, as reflected by the better accuracy of **mb2** than **std**. Furthermore, the comparison between **mb1** and **mb2** indicates that the improvement is not simply due to an overall increase in accuracy but is specifically related to the long-range description. Although the introduction of short-range MB corrections in **mb1** also improves the overall performance of the training, it does not improve the prediction of these two defective structures. Interestingly, the **mb3** potential with multiple MB corrections does provide the best performance, suggesting the beneficial coupling between different MB correction terms.

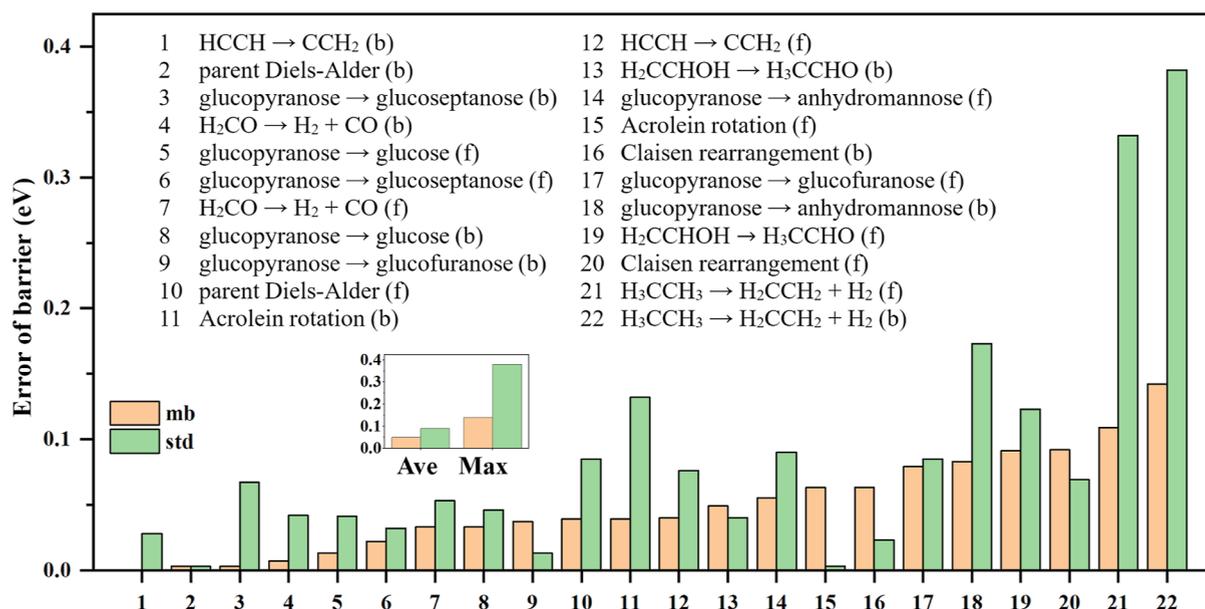
We may understand the results in Figure 3 as follows. In defective structures, the presence of the vacancies for the P-termini changes not only the local electron density but also the charge density at the conduction band minimum and the valence band maximum, which are of a long-range nature. The standard AC-MLP cannot decompose these defect-related interactions among all atoms properly due to the cutoff range of the local environment. As a result, atoms that are distal to the defects are slow to respond and thus are underestimated in energy (the long-range electrostatic interaction is erroneously reduced in defective systems). In contrast, the long-range MB functions allow atoms to “see” directly the long-range defects and respond quickly in energy and force.

**3.3. Reaction Prediction.** The high-energy regions on PES are inevitably less visited during PES exploration owing to both the Boltzmann law and the rapid expansion of the configurational space on going up in energy. For investigating chemical reactions, this could lead to a severe problem for MLP to predict reaction kinetics since the reaction barrier can be poorly described when the accuracy in describing the initial and transition states is inconsistent. To address this issue, one possible approach is to develop a specified MLP for the target reaction. In such an approach, the data set is expanded to include as many as possible data points that cover the entire reaction. For example, Rivero et al. has shown that around 200,000 structures are required to train a MLP to reproduce a

single Diels–Alder reaction between 2,3-dibromo-1,3-butadiene and maleic anhydride.<sup>37</sup> Obviously, the number of data points required to accurately describe a specific reaction can be so large that it is not feasible for general-purpose reaction prediction for unknown reactions. In general, it is a key challenge to train a robust MLP for predicting reaction energetics with relatively fewer data points on PES as practiced in our group using SSW global PES sampling.<sup>38</sup> Below we will show that the introduction of MB corrections does improve MLP in predicting the reaction transition regions using the Claisen rearrangement reaction of allyl vinyl ether to 4-pentenal as the example.

Figure 4A shows two data sets, each with 10,000 structures collected from our SSW and MD simulations (500 K), respectively, projected onto a two-dimensional surface defined by C–O distance (distance between C3 and O1, inset in Figure 4A) and C–C distance (distance between C1 and C5). It is obvious that the SSW trajectory (red) is much more effective in generating reactive data than MD data (black), although the transition state region data are still much less populated compared to the minima regions. It is worth mentioning that no further biased data selection is implemented to better sample the Claisen reaction since the purpose of this example is to illustrate the performance of MLP in a typical situation of global PES exploration, where in a global PES data set the transition state region data is often much less in density.<sup>39</sup>

We then trained the MLP for the MD data set and the SSW data set. Not surprisingly, the MLP trained from the MD data set cannot be utilized for predicting the reaction due to the complete lack of reaction data in the MD data set—the transition state search fails, suggesting a poor definition of the reaction PES. For MLPs trained based on the SSW data set, the standard NN potential utilizes a network size 291-64-86-1 and the MBNN potential has the equivalent network size of 291-64-64-23 with additional 22 many-body functions [13 two-body, 6 three-body, and 3 four-body functions (see Table S3 for many-body function details)]. Apparently, a large number of many-body functions are required to provide the



**Figure 5.** MBNN (mb) performance for predicting the reaction barrier of 22 organic reactions in comparison with that of standard NN (std). The names of the reactions are inserted in the graph, where (f) and (b) indicate the forward and backward reaction, respectively. The reactions are ranked by the error of the MBNN results with respect to the DFT results. The value and maximum value of the error of MBNN and standard NN potential are inserted.

best extrapolation to the transition state region since the training data set population for the transition state (higher energy, low data density) and the minima region (lower energy, high data density) is about 500 and 9500, respectively.

The performances of two MLPs are then compared by evaluating the pathway of the Claisen reaction. The results are shown in Figure 4B, which utilize 31 structures from DFT calculations along the reaction channel as the benchmark. The MAE of standard NN and MBNN are 2.58 and 1.42 meV/atom, and the RMSE are 4.65 and 2.11 meV/atom, respectively (see Table S4). Obviously, the overall accuracy in energy of MBNN is significantly improved, particularly in the transition state region of the reaction (data point 14–20) compared to standard NN, where the RMSE of standard NN is 8.62 meV/atom and MBNN is 3.40 meV/atom. In addition, the force prediction is also generally better in the mb-SSW potential. We note that the larger error at the transition state region compared to the minima region, particular for the standard NN, is expected due to the low data density at the transition state. The same phenomenon was also observed in previous works.<sup>2,38</sup>

To illustrate the general improvement of G-MBNN in activity prediction for a wide range of reactions, we further generated a C–H–O three-element G-MBNN potential and compared its performance with the standard G-NN potential. The training data set used in fitting the MLPs is a subset of global PES data taken from our previous work<sup>38,40</sup> in exploring the full reaction network of glucose pyrolysis, which contains 53,727 structures with up to C–H–O three elements (the data set is openly available<sup>41</sup>). These structures were collected from SSW global optimization trajectories of QM9 molecules and glucose<sup>38</sup> with and without periodic boundary conditions. The network architecture of standard G-NN and G-MBNN has the equivalent network size, being 345-64-86-1 and 345-64-78-8, respectively, where G-MBNN has an additional 7 many-body functions (3 two-body, 3 three-body, and 1 four-body functions).

The performances of G-MBNN and standard G-NN in predicting the barriers of 22 organic reactions are shown in Figure 5, where the barriers are compared with the DFT calculation results. These reactions are the forward and backward reactions of 11 distinct reactions as listed in Figure 5 (all data are detailed in Table S5), which are taken from the Baker test system and glucose decomposition.<sup>42,43</sup> As shown in Figure 5, the G-MBNN outperforms standard G-NN with the average and maximum errors in the barrier being 0.05 and 0.14 eV, respectively, being significantly smaller than the corresponding errors of 0.09 and 0.38 eV obtained from standard G-NN potential. This confirms the better extrapolation of the G-MBNN potential to the transition state region of PES and thus provides higher accuracy for chemical reactions. It should be noted that the comparison here is based on the same, relatively small parameter space—when the network size increases, we observe that the difference in predicting the barriers between G-MBNN and standard G-NN becomes smaller, although G-MBNN is consistently better than standard NN.

#### 4. CONCLUSIONS

This work develops a new MLP architecture (G-MBNN) by incorporating explicit MB corrections in the derivation of the total energy. A set of two-body, three-body, and four-body MB functions are proposed as the MB functions, where the parameters are the output of atom-centered NN. Compared to the traditional atom-centered MLP, the G-MBNN provides an easy access framework to delicate energy terms between particular atoms, such as very short repulsion, long-range attraction, and angle-dependent covalent interactions. The G-MBNN is targeted to fit the global PES data set obtained from SSW global PES sampling, the same as the G-NN potential utilized in our previous work. By testing on a range of complex PES systems, we demonstrate that the G-MBNN potentials generally improve on the fitting speed, the description of long-range interactions, and a better description of chemical bond making/breaking. We emphasize that the G-MBNN architec-

ture offers a flexible ML framework, not only for energy and force (extensive property) prediction in MLP but also for intensive properties predictions, which can be done by modifying the MB function forms. The potential of the MBNN architecture for property predictions will be explored in our coming works.

## ■ ASSOCIATED CONTENT

### SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jctc.3c00873>.

Training information on Li–Co–O; training information on Ti–O; MB functions in fitting Claisen rearrangement; test results of Claisen rearrangement; and benchmark results for 11 distinct organic reactions (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Authors

**Cheng Shang** – Collaborative Innovation Center of Chemistry for Energy Material (iChem), Shanghai Key Laboratory of Molecular Catalysis and Innovative Materials, Key Laboratory of Computational Physical Science, Department of Chemistry, Fudan University, Shanghai 200433, China; [orcid.org/0000-0001-7486-1514](https://orcid.org/0000-0001-7486-1514); Email: [cshang@fudan.edu.cn](mailto:cshang@fudan.edu.cn)

**Zhi-Pan Liu** – Collaborative Innovation Center of Chemistry for Energy Material (iChem), Shanghai Key Laboratory of Molecular Catalysis and Innovative Materials, Key Laboratory of Computational Physical Science, Department of Chemistry, Fudan University, Shanghai 200433, China; Key Laboratory of Synthetic and Self-Assembly Chemistry for Organic Functional Molecules, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Shanghai 200032, China; Shanghai Qi Zhi Institution, Shanghai 200030, China; [orcid.org/0000-0002-2906-5217](https://orcid.org/0000-0002-2906-5217); Email: [zpliu@fudan.edu.cn](mailto:zpliu@fudan.edu.cn)

### Authors

**Pei-Lin Kang** – Collaborative Innovation Center of Chemistry for Energy Material (iChem), Shanghai Key Laboratory of Molecular Catalysis and Innovative Materials, Key Laboratory of Computational Physical Science, Department of Chemistry, Fudan University, Shanghai 200433, China; [orcid.org/0000-0003-2147-2472](https://orcid.org/0000-0003-2147-2472)

**Zheng-Xin Yang** – Collaborative Innovation Center of Chemistry for Energy Material (iChem), Shanghai Key Laboratory of Molecular Catalysis and Innovative Materials, Key Laboratory of Computational Physical Science, Department of Chemistry, Fudan University, Shanghai 200433, China

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.jctc.3c00873>

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

This work received financial support from the National Science Foundation of China (12188101, 22033003, 91945301, 91745201, 22122301, 92061112, and 92145302), the Fundamental Research Funds for the Central Universities

(20720220011), the National Key Research and Development Program of China (2018YFA0208600), and the Tencent Foundation for XPLOER PRIZE.

## ■ REFERENCES

- (1) Ko, T. W.; Finkler, J. A.; Goedecker, S.; Behler, J. General-Purpose Machine Learning Potentials Capturing Nonlocal Charge Transfer. *Acc. Chem. Res.* **2021**, *54* (4), 808–817.
- (2) Unke, O. T.; Chmiela, S.; Sauceda, H. E.; Gastegger, M.; Poltavsky, I.; Schütt, K. T.; Tkatchenko, A.; Müller, K. R. Machine Learning Force Fields. *Chem. Rev.* **2021**, *121* (16), 10142–10186.
- (3) Deringer, V. L.; Bartók, A. P.; Bernstein, N.; Wilkins, D. M.; Ceriotti, M.; Csányi, G. Gaussian Process Regression for Materials and Molecules. *Chem. Rev.* **2021**, *121* (16), 10073–10141.
- (4) Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Phys. Rev. Lett.* **2007**, *98* (14), 146401.
- (5) Smith, J. S.; Isayev, O.; Roitberg, A. E. ANI-1: An Extensible Neural Network Potential with DFT Accuracy at Force Field Computational Cost. *Chem. Sci.* **2017**, *8* (4), 3192–3203.
- (6) Huang, S.-D.; Shang, C.; Zhang, X.-J.; Liu, Z.-P. Material Discovery by Combining Stochastic Surface Walking Global Optimization with a Neural Network. *Chem. Sci.* **2017**, *8* (9), 6327–6337.
- (7) Huang, S.-D.; Shang, C.; Kang, P.-L.; Liu, Z.-P. Atomic Structure of Boron Resolved Using Machine Learning and Global Sampling. *Chem. Sci.* **2018**, *9* (46), 8644–8655.
- (8) Bartók, A. P.; Payne, M. C.; Kondor, R.; Csányi, G. Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons. *Phys. Rev. Lett.* **2010**, *104* (13), 136403.
- (9) Bartók, A. P.; Kondor, R.; Csányi, G. On Representing Chemical Environments. *Phys. Rev. B* **2013**, *87* (18), 184115.
- (10) Schütt, K. T.; Arbabzadah, F.; Chmiela, S.; Müller, K. R.; Tkatchenko, A. Quantum-Chemical Insights from Deep Tensor Neural Networks. *Nat. Commun.* **2017**, *8* (1), 13890.
- (11) Schütt, K. T.; Sauceda, H. E.; Kindermans, P.-J.; Tkatchenko, A.; Müller, K.-R. SchNet – A Deep Learning Architecture for Molecules and Materials. *J. Chem. Phys.* **2018**, *148* (24), 241722.
- (12) Zhang, Y.; Xia, J.; Jiang, B. Physically Motivated Recursively Embedded Atom Neural Networks: Incorporating Local Completeness and Nonlocality. *Phys. Rev. Lett.* **2021**, *127* (15), 156002.
- (13) Han, J.; Zhang, L.; Car, R.; E, W. Deep Potential: A General Representation of a Many-Body Potential Energy Surface. *Commun. Comput. Phys.* **2018**, *23*, 629.
- (14) Zhang, L.; Han, J.; Wang, H.; Car, R.; E, W. Deep Potential Molecular Dynamics: A Scalable Model with the Accuracy of Quantum Mechanics. *Phys. Rev. Lett.* **2018**, *120* (14), 143001.
- (15) Dékány, A. A.; Czako, G. Exploring the Versatile Reactivity of the F– + SiH<sub>3</sub>Cl System on a Full-Dimensional Coupled-Cluster Potential Energy Surface. *J. Chem. Phys.* **2023**, *158* (22), 224303.
- (16) Czako, G.; Bowman, J. M. Dynamics of the Reaction of Methane with Chlorine Atom on an Accurate Potential Energy Surface. *Science* **2011**, *334* (6054), 343–346.
- (17) Bowman, J. M.; Czako, G.; Fu, B. High-Dimensional Ab Initio Potential Energy Surfaces for Reaction Dynamics Calculations. *Phys. Chem. Chem. Phys.* **2011**, *13* (18), 8094–8111.
- (18) Behler, J.; Csányi, G. Machine Learning Potentials for Extended Systems: A Perspective. *Eur. Phys. J. B* **2021**, *94* (7), 142.
- (19) Behler, J. Atom-Centered Symmetry Functions for Constructing High-Dimensional Neural Network Potentials. *J. Chem. Phys.* **2011**, *134* (7), 074106.
- (20) Unke, O. T.; Meuwly, M. PhysNet: A Neural Network for Predicting Energies, Forces, Dipole Moments, and Partial Charges. *J. Chem. Theory Comput.* **2019**, *15* (6), 3678–3693.
- (21) Ko, T. W.; Finkler, J. A.; Goedecker, S.; Behler, J. A Fourth-Generation High-Dimensional Neural Network Potential with Accurate Electrostatics Including Non-Local Charge Transfer. *Nat. Commun.* **2021**, *12* (1), 398.

- (22) Ghasemi, S. A.; Hofstetter, A.; Saha, S.; Goedecker, S. Interatomic Potentials for Ionic Systems with Density Functional Accuracy Based on Charge Densities Obtained by a Neural Network. *Phys. Rev. B* **2015**, *92* (4), 045131.
- (23) Xie, X.; Persson, K. A.; Small, D. W. Incorporating Electronic Information into Machine Learning Potential Energy Surfaces via Approaching the Ground-State Electronic Energy as a Function of Atom-Based Electronic Populations. *J. Chem. Theory Comput.* **2020**, *16* (7), 4256–4270.
- (24) Zhang, L.; Wang, H.; Muniz, M. C.; Panagiotopoulos, A. Z.; Car, R.; E, W. A Deep Potential Model with Long-Range Electrostatic Interactions. *J. Chem. Phys.* **2022**, *156* (12), 124107.
- (25) Kang, P.; Shang, C.; Liu, Z. Recent Implementations in LASP 3.0: Global Neural Network Potential with Multiple Elements and Better Long-Range Description. *Chin. J. Chem. Phys.* **2021**, *34* (5), 583–590.
- (26) Huang, S.-D.; Shang, C.; Kang, P.-L.; Zhang, X.-J.; Liu, Z.-P. LASP: Fast Global Potential Energy Surface Exploration. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2019**, *9* (6), No. e1415.
- (27) Daw, M. S.; Foiles, S. M.; Baskes, M. I. The Embedded-Atom Method: A Review of Theory and Applications. *Mater. Sci. Rep.* **1993**, *9* (7–8), 251–310.
- (28) Kang, P.-L.; Shang, C.; Liu, Z.-P. Large-Scale Atomic Simulation via Machine Learning Potentials Constructed by Global Potential Energy Surface Exploration. *Acc. Chem. Res.* **2020**, *53* (10), 2119–2129.
- (29) Kovács, D. P.; Oord, C. v. d.; Kucera, J.; Allen, A. E. A.; Cole, D. J.; Ortner, C.; Csányi, G. Linear Atomic Cluster Expansion Force Fields for Organic Molecules: Beyond RMSE. *J. Chem. Theory Comput.* **2021**, *17* (12), 7696–7711.
- (30) Kresse, G.; Furthmüller, J. Efficiency of Ab-Initio Total Energy Calculations for Metals and Semiconductors Using a Plane-Wave Basis Set. *Comput. Mater. Sci.* **1996**, *6* (1), 15–50.
- (31) Kresse, G.; Joubert, D. From Ultrasoft Pseudopotentials to the Projector Augmented-Wave Method. *Phys. Rev. B* **1999**, *59* (3), 1758–1775.
- (32) Shang, C.; Liu, Z.-P. Stochastic Surface Walking Method for Structure Prediction and Pathway Searching. *J. Chem. Theory Comput.* **2013**, *9* (3), 1838–1845.
- (33) Zhang, P.; Shang, C.; Liu, Z.; Yang, J.-H.; Gong, X.-G. Origin of performance degradation in high-delithiation  $\text{Li}_x\text{CoO}_2$ : insights from direct atomic simulations using global neural network potentials. *J. Mater. Chem. A* **2023**, *11* (10), 5370–5379.
- (34) Nocedal, J. Updating Quasi-Newton Matrices with Limited Storage. *Math. Comput.* **1980**, *35* (151), 773–782.
- (35) Liu, D. C.; Nocedal, J. On the Limited Memory BFGS Method for Large Scale Optimization. *Math. Program.* **1989**, *45* (1–3), 503–528.
- (36) Ma, S.; Huang, S.-D.; Fang, Y.-H.; Liu, Z.-P. TiH Hydride Formed on Amorphous Black Titania: Unprecedented Active Species for Photocatalytic Hydrogen Evolution. *ACS Catal.* **2018**, *8* (10), 9711–9721.
- (37) Rivero, U.; Unke, O. T.; Meuwly, M.; Willitsch, S. Reactive Atomistic Simulations of Diels-Alder Reactions: The Importance of Molecular Rotations. *J. Chem. Phys.* **2019**, *151* (10), 104301.
- (38) Kang, P.-L.; Shang, C.; Liu, Z.-P. Glucose to 5-Hydroxymethylfurfural: Origin of Site-Selectivity Resolved by Machine Learning Based Reaction Sampling. *J. Am. Chem. Soc.* **2019**, *141* (51), 20525–20536.
- (39) Gastegger, M.; Marquetand, P. High-Dimensional Neural Network Potentials for Organic Reactions and an Improved Training Algorithm. *J. Chem. Theory Comput.* **2015**, *11* (5), 2187–2198.
- (40) Kang, P.-L.; Shi, Y.-F.; Shang, C.; Liu, Z.-P. Artificial Intelligence Pathway Search to Resolve Catalytic Glycerol Hydrogenolysis Selectivity. *Chem. Sci.* **2022**, *13* (27), 8148–8160.
- (41) The dataset can be download from [http://www.lasphub.com/supportings/CHO\\_train.tgz](http://www.lasphub.com/supportings/CHO_train.tgz) (accessed Aug 8, 2023).
- (42) Shang, C.; Liu, Z.-P. Constrained Broyden Minimization Combined with the Dimer Method for Locating Transition State of Complex Reactions. *J. Chem. Theory Comput.* **2010**, *6* (4), 1136–1144.
- (43) Zhang, X.-J.; Liu, Z.-P. Reaction Sampling and Reactivity Prediction Using the Stochastic Surface Walking Method. *Phys. Chem. Chem. Phys.* **2015**, *17* (4), 2757–2769.