JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

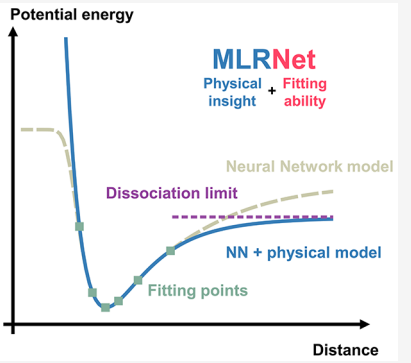
MLRNet: Combining the Physics-Motivated Potential Models with Neural Networks for Intermolecular Potential Energy Surface Construction

You Li, Yu Zhai, and Hui Li*

Cite This: J. Chem. Theory Comput. 2023, 19, 1421−1431 Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: A physics-based machine learning model called MLRNet has been developed to construct the high-accuracy two-body intermolecular potential energy surface (IPES). The outputs of the neural network are integrated into the physically realistic Morse/long-range (MLR) function, which ensures that the MLRNet has meaningful extrapolation at both short and long ranges and solves the asymptotic problem in common neural network potential (NNP) models. The neural network representation of the MLR parameters is more flexible and more efficient than the polynomial expansion in the conventional mdMLR model, especially for systems containing nonrigid monomer(s). The present work illustrates the basic framework of the current MLRNet model, including (i) how to combine the physically meaningful MLR function with different possible NN structures, (ii) the preservation of permutation symmetry, and (iii) the predetermination of the long-range function u_{LR}. We choose two realistic systems to demonstrate the performance of MLRNet: the three-dimensional IPES of CO_2−He including the CO_2 antisymmetric vibration Q_3 and the six-dimensional IPES of the H_2O−Ar system. In both cases, the fitting errors of the MLRNet are several times smaller than those of the conventional mdMLR model. Both short-range and long-range extrapolation tests were performed to illustrate the extrapolation ability of the MLRNet and its damping function version. Moreover, for the 6-D H_2O−Ar system, the MLRNet only needs 1596 trainable parameters, which is almost equal to the number needed for the 5-D mdMLR model (1509) and half that needed for the PIP-NN model (3501) within similar accuracy, which illustrates the model efficiency in high-dimensional IPES fitting.



1. INTRODUCTION

The potential energy surface (PES) is fundamentally important in atomic-scale computational studies, and a reliable PES is one of the prerequisites for chemical dynamics simulations. Although the PES can be obtained by direct electronic structure calculations during dynamic simulations, the analytical PES constructed by fitting or interpolation of some single-point energies (SPEs) is much more efficient. Traditionally, PES fitting has been based on physically meaningful potential functions. These models are often limited by their functional forms and only have good numerical accuracy for specific systems. Nowadays, many machine learning (ML) PES models, especially neural network (NN) ones, have been developed and become more and more popular due to their convenience in the application and universality.[1−3] Many high-accuracy ML PESs have been constructed for small-molecule quantum dynamics simulations and vibrational spectroscopy calculations.[3−8]

When dealing with a nonbonded system, the total potential energy is usually split into two components: the intramolecular potential for the isolated molecule monomers and the intermolecular potential energy surface (IPES) that describes the interaction energy between molecules. The separation of pure intermolecular interactions provides additional physical insight. For instance, the long-range interaction can be divided into three contributions: electrostatic, induction, and dispersion, all of which can be accurately approximated by a series in inverse powers of distance R based on perturbation theory.[9,10] Compared to physics-motivated models, the general machine learning PES models usually need more fitting points to "learn" the correct asymptotic behavior and have poor extrapolation outside the dataset region.[3−5] A common solution is to employ a switch function to linearly connect the ML potential and predetermined asymptotic potential.[11,12] However, this switch model is sensitive to the crossover position, and the energy around the switch region would still require data. Another approach is to integrate the ML model

Received: October 22, 2022
Published: February 24, 2023

into a certain physics-motivated function, e.g., fitting the physical parameters by a neural network.[13−18] Then the problem becomes how to find a robust, flexible, and asymptotic sensible physical function for the simulated system. If the physical models are not suitable to represent the global interaction energy, we still need additional effort to compromise the deviation between the given physical function and the true ab initio energies.

The recently developed Morse/long-range (MLR) model is a physics-motivated potential model for atom−atom interactions[19−22] that is flexible enough for midrange fitting while extrapolating in physically realistic behavior in the long- and short-distance regions.[22,23] Its extended version, the multi-dimensional MLR (mdMLR) model,[24] has achieved success in fitting IPESs for atom−molecule[25−28] and molecule−molecule[29−32] systems within spectroscopic accuracy and retains the flexibility and sensible extrapolation that have been illustrated in dynamic simulations.[33−36] As with many other model potentials, the mdMLR model suffers from high-dimensional fitting problems, especially when intramolecular degrees of freedom are included.[24] One efficient way is to average over the intramolecular vibrations under the weak coupling condition. However, this method needs direct product grids of ab initio points, and thus, the calculation cost would increase exponentially with the number of intramolecular degrees of freedom considered.

In the previous studies, the physically meaningful parameters in the mdMLR model are expanded in terms of angular basis functions (e.g., orthogonal polynomials) or their direct products with basis functions of intramolecular coordinates if intramolecular motions are taken into consideration explicitly. In fact, however, there is no constraint on the representation of the MLR parameters, and it is feasible to employ other mathematical approaches for MLR potential fitting. As mentioned above, the neural network model is efficient in high-dimensional fitting and can be integrated into other models easily, but it lacks physical knowledge in the asymptotic region. Thus, a combination of the mdMLR model and an NN PES model can become a promising method having both good high-dimensional fitting ability and reliable asymptotic behavior. From the view of the neural network model, the MLR function works like a pretrained layer(s) with frozen weights and bias, which connects the NN outputs (those physical parameters) and the final target interaction energies. As a result, the neural network should approach the target function more easily if the physics-inspired MLR function itself matches the real IPES well.

In the present work, we aim to construct a physics-based machine learning model by integrating the neural network model into the MLR function. Meanwhile, the conventional mdMLR format is modified to a more efficient representation. The final model, named MLRNet, can represent high-dimensional IPESs with explicit intramolecular degrees of freedom like other ML models while preserving physically sensible asymptotic behavior like the MLR model. If proper high-dimensional NN PES models are chosen,[17,18] MLRNet would be promising for fitting the IPESs of large dimer systems. In order to test the fitting ability of MLRNet, several IPESs for the chosen two systems have been constructed. The first system is the three-dimensional IPES of CO_2−He including the CO_2 antisymmetric mode. The previous mdMLR model achieved a root-mean-square deviation (RMSD) of 0.032 cm^{-1} with only 55 trainable parameters, and its long-range behavior played an important role in dynamic simulation.[33] Thus, it is a good reference model to check both the fitting and long-range extrapolation accuracy. The second one is the six-dimensional IPES of H_2O−Ar including all H_2O normal modes, which remains hard for conventional mdMLR fitting. Also, there is a very new machine learning IPES of H_2O−Ar constructed using the permutationally invariant polynomials neural network (PIP-NN) method.[37,38] We will demonstrate that MLRNet can determine the high-dimensional systems even more accurately and efficiently than the vibrationally averaged method mdMLR model and general PIP-NN and obtain an accurate short-range potential with an appropriate damping function and uncertainty (weight) function.

This paper is organized as follows. Section 2 describes the idea of the mdMLR model and the basic framework of MLRNet, including the possible network structure for MLR parameter fitting, the strategy for permutation symmetry, and the effect of the long-range function. Section 3 discusses the performance of MLRNet for two realistic systems. The results are compared to the conventional mdMLR model and PIP-NN. Finally, the conclusion and outlook of the MLRNet model are given in section 4.

2. METHODOLOGY

Before discussing MLRNet, it is worthwhile to understand some basic concepts of the mdMLR model, and a more comprehensive discussion can be found in our recent review.[24] The intermolecular interaction between two molecules A and B can be written as

$$\Delta V_{AB}(R, \mathbf{\Omega}, \mathbf{Q}) = V_{tot}(R, \mathbf{\Omega}, \mathbf{Q}) - V_A(\mathbf{Q}_A) - V_B(\mathbf{Q}_B) \tag{1}$$

where $\mathbf{Q}_A$, $\mathbf{Q}_B$ and $\mathbf{Q} \equiv (\mathbf{Q}_A, \mathbf{Q}_B)$ represent the intramolecular configuration, R is distance between the two molecular centers of mass (COMs), and $\mathbf{\Omega}$ denotes the relative orientation of the two molecules. V_A, V_B, and V_{tot} are the potential energies of the isolated A and B molecules and the total energy of the dimer, respectively. The central *ansatz* of the mdMLR model is the IPES can be approximated by the following Morse-like potential formula:

$$V_{mdMLR}(R, \mathbf{\Omega}, \mathbf{Q}) = \mathcal{D}_e(\mathbf{\Omega}, \mathbf{Q})\left\{1 - \frac{u_{LR}(R, \mathbf{\Omega}, \mathbf{Q})}{u_{LR}(R_e, \mathbf{\Omega}, \mathbf{Q})} e^{-\beta(R,\mathbf{\Omega},\mathbf{Q})\cdot y_p^{eq}(R,\mathbf{\Omega},\mathbf{Q})}\right\}^2 - \mathcal{D}_e(\mathbf{\Omega}, \mathbf{Q}) + V_{lim} \tag{2}$$

where V_{lim} is the asymptotic energy ($V_{lim} = \lim_{R\to+\infty} V_{mdMLR}$), which is usually set as zero, and $\mathcal{D}_e$ and R_e are the explicit well depth and equilibrium distance, respectively. The function u_{LR} represents the "long-range interaction" and can be written as

$$u_{LR}(R, \mathbf{\Omega}, \mathbf{Q}) = \sum_{i=1}^{last} D_{m_i}(R; s, \rho)\frac{C_{m_i}(\mathbf{\Omega}, \mathbf{Q})}{R^{m_i}} \tag{3}$$

in which the long-range coefficients $\{C_{m_i}\}$ are usually defined by perturbation theory or fitted in the long-range region and an appropriate damping function $D_{m_i}(R; s, \rho)$ can improve both the short-range and long-range behavior.[22] We will discuss the long-range function in more detail in section 2.3. $y_p(R, \mathbf{\Omega}, \mathbf{Q})$ is

https://doi.org/10.1021/acs.jctc.2c01049
J. Chem. Theory Comput. 2023, 19, 1421−1431

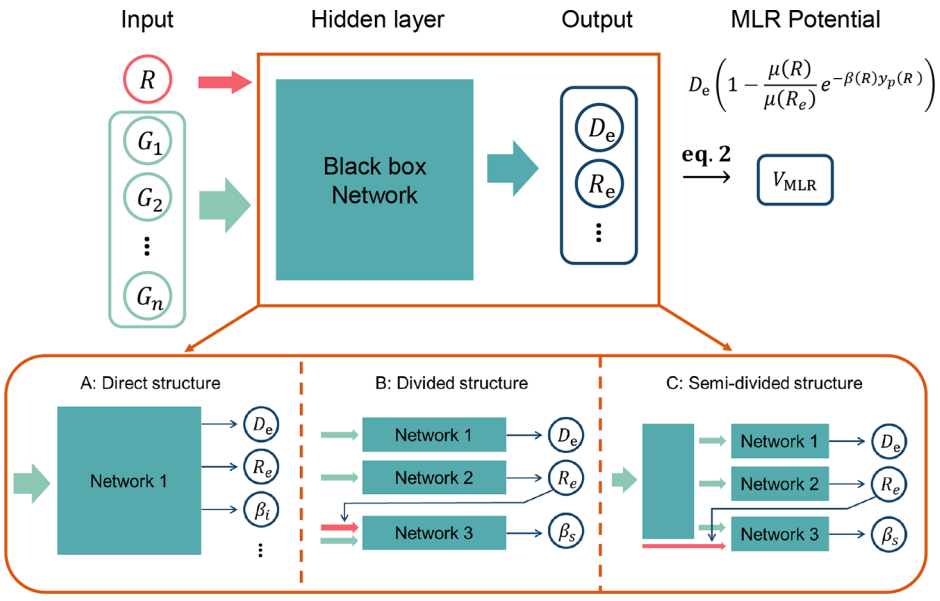


Figure 1. Basic MLRNet structures discussed in this work. The input vector is the COM distance R with PIPs/FI $G^{R=R_0}$. The output of MLRNet must contain at least these MLR parameters: the well depth $\mathcal{D}_e$, the equilibrium distance R_e, and the shape function β_s or a set of expansion coefficients $\{\beta_i\}$. Three NN architectures are discussed, while the detailed structures of the (sub)networks are displayed as "black boxes". The "direct structure" uses only one network to represent all of the MLR parameter functions directly. The second one, "divided structure", improves the representation of β and uses independent subnetworks to express the MLR parameters. If the subnetworks share some NN parameters, we get the third kind of structure, the "semi-divided structure". The "divided structure" is recommended for completely independent MLR parameter functions and is the only structure employed in this work.

a dimensionless exponent radial function which transforms distance $R \in [0, +\infty]$ to $y_p \in [-1, 1]$:

$$y_p^{\text{ref}}(R, \mathbf{\Omega}, \mathbf{Q}) = \frac{R^p - R_{\text{ref}}^p(\mathbf{\Omega}, \mathbf{Q})}{R^p + R_{\text{ref}}^p(\mathbf{\Omega}, \mathbf{Q})} \tag{4}$$

where the parameter R_{ref} is conventionally expressed as a user-defined constant f_{ref} multiplied by R_e and y_p^{eq} means $R_{\text{ref}} = R_e$. The integer power p must be greater than $(m_{\text{last}} - m_1)$ for proper long-range behavior. Finally, the exponent coefficient function β is usually written as

$$\beta(R, \mathbf{\Omega}, \mathbf{Q}) = y_p^{\text{ref}}(R)\beta_\infty + [1 - y_p^{\text{ref}}(R)]\sum_{i=0}^{N_\beta} \beta_i(\mathbf{\Omega}, \mathbf{Q}) (y_q^{\text{ref}}(R))^i \tag{5}$$

where $\beta_\infty = \lim_{R\to\infty} \beta = \ln[2\mathcal{D}_e/u_{\text{LR}}(R_e)]$ is the asymptotic value of β and the $\{\beta_i\}$ are power expansion coefficients.

In the conventional mdMLR model, these physically meaningful functions are expressed as sums of direct products of orthogonal basis functions:

$$Y(\mathbf{\Omega}, \mathbf{Q}) = \sum_{\substack{i_1,i_2,\ldots,i_m \\ j_1,j_2,\ldots,j_n}} c_{i_1 i_2 \ldots i_m j_1 j_2 \ldots j_n} f_{Q_1}^{(i_1)} f_{Q_2}^{(i_2)} \ldots f_{Q_m}^{(i_m)} g_{\Omega_1}^{(j_1)} g_{\Omega_2}^{(j_2)} \ldots g_{\Omega_n}^{(j_n)} \tag{6}$$

where $Y(\mathbf{\Omega}, \mathbf{Q})$ represents one of the possible MLR parameter functions $\mathcal{D}_e(\mathbf{\Omega}, \mathbf{Q})$, $R_e(\mathbf{\Omega}, \mathbf{Q})$, and $\beta_i(\mathbf{\Omega}, \mathbf{Q})$ and $f_{Q_k}^{(i_k)}$ and $g_{\Omega_l}^{(j_l)}$ are expansion basis functions for normal mode Q_k and angles Ω_l. Although this basis expansion method performed well in previous studies, it has at least the following shortcomings: (i) The size of the expansion coefficients/basis grows too fast when the intramolecular degrees of freedom are involved. (ii) The basis set needs to be chosen empirically. An inappropriate basis set will limit the fitting capability and hence reduce the fitting accuracy. (iii) The rigorous permutation symmetry of IPES is ensured by the constraints on the special term of the manual-determined expansion basis. When the numbers of identical atoms and permutation-equal configurations increase, it is hard to distinguish and eliminate all of these symmetry-avoiding expansions.

In order to go beyond the conventional mdMLR model, we propose a neural network representation of the mdMLR potential. The basic idea is to build (sub)networks to represent physically meaningful MLR parameters in the mdMLR model and transform the NN output to the predicted energy based on eq 2.

2.1. Neural Network Representation of the mdMLR Function. There are many kinds of neural network models nowadays for high-dimensional function fitting or other problems. However, all of these models can be regarded as "black box" functions: $\mathbf{Y} = \mathbf{f}_{\text{NN}}(\mathbf{X})$ where $\mathbf{X}$ and $\mathbf{Y}$ are the input and output vectors of the neural network, respectively. In this section, we only focus on the neural network representation of the mdMLR potential displayed in eq 2 and simplify any (sub)network as $\mathbf{f}_{\text{NN}}$. Theoretically, all of the MLR parameters can be approximated by the NN, but the long-range function u_{LR} is usually defined in advance, and the integer powers p and q are user-defined hyperparameters. Thus, only $\mathcal{D}_e(\mathbf{\Omega}, \mathbf{Q})$, $R_e(\mathbf{\Omega}, \mathbf{Q})$ (sometimes including the reference distance R_{ref}), and $\{\beta_i(\mathbf{\Omega}, \mathbf{Q})\}$ need to be optimized in global energy fitting. It is straightforward to use different elements in a single network output vector to represent these MLR parameters (named "direct structure"):

https://doi.org/10.1021/acs.jctc.2c01049
J. Chem. Theory Comput. 2023, 19, 1421−1431

$$\mathbf{f}_{\mathrm{NN}}(\mathbf{\Omega}, \mathbf{Q}) = \begin{pmatrix} \mathcal{D}_e \\ R_e \\ \beta_1 \\ \cdots \\ \beta_{N_\beta} \end{pmatrix} \tag{7}$$

and then directly substitute these NN outputs with the independent variable R into eq 2 to calculate the mdMLR energy. This kind of representation can be more efficient if we notice that the sum in eq 5 is a polynomial expansion of the β_s function,

$$\beta_s(y_q^{\mathrm{ref}}(R), \mathbf{\Omega}, \mathbf{Q}) \approx \sum_{i=0}^{N_\beta} \beta_i(\mathbf{\Omega}, \mathbf{Q})(y_q^{\mathrm{ref}}(R))^i \tag{8}$$

and we can directly use a neural network to approximate one β_s function instead of several expansion coefficient functions $\{\beta_i\}$. This strategy also improves the total function representation capability since the β function determines the potential shape and the β_s network can directly approximate function shape information instead of fitting polynomial expansion coefficients in high order. Then the corresponding improved NN representation (named "divided structure") can be written as

$$\mathbf{f}_{\mathrm{NN}}(R, \mathbf{\Omega}, \mathbf{Q}) = \begin{pmatrix} f_{\mathrm{NN}}^{(1)}(\mathbf{\Omega}, \mathbf{Q}) \\ f_{\mathrm{NN}}^{(2)}(\mathbf{\Omega}, \mathbf{Q}) \\ f_{\mathrm{NN}}^{(3)}(y_q^{\mathrm{ref}}(R), \mathbf{\Omega}, \mathbf{Q}) \end{pmatrix} = \begin{pmatrix} \mathcal{D}_e \\ R_e \\ \beta_s \end{pmatrix} \tag{9}$$

where $f_{\mathrm{NN}}^{(1)}(\mathbf{\Omega}, \mathbf{Q})$ and $f_{\mathrm{NN}}^{(2)}(\mathbf{\Omega}, \mathbf{Q})$ denote subnetworks for $\mathcal{D}_e$ and R_e, respectively, and the third element $f_{\mathrm{NN}}^{(3)}(y_q^{\mathrm{ref}}(R), \mathbf{\Omega}, \mathbf{Q})$ represents a special subnetwork for β_s, where one of inputs is the radial function $y_q^{\mathrm{ref}}(R)$ that is dependent on the R_e calculated by the second subnetwork $f_{\mathrm{NN}}^{(2)}(\mathbf{\Omega}, \mathbf{Q})$ (see "divided structure" in Figure 1).

In practice, the three subnetworks in eq 9 can share some parameters or layers performed as a "structure feature transformer", which could be named the "semi-divided structure". However, the "divided structure" is more convenient to control the different MLR parameters, which means the weights and biases in different subnetworks can have different learning rates during the training step. If we only want to optimize some of the MLR parameter subnetworks, the others can be totally frozen. Thus, the divided structure is more recommended in practice. A summary of the MLRNet structure discussed above can be found in Figure 1. In Applications and Discussion, the MLRNet structures are always divided structures, and the subnetworks are fully connected neural networks including two hidden layers:

$$f_{\mathrm{NN}}^{(i)}(\mathbf{X}) = \mathbf{b}_3 + \mathbf{W}_3 \cdot f_2(\mathbf{b}_2 + \mathbf{W}_2 \cdot f_1(\mathbf{b}_1 + \mathbf{W}_1\mathbf{X})) \tag{10}$$

where f_1 and f_2 are both hyperbolic tangent functions for nonlinear activation and $\mathbf{W}_i$ and $\mathbf{b}_i$ are the weights and biases in the corresponding hidden layers.

2.2. Permutation Symmetry. Rigorous permutation symmetry is important in PES fitting. The inputs of the mdMLR model are the distance between molecular COMs R, the intramolecular normal modes $\mathbf{Q}$, and the angular coordinates $\mathbf{\Omega}$ that describe monomer rotation. Obviously, only the distance R is a permutation-invariant variable, while the others are usually not. One of the most important tasks for MLRNet model construction is to find proper descriptors and NN structure to ensure the permutation symmetry while the variable R remains an explicitly independent degree of freedom. Here, we propose a simple but efficient descriptor $X = (R, G^{R_0})$, where G^{R_0} denotes the R-independent permutationally invariant polynomials (PIPs)[39−41] or fundamental invariants (FIs)[42,43] calculated in the $R = R_0$ configuration:

$$G^{R_0} = \hat{S}\prod_{i<j}^{N} p_{ij}^{l_{ij}}(R_0, \mathbf{\Omega}, \mathbf{Q}) \tag{11}$$

where p_{ij} is a monomial of the atomic distance r_{ij} and $\hat{S}$ is the symmetrization operator that contains all possible permutation operations. For the permutation-equal configurations, their R-independent PIPs G^{R_0} must be the same, hence preserving the permutation invariance of the given system. It should be noticed that the intermolecular potential in eq 1 is more suitable for the nonreaction system where no atom permutation occurs between two molecules. Thus, the total number of PIPs can be reduced by only considering the intramolecular permutation and the whole exchange of two identical molecules when applying the operator $\hat{S}$.[44]

Although the PIP and FI methods provide a complete description of the system, the number of invariant polynomials increases very rapidly for a large system. Another popular NN potential model, the so-called high-dimensional NN potential (HDNNP),[17,18] can handle permutation symmetry for a large number of identical atoms. It is very easy to use any HDNNP model to construct MLRNet only if the separation R is fixed while calculating the HDNNP descriptors. The summation of HDNNP outputs is permutation-invariant, which can directly represent any MLR parameter Y:

$$Y(\mathbf{X}) = \sum_{i}^{N_{\mathrm{atoms}}} y_i(\mathbf{X}) \tag{12}$$

where $\mathbf{X}$ denotes the HDNNP inputs and y_i is the sublayer output for the ith atom. In this work, only the PIP and FI methods were tested, and the HDNNP version of MLRNet is still under development.

2.3. Long-Range Function u_{LR}. The mdMLR potential has a physically realistic extrapolation when the long-range function u_{LR} is chosen appropriately. It has been illustrated[19,22] that for an integer power $p > m_{\mathrm{last}} - m_1$, the long-range behavior of MLR potential would be

$$V_{\mathrm{MLR}}(R \to +\infty) \simeq -u_{\mathrm{LR}}(R) + V_{\mathrm{lim}} \tag{13}$$

and the short-range limit is

$$V_{\mathrm{MLR}}(R \to 0) \simeq \mathcal{D}_e\left\{\frac{u_{\mathrm{LR}}(R)}{u_{\mathrm{LR}}(R_e)}\right\}^2 \exp(2\beta(0)) - \mathcal{D}_e + V_{\mathrm{lim}} \propto \{u_{\mathrm{LR}}(R)\}^2 \tag{14}$$

The short-range limit of the original MLR version (no damping function involved) will be proportional to the square of the highest-order inverse-power term $R^{2m_{\mathrm{last}}}$ and become physically wrong. However, with a reasonable damping function $D_{m_i}(R; s, \rho)$, the short-range limit becomes R^{2s}.

These results can be extended to the multidimensional version in eq 2 since they always hold for every one-dimensional curve along distance R. For a given dimer system,

https://doi.org/10.1021/acs.jctc.2c01049
J. Chem. Theory Comput. 2023, 19, 1421−1431

the long-range intermolecular interaction can be approximated with multipole expansions,[9,10] and the long-range coefficient $C_{m_i}(\mathbf{\Omega}, \mathbf{Q})$ in eq 3 can be determined by fitting long-range energies or directly obtained from experimental data. However, the short-range limit is more complicated and usually needs some hyperparameter test for correct asymptotic behavior. The recommend value is −1 for s and $\rho_{AB} = 2\rho^A\rho^B/(\rho^A + \rho^B)$ for ρ, where $\rho^{A/B} = (I^{A/B}{}_p/I^H{}_p)$ is the ratio of the ionization potential of molecule/atom A or B to that for a ground-state hydrogen atom. Following previous mdMLR potential models, the damping function that controls the short-range limit is assumed to be independent of $(\mathbf{\Omega}, \mathbf{Q})$. Thus, the hyperparameters are determined by comparing the performance of different (s, ρ_{AB}) combinations in a 1-D potential energy curve (PEC) $V(R; \mathbf{\Omega} = \mathbf{\Omega}_0, \mathbf{Q} = \mathbf{Q}_0)$ fitting test with the betaFIT program.[45]

3. APPLICATIONS AND DISCUSSION

3.1. CO_2−He System. The first study of the mdMLR potential model is the three-dimensional CO_2−He IPES constructed by Li and Le Roy,[27] which we call "Li2008" below. This work not only determined the basic framework of the fitting and morphing procedure of the mdMLR model but also illustrated the high fitting accuracy and good performance in spectroscopic calculations of the corresponding IPES. The fitting result of 2832 points (including permutation-equal configurations) has a root-mean-square deviation (RMSD) of about 0.032 cm^{-1} with only 55 parameters. It is usually considered hard work to achieve such high fitting accuracy by a neural network, especially when limiting the number of NN trainable parameters.

In order to test the performance of MLRNet in the low-dimensional but high-accuracy-required system, the three-dimensional IPES of the CO_2−He system was constructed. The ab initio points were totally the same as those in the previous work for "Li2008" model training,[27] and the total grids were divided into the training set (90%) and the validation set (10%), while the early stopping algorithm was employed to avoid overfitting. The long-range function and hyperparameters were also consistent in "Li2008" model, where $R_{ref} = R_e$, $p = 3$, and $q = 3$. The damping function is not used in this case to stay with the same function formula as in the previous study. The MLRNet structure was the "divided structure" with fully connected networks discussed in the Methodology. As discussed in section 2, we use a reduced set of R-independent PIPs calculated in $R_0 \equiv \overline{R_e}$ to preserve the permutation symmetry:

$$G_1^{R_0} = r_{CO1}\cdot r_{CO2}$$
$$G_2^{R_0} = \frac{r_{HeO1} + r_{HeO2}}{2}$$
$$G_3^{R_0} = r_{HeO1}\cdot r_{HeO2} \quad (15)$$

where $\overline{R_e}$ = 3.62 Å is the average R_e of the previous model "Li2008" and r_{AB} is simply the distance between the two atoms A and B. It should be noted that the inputs of the neural network usually need to be normalized to $[-1, +1]$, and the maximum−minimum normalization method was employed to normalize R-independent PIPs in this work:

$$x_{normal} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (16)$$

where x_{min} and x_{max} are the minimum and maximum values, respectively, of the corresponding PIP in the total dataset. These normalized PIPs and distance R form the input of MLRNet. Empirically, it is important to set the default tensor type to double precision, use the second-order optimizer L-BFGS for NN parameter optimization, and weight the potential with the inverse of uncertainty $u_i \equiv u(V_i)$ just like conventional MLR fitting:

$$V_i' = V_i/u_i$$
$$u_i = \begin{cases} 1.0 & V < 0.0 \\ (V + V_{1/2})/V_{1/2} & V \geq 0.0 \end{cases} \quad (17)$$

where $V_{1/2}$ is constant during training and is 5.0 cm^{-1} in this case. The subscript "1/2" indicates that the weighted potential V' is equal to $V/2$ when $V = V_{1/2}$. The (dimensionless) RMSD that evaluates the fitting quality is defined as

$$\text{RMSD} \equiv \overline{dd} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[V'_{true}(\boldsymbol{x}_i) - V'_{pred}(\boldsymbol{x}_i)]^2} \quad (18)$$

The model implementation was based on PyTorch and the PyTorch-Lightning package.[46,47]

After some tests for hidden layer structures and other NN training hyperparameters, the final MLRNet architecture is "3−4−5−1" for $\mathcal{D}_e$ and R_e and "4−5−6−1" for β_s, where "n_{in}−n_1−n_2−n_{out}" denotes a neural network with n_i neurons in the ith hidden layer which fully connects to the next layer (if it exists). Although the total number of trainable parameters is 162, which is about triple that of the conventional mdMLR model, the RMSD is decreased by an order of magnitude: 0.005, 0.007, and 0.005 cm^{-1} for the training set, validation set, and total set, respectively. The error distributions of the MLRNet model and mdMLR model are displayed in Figure 2, where it can be seen that points of the MLRNet are more concentrated around zero, especially in the repulsive region. The better fitting accuracy of the MLRNet model is not surprising since the MLRNet model not only contains more fitting parameters but also takes the high-order expansion of β_i into account by our improved β network discussed in eq 8.

One of the significant advantages of the MLRNet is the physically realistic extrapolation, especially giving accurate long-range potential energy values outside the dataset. In order to validate the long-range accuracy and extrapolation ability, several MLRNet models were constructed using the dataset with distances $R < R_{cut}$ (divided into the training set and validation set during model training, too). Then the ab initio points $R \geq R_{test}$ were used to detect the accuracy in both the training and extrapolation regions. Figure 3 displays the error distribution of these models along distance R_{test}. The MLRNet models provides very good extrapolation accuracy and are much more accurate than the long-range function u_{LR} in every $R_{test} \leq 10.0$ Å, except when the size of the dataset is too small (R_{cut} = 4.4 Å), which shows that the MLRNet could provide more realistic extrapolation than the linear switch model (which almost relies on a long-range function after crossover points). The asymptotic behavior of the MLR function can be regarded as smooth variation along the distance R from the minimum to the asymptotic limit, and the information around the minimum is important in extrapolation. The incorrect extrapolation of the R_{cut} = 4.4 Å model is mostly caused by missing too many training data from the minimum of the 1-D

https://doi.org/10.1021/acs.jctc.2c01049
J. Chem. Theory Comput. 2023, 19, 1421−1431

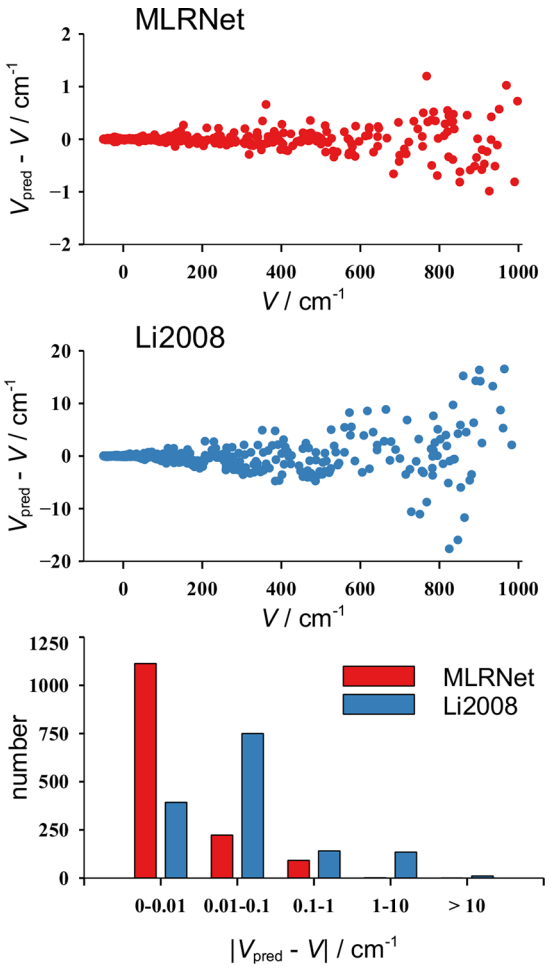


Figure 2. Fitting error distributions of the MLRNet model constructed in this work (red) and the conventional mdMLR model "Li2008" (blue) for the CO_2−He system. With more trainable parameters and a better representation of the β function, the MLRNet model achieves a much better error distribution. The maximum deviation is located in the short-range repulsive region and is still below 2 cm^{-1}, which is much lower than that of the previous model (about 20 cm^{-1}).

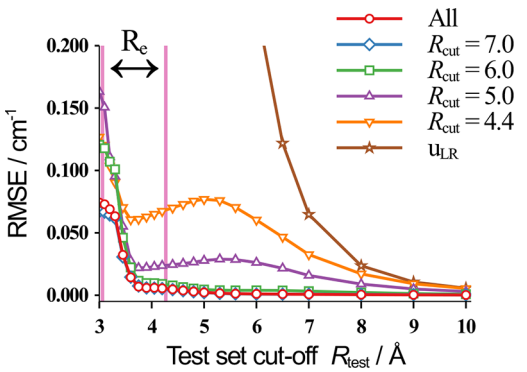


Figure 3. Root-mean-square error (energy in cm^{-1} and distance in Å) of the MLRNet models constructed with different datasets $R < R_{cut}$. The x axis is the distance cutoff of test points $R \geq R_{test}$ that illustrate the error distribution along R. For example, the blue line with diamond markers displays the error distribution of the model that was trained with $R < 7.0$ Å dataset, and its point at $R_{test} = 7.0$ Å is the extrapolation accuracy of the long-range region $R \geq 7.0$ Å. The two perpendicular pink lines describe the range of R_e. The incorrect extrapolation of the $R_{cut} = 4.4$ Å model is mostly caused by missing too many training points which are close to the R_e region.

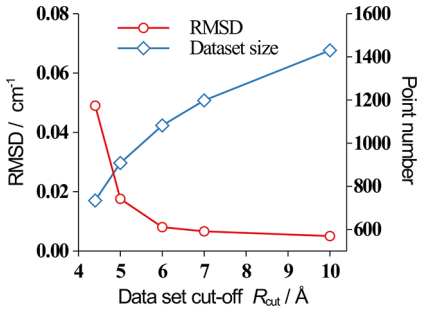


Figure 4. Total dataset RMSDs of different MLRNet models (red circles) and the sizes of the $R < R_{cut}$ datasets (blue diamonds) with respect to the dataset cutoff distance R_{cut}.

potential that cut along R (the maximum of R_e is close to 4.3 Å; see the R_e region in Figure 3) to the asymptotic region.

In Figure 4, the total dataset RMSDs of these models and the sizes of the corresponding distance-cutoff datasets are displayed. In the case of $R_{cut} = 5.0$ Å, even missing all of the ab initio points for which $R \geq 5.0$ Å, the MLRNet model still gives a good accuracy: RMSD = 0.027 cm^{-1} for all extrapolated points and RMSD = 0.017 cm^{-1} for the total dataset. Meanwhile, the dataset decreased by 36% while the RMSD of the total dataset increased by only 0.012 cm^{-1}, which illustrates the potential of the MLRNet model in grid saving. If a proper sampling procedure is chosen, especially the algorithm based on the potential function formula,[48] it promises to save much time in the ab initio calculation without significant decreases in accuracy.

3.2. H_2O−Ar System. The H_2O−Ar complex is an important model for studying the dynamics of all intra- and intermolecular vibrations. The previous mdMLR model, denoted here as "Hou2016", employed the vibration-averaged method to avoid the high-dimensional fitting problem caused by intramolecular coordinates, and the quality of the resulting state-specific PESs were testified by further ro-vibration calculations.[49] The three-dimensional mdMLR fitting to the 442 vibrationally averaged points for the (0,0,0) state achieved an RMSD of 0.121 cm^{-1} with only 58 parameters. Very recently, a full-dimensional IPES of the H_2O−Ar complex, denoted here as "Zhou2022", was constructed by the PIP-NN method,[38] and the fitting RMSE was 0.284 cm^{-1} for potential energy up to 10000 cm^{-1}. In this section, different MLRNet fitting strategies for constructing a full-dimensional IPES of the H_2O−Ar system will be discussed. The first subsection provides a comparison of the MLRNet model and the mdMLR model ("Hou2016"), and the next two subsections introduce the damping function version, denoted as MLRNet-D, and illustrate the role of the uncertainty function for accuracy improvement, especially in the short-range region.

MLRNet without Damping Functions. We first constructed the 6-D IPES that does not consider the damping function, mostly to ensure the same condition as the previous model, "Hou2016". The ab initio points are those calculated in the previous study prior to vibrational averaging.[49] The hyperparameters ($p = 3$, $q = 3$) are equal to those of the previous work, but the reference distance R_{ref} is one of the outputs of the R_e subnetwork for better optimization. The R-independent FIs for the A_2BC system are chosen for rigorous permutation

https://doi.org/10.1021/acs.jctc.2c01049
J. Chem. Theory Comput. 2023, 19, 1421−1431

symmetry.[42,43] The weight function is still defined as in eq 17 with $V_{1/2}$ = 1.0 cm^{-1}, following that of "Hou2016".

The other NN training details are similar to those of the CO_2−He system. After several tests, the final architecture of the MLRNet model is "7−17−17−1" for $\mathcal{D}_e$, "7−17−17−2" for R_e and R_{ref}, and "8−19−19−1" for β. The fitting RMSDs are 0.036, 0.040, and 0.037 cm^{-1} for the training set, validation set, and total set, respectively, which are even better than for the 3-D mdMLR model "Hou2016". The error distribution displayed in Figure 5 illustrates that the maximum deviation of

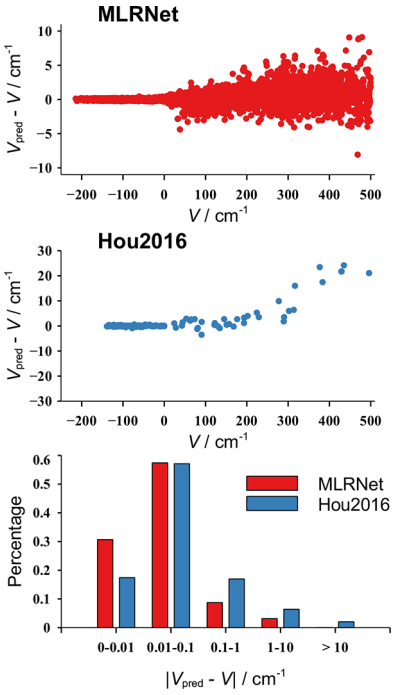


Figure 5. Fitting error distributions of the 6-D MLRNet model (red) and the vibrational-averaged 3-D mdMLR model "Hou2016" (blue) for the H_2O−Ar system. The maximum deviation of the MLRNet model is about 3 times smaller than that obtained with the mdMLR model.

the MLRNet model is about 3 times smaller than that obtained with the mdMLR model. More importantly, our 6-D MLRNet models (and the other MLRNet-D model discussed later) require fewer trainable parameters than both mdMLR and PIP-NN. The MLRNet only contains 1509 trainable parameters, while the 5-D mdMLR model for H_2O−N_2 has already required 1625 parameters[29] and the PIP-NN for the same system has 3501 NN parameters.[38] This result exactly shows the power of combining the physical potential model with machine learning. On the one hand, the physically meaningful function can help the neural network "understand" the function shape and reduce the number of trainable parameters, and on the other hand, the neural network can efficiently figure out the high-dimensional fitting problem and increase the accuracy.

MLRNet-D: MLRNet with Damping Functions. As discussed in Methodology, an appropriate damping function could provide physically realistic short-range extrapolation. In this subsection, we will illustrate it by the case of the H_2O−Ar system. Here the notation MLRNet-D denotes the MLRNet model that incorporates the damping function. The Tang−Toennies(TT) function[22,50] is chosen:

$$D_m^{TT}(R; s, \rho) = 1 - \exp[-b^{TT}(s)(\rho r)] \sum_{k=0}^{m-1+s} \frac{[b^{TT}(s)(\rho r)]^k}{k!} \quad (19)$$

and its hyperparameters $s = -1$, $\rho = 1.18$ were determined after one-dimensional PEC fitting tests using the betaFIT program,[45] while $b(s = -1) = 2.44$ is a system-independent constant.[22,51] Using the same training data and fitting procedure, MLRNet-D provides higher accuracy than MLRNet: the RMSDs are 0.028, 0.027, and 0.028 cm^{-1} for the training set, validation set, and total set, respectively. To further explore the model performance in a strongly repulsive extrapolation, 1500 additional short-range potential energies outside the original dataset were calculated by the same ab initio method as for "Hou2016".[49] The configurations of these points are determined by the product of three distance grids $\{R_i\}$ = {2.0 Å, 2.5 Å, 2.8 Å} and 500 $\{(\mathbf{\Omega}_i, \mathbf{Q}_i)\}$ grids that were randomly selected from the $\{(\mathbf{\Omega}_i, \mathbf{Q}_i)\}$ grids in the original total dataset. Figure 6 and Table 1 display the short-range

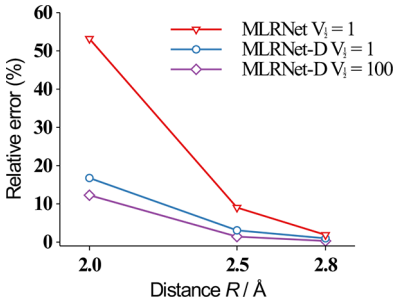


Figure 6. Extrapolation relative root-mean-square errors of the MLRNet model (red triangles), the MLRNet-D model with origin uncertainty (blue circles), and the MLRNet-D with lower repulsive-region uncertainty by increasing $V_{1/2}$ from 1.0 to 100.0 (purple diamonds). The damping function makes the short-range limit more reasonable, and larger $V_{1/2}$ increases the weight of short-range points, both of which increase the short-range extrapolation accuracy.

extrapolation errors of MLRNet and MLRNet-D. It is obvious that MLRNet-D performs better than MLRNet since the short-range limit of MLRNet-D is R^{-2}, which is much more sensible than that of MLRNet (R^{-8} in this case). The relative RMSE of MLRNet-D is below 16.76% for the dataset with R = 2.0 Å (where the mean energy is 29352 cm^{-1} and the maximum energy is 132138 cm^{-1}), which is a very acceptable extrapolation accuracy for an NN potential model that was trained below 500 cm^{-1}. In fact, the short-range extrapolation can further be improved by changing the uncertainty function and adding high-energy points, which will be discussed below.

Uncertainty Function for Better Repulsive Accuracy. The conventional MLR fitting normally employs an uncertainty function that increases extremely fast in the repulsive region, which makes the model focus on the potential well and accelerates convergence. However, it is important to achieve a small fitting error in the repulsive region if further dynamic simulation requires it. In this condition, we propose to decrease the uncertainty of short-range points by increasing $V_{1/2}$.

Two fitting tests of the H_2O−Ar system were performed to access the short-range accuracy of the high-$V_{1/2}$ strategy. In the first test, we still used the same ab initio points as for the total

https://doi.org/10.1021/acs.jctc.2c01049
J. Chem. Theory Comput. 2023, 19, 1421−1431

Table 1. Root-Mean-Square Errors of the Fitting Dataset and Extra-Short-Range Dataset[a]

	fitting RMSE		short-range RMSE (relative RMSE)		
model ($V_{1/2}$, E_{cut})	total set	$V < 0.0$	R = 2.0	R = 2.5	R = 2.8
MLRNet (1, 500)	0.421	0.039	14577.70 (53.17%)	381.23 (9.03%)	28.77 (1.89%)
MLRNet-D (1, 500)	0.279	0.019	5395.10 (16.76%)	139.25 (3.08%)	11.51 (0.98%)
MLRNet-D (100, 500)	0.059	0.031	4852.40 (12.26%)	115.09 (1.43%)	8.72 (0.30%)
MLRNet-D (2000, 10^4)	0.165	0.057	2142.42 (3.83%)	21.75 (0.16%)	0.52 (0.03%)

[a]Energy in cm^{-1} and distance in Å. ($V_{1/2}$, E_{cut}) denotes the uncertainty parameter and cutoff energy for the training dataset.

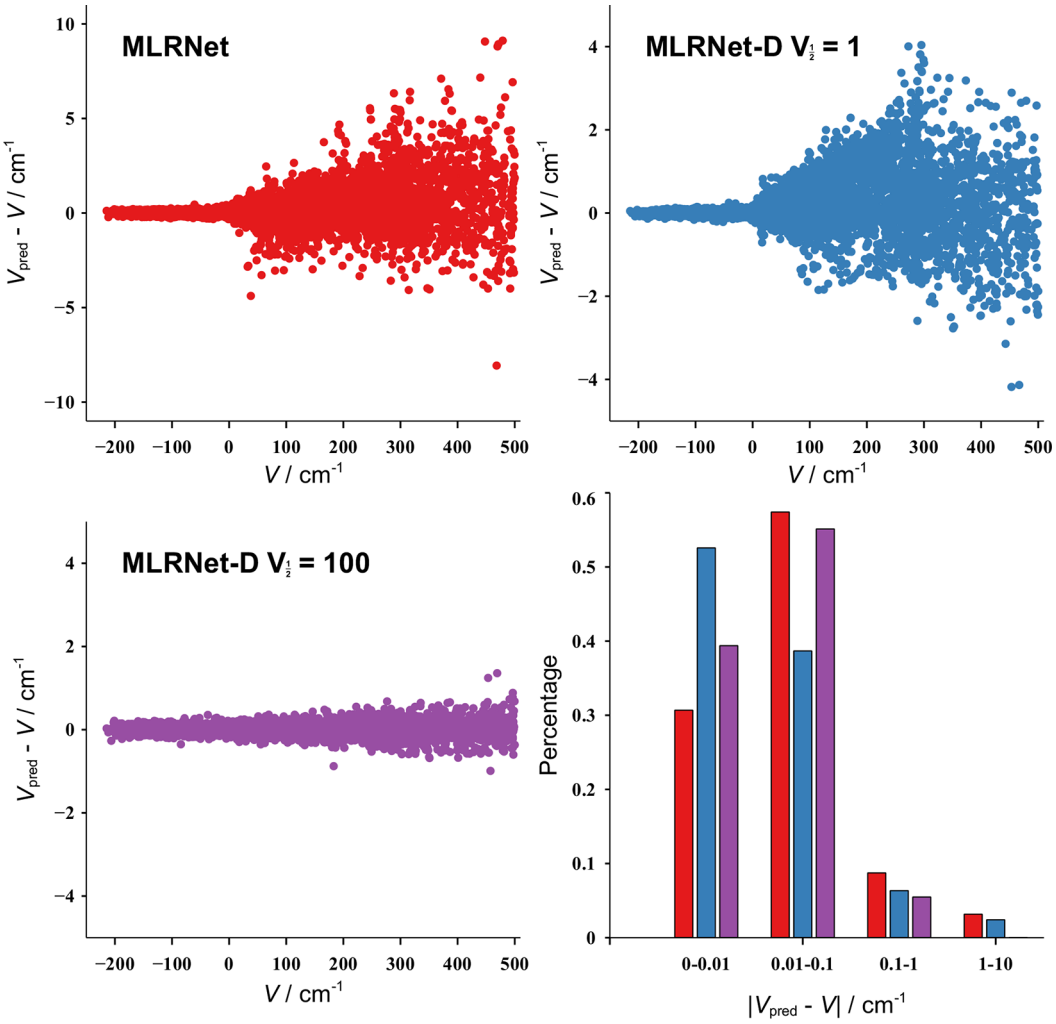


Figure 7. Error distribution of the MLRNet model (red) and the damping-function-version MLRNet-D models with different uncertainty parameters (blue, $V_{1/2}$ = 1; purple, $V_{1/2}$ = 100). The cutoff energies of the dataset, E_{cut}, are both 500 cm^{-1}.

dataset but changed $V_{1/2}$ from 1.0 to 100.0. In the other test, additional short-range points with energy less than 10000 cm^{-1} were included in the total dataset, and $V_{1/2}$ was changed to 2000.0. The uncertainty functions in both cases become 6.0 for maximum potential, which is similar to that of the PIP-NN model "Zhou2022" (u_i = 5.0 in 10000 cm^{-1}). Table 1 compares the fitting accuracies of those tests and the MLRNet/MLRNet-D results with the weight function. The unweighted RMSE of MLRNet-D significantly decreases from 0.279 to 0.059 cm^{-1} when $V_{1/2}$ is increased from 1.0 to 100.0. The error distributions displayed in Figure 7 more clearly demonstrate the short-range deviation. Moreover, the increased short-range fitting accuracy also makes extrapolation more reliable, as shown in Figure 6. In the second test, the RMSE (0.165 cm^{-1}) and the error distribution (displayed in Figure 8) of the MLRNet-D model are close to the PIP-NN results (RMSE = 0.284 cm^{-1}), while the relative error is only 3.83% for the R = 2.0 dataset. These results illustrate that the MLRNet-D model can achieve high fitting accuracy and sensible extrapolation in the repulsive region with high-$V_{1/2}$ uncertainty functions.

It should be noted that although a higher uncertainty parameter $V_{1/2}$ provides better short-range behavior, it harms the accuracy of the potential well region since the relative uncertainty of those points is decreased. This result can be demonstrated by comparing the fitting error of the $V < 0.0$ region in Table 1, where MLRNet-D with conventional uncertainty has a smaller error than the high-$V_{1/2}$ model (0.019

https://doi.org/10.1021/acs.jctc.2c01049
J. Chem. Theory Comput. 2023, 19, 1421−1431

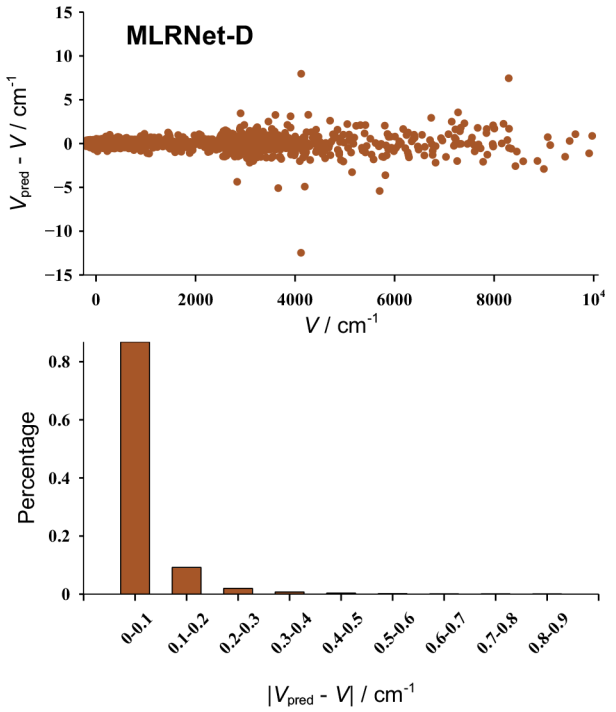


Figure 8. Error distribution of MLRNet-D with uncertainty parameter $V_{1/2}$ = 2000 for the dataset with energies up to 10000 cm^{-1}. The fitting error of MLRNet-D is very close to the PIP-NN result for the same system,[38] which illustrates the fitting ability of MLRNet-D in the short-range region.

cm^{-1} vs 0.031 cm^{-1}). In practical fitting, people should choose the appropriate weight function for their purpose.

4. CONCLUSION

We have discussed the framework of the physics-motivated MLRNet model and validated the fitting accuracy and extrapolation ability of MLRNet(-D) by several tests. The results of the application systems showed that MLRNet can achieve higher accuracy and better high-dimensional fitting efficiency than the conventional mdMLR model, while the extrapolation retains physical sense. For the CO_2−He system, the fitting RMSD of MLRNet was only 0.005 cm^{-1}, which is about an order of magnitude lower than that for the previous mdMLR model "Li2008". The extrapolation of the region $R \geq$ 5.0 Å was only 0.027 cm^{-1} for the model trained with the $R <$ 5.0 Å dataset. For the H_2O−Ar system, MLRNet achieved an RMSD of 0.037 cm^{-1}, and the damping function version MLRNet-D had an even smaller error (RMSD = 0.028 cm^{-1}). For the MLRNet-D model trained with the $V <$ 500.0 cm^{-1} dataset, the relative RMSE of the additional short-range set R = 2.0 Å ($V \in$ (29352, 132138) cm^{-1}) was only 12.26%, which showed physically sensible short-range extrapolation. Compared with the general machine learning model PIP-NN, MLRNet-D displayed a similar fitting performance for points with energy up to 10000 cm^{-1}. However, MLRNet-D had a greater fitting efficiency and only required half of the trainable parameters.

The aim of this work has been concerned not only with the introduction of a high-accuracy intermolecular PES model but also with a discussion of the combination of a popular machine learning model with a physically realistic model potential. As mentioned in the Introduction, this type of physically motivated ML model can inherit both the physics knowledge of the model potential and the fitting ability of the ML technique. Compared to other ML techniques, one of the most important advantages of neural networks is that they can easily be integrated with other potential models. At least in the MLRNet case, an appropriate neural network representation of the MLR function is an accurate, high-dimensionally efficient, and also extrapolation-sensible framework for IPES fitting. In addition, the high-dimensional neural network potential (HDNNP) version of MLRNet is under development for systems containing tens of atoms. There are, of course, some shortcomings of the current MLRNet model that should be mentioned. The biggest problem comes from the constrained physical function format. Although the MLR function provides a sensible extrapolation, it limits the entire model to be suitable only for two-body interactions. For many-body interactions in higher order, the MLRNet model can hardly perform well since the MLR function itself considers only two-body interactions so far. However, machine learning models like PIP-NN do not have this limitation and can handle complex reaction systems easily. More flexible (or other target-specific) physical potential functions need to be developed to ensure that the combination model still available.

■ ASSOCIATED CONTENT

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

■ AUTHOR INFORMATION

Corresponding Author

Hui Li − *Institute of Theoretical Chemistry, College of Chemistry, Jilin University, Changchun 130023, P. R. China;* orcid.org/0000-0002-2689-2241; Email: Prof_huili@jlu.edu.cn

Authors

You Li − *Institute of Theoretical Chemistry, College of Chemistry, Jilin University, Changchun 130023, P. R. China;* orcid.org/0000-0003-4882-8459

Yu Zhai − *Institute of Theoretical Chemistry, College of Chemistry, Jilin University, Changchun 130023, P. R. China;* orcid.org/0000-0002-5065-688X

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.2c01049

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We appreciate F. R. W. McCourt (University of Waterloo) for careful reading of and helpful suggestions on this manuscript. This research was sponsored by the 2020-JCJQ Project (GFJQ2126-007) and the National Natural Science Foundation of China (Grants 22073035, 21773081, and 21533003).

■ REFERENCES

(1) Hornik, K.; Stinchcombe, M.; White, H. Multilayer feedforward networks are universal approximators. *Neural Networks* **1989**, *2*, 359−366.

(2) Barron, A. Universal approximation bounds for superpositions of a sigmoidal function. *IEEE Trans. Inf. Theory* **1993**, *39*, 930−945.

(3) Behler, J. Neural network potential-energy surfaces in chemistry: a tool for large-scale simulations. *Phys. Chem. Chem. Phys.* **2011**, *13*, 17930−17955.
(4) Manzhos, S.; Carrington, T., Jr. Neural network potential energy surfaces for small molecules and reactions. *Chem. Rev.* **2021**, *121*, 10187−10217.
(5) Dral, P. O. Quantum chemistry in the age of machine learning. *J. Phys. Chem. Lett.* **2020**, *11*, 2336−2347.
(6) Jiang, B.; Li, J.; Guo, H. High-fidelity potential energy surfaces for gas-phase and gas−surface scattering processes from machine learning. *J. Phys. Chem. Lett.* **2020**, *11*, 5120−5131.
(7) Meuwly, M. Machine learning for chemical reactions. *Chem. Rev.* **2021**, *121*, 10218−10239.
(8) 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*, 10073−10141.
(9) Buckingham, A. D.; Fowler, P. W.; Hutson, J. M. Theoretical studies of van der Waals molecules and intermolecular forces. *Chem. Rev.* **1988**, *88*, 963−988.
(10) Stone, A. *The Theory of Intermolecular Forces*; Oxford University Press, 2013.
(11) Li, A.; Guo, H. A Full-Dimensional Global Potential Energy Surface of $H_3O^+(\tilde{a}^3A)$ for the $OH^+(\tilde{X}^3\Sigma^-) + H_2\ (\tilde{X}^1\Sigma_g^+) \rightarrow H(^2S) + H_2O^+(\tilde{X}^2B_1)$ Reaction. *J. Phys. Chem. A* **2014**, *118*, 11168−11176.
(12) Yang, D.; Zuo, J.; Huang, J.; Hu, X.; Dawes, R.; Xie, D.; Guo, H. A Global Full-Dimensional Potential Energy Surface for the K_2Rb_2 Complex and Its Lifetime. *J. Phys. Chem. Lett.* **2020**, *11*, 2605−2610.
(13) Stoppelman, J. P.; McDaniel, J. G. Physics-based, neural network force fields for reactive molecular dynamics: Investigation of carbene formation from [EMIM+][OAc]. *J. Chem. Phys.* **2021**, *155*, 104112.
(14) Bereau, T.; DiStasio, R. A.; Tkatchenko, A.; von Lilienfeld, O. A. Non-covalent interactions across organic and biological subsets of chemical space: Physics-based potentials parametrized from machine learning. *J. Chem. Phys.* **2018**, *148*, 241706.
(15) Konrad, M.; Wenzel, W. CONI-Net: Machine learning of separable intermolecular force fields. *J. Chem. Theory Comput.* **2021**, *17*, 4996−5006.
(16) James, A.; John, C.; Melekamburath, A.; Rajeevan, M.; Swathi, R. S. A journey toward the heaven of chemical fidelity of intermolecular force fields. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2022**, *12*, No. e1599.
(17) Behler, J. Four generations of high-dimensional neural network potentials. *Chem. Rev.* **2021**, *121*, 10037−10072.
(18) Gokcan, H.; Isayev, O. Learning molecular potentials with neural networks. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2022**, *12*, No. e1564.
(19) Le Roy, R. J.; Huang, Y.; Jary, C. An accurate analytic potential function for ground-state N_2 from a direct-potential-fit analysis of spectroscopic data. *J. Chem. Phys.* **2006**, *125*, 164310.
(20) Le Roy, R. J.; Henderson, R. D. E. A new potential function form incorporating extended long-range behaviour: application to ground-state Ca_2. *Mol. Phys.* **2007**, *105*, 663−677.
(21) Le Roy, R. J.; Dattani, N. S.; Coxon, J. A.; Ross, A. J.; Crozet, P.; Linton, C. Accurate analytic potentials for $Li_2(X\sum_{1g}^{+})$ and $Li_2(A\sum_{1u}^{+})$ from 2 to 90 Å, and the radiative lifetime of Li(2p). *J. Chem. Phys.* **2009**, *131*, 204309.
(22) Le Roy, R. J.; Haugen, C. C.; Tao, J.; Li, H. Long-range damping functions improve the short-range behaviour of 'MLR' potential energy functions. *Mol. Phys.* **2011**, *109*, 435−446.
(23) Dawes, R.; Quintas-Sánchez, E. The Construction of Ab Initio-Based Potential Energy Surfaces. *Rev. Comput. Chem.* **2018**, *31*, 199−263.
(24) Zhai, Y.; Li, H.; Le Roy, R. J. Constructing high-accuracy intermolecular potential energy surface with multi-dimension Morse/Long-Range model. *Mol. Phys.* **2018**, *116*, 843−853.
(25) Wang, L.; Xie, D.-Q.; Le Roy, R. J.; Roy, P.-N. A new four-dimensional ab initio potential energy surface for NO-He and vibrational band origin shifts for the N_2O-He_N clusters with N = 1−40. *J. Chem. Phys.* **2012**, *137*, 104311.
(26) Zhang, X.-L.; Li, H. Three-dimensional ab initio potential energy surface and predicted spectra for the CH_4-Ne complex. *Chin. J. Chem. Phys.* **2021**, *34*, 874−882.
(27) Li, H.; Le Roy, R. J. Analytic three-dimensional 'MLR' potential energy surface for CO_2-He, and its predicted microwave and infrared spectra. *Phys. Chem. Chem. Phys.* **2008**, *10*, 4128−4137.
(28) Hou, D.; Yang, J.-T.; Zhai, Y.; Zhang, X.-L.; Liu, J.-M.; Li, H. Analytic intermolecular potential energy surface and first-principles prediction of the rotational profiles for a symmetric top ion-atom complex: A case study of H_3O^+-Ar. *J. Chem. Phys.* **2020**, *152*, 214302.
(29) Wang, L.; Zhang, X.-L.; Zhai, Y.; Nooijen, M.; Li, H. Explicitly correlated ab initio potential energy surface and predicted rovibrational spectra for H_2O−N_2 and D_2O−N_2 complexes. *J. Chem. Phys.* **2020**, *153*, 054303.
(30) Zhang, X.-L.; Ma, Y.-T.; Zhai, Y.; Li, H. Analytic Morse/long-range potential energy surfaces and "adiabatic-hindered-rotor" treatment for a symmetric top-linear molecule dimer: A case study of CH_3F-H_2. *J. Chem. Phys.* **2018**, *148*, 124302.
(31) Liu, J.-M.; Zhai, Y.; Zhang, X.-L.; Li, H. Intermolecular configurations dominated by quadrupole−quadrupole electrostatic interactions: explicit correlation treatment of the five-dimensional potential energy surface and infrared spectra for the CO-N_2 complex. *Phys. Chem. Chem. Phys.* **2018**, *20*, 2036−2047.
(32) Liu, J.-M.; Zhai, Y.; Li, H. Explicit correlation treatment of the six-dimensional potential energy surface and predicted infrared spectra for OCS-H_2. *J. Chem. Phys.* **2017**, *147*, 044313.
(33) Li, H.; Blinov, N.; Roy, P.-N.; Le Roy, R. J. Path-integral Monte Carlo simulation of $\nu 3$ vibrational shifts for CO_2 in $(He)_n$ clusters critically tests the He−CO_2 potential energy surface. *J. Chem. Phys.* **2009**, *130*, 144305.
(34) Li, H.; Le Roy, R. J.; Roy, P.-N.; McKellar, A. R. W. Molecular superfluid: nonclassical rotations in doped para-hydrogen clusters. *Phys. Rev. Lett.* **2010**, *105*, 133401.
(35) Raston, P. L.; Jäger, W.; Li, H.; Le Roy, R. J.; Roy, P.-N. Persistent Molecular Superfluid Response in Doped Para-Hydrogen Clusters. *Phys. Rev. Lett.* **2012**, *108*, 253402.
(36) Li, H.; Ma, Y.-T. An intramolecular vibrationally excited intermolecular potential for He-OCS: Globally tested by simulation of vibrational shifts for OCS in He_n n = 1 100 clusters. *J. Chem. Phys.* **2012**, *137*, 234310.
(37) Yang, D.-Z.; Liu, L.; Xie, D.-Q.; Guo, H. Full-dimensional quantum studies of vibrational energy transfer dynamics between H_2O and Ar: theory assessing experiment. *Phys. Chem. Chem. Phys.* **2022**, *24*, 13542−13549.
(38) Zhou, Y.-Z.; Xie, D.-Q.; Liu, Q.; Wang, J.-Y. A full-dimensional ab initio intermolecular potential energy surface and dipole moment surfaces for H_2O-Ar. *Curr. Chin. Sci.* **2022**, *2*, 325−334.
(39) Braams, B. J.; Bowman, J. M. Permutationally invariant potential energy surfaces in high dimensionality. *Int. Rev. Phys. Chem.* **2009**, *28*, 577−606.
(40) Bowman, J. M.; Czakó, G.; Fu, B. High-dimensional ab initio potential energy surfaces for reaction dynamics calculations. *Phys. Chem. Chem. Phys.* **2011**, *13*, 8094.
(41) Jiang, B.; Guo, H. Permutation invariant polynomial neural network approach to fitting potential energy surfaces. *J. Chem. Phys.* **2013**, *139*, 054112.
(42) Shao, K.; Chen, J.; Zhao, Z.; Zhang, D. H. Communication: Fitting potential energy surfaces with fundamental invariant neural network. *J. Chem. Phys.* **2016**, *145*, 071101.
(43) Fu, B.; Zhang, D. H. Ab Initio Potential Energy Surfaces and Quantum Dynamics for Polyatomic Bimolecular Reactions. *J. Chem. Theory Comput.* **2018**, *14*, 2289−2303.
(44) Paukku, Y.; Yang, K. R.; Varga, Z.; Truhlar, D. G. Global ab initio ground-state potential energy surface of N_4. *J. Chem. Phys.* **2013**, *139*, 044309.

https://doi.org/10.1021/acs.jctc.2c01049
J. Chem. Theory Comput. 2023, 19, 1421−1431

(45) Le Roy, R. J.; Pashov, A. betaFIT: A computer program to fit pointwise potentials to selected analytic functions. *J. Quant. Spectrosc. Radiat. Transfer* **2017**, *186*, 210−220.
(46) Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; Desmaison, A.; Kopf, A.; Yang, E.; DeVito, Z.; Raison, M.; Tejani, A.; Chilamkurthy, S.; Steiner, B.; Fang, L.; Bai, J.; Chintala, S. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In *Advances in Neural Information Processing Systems 32*; Curran Associates, 2019; pp 8024−8035.
(47) Falcon, W.; Borovec, J.; Wälchli, A.; Eggert, N.; Schock, J.; Jordan, J.; Skafte, N.; Ir1dXD; Bereznyuk, V.; Harris, E.; Murrell, T.; Yu, P.; Præsius, S.; Addair, T.; Zhong, J.; Lipin, D.; Uchida, S.; Bapat, S.; Schröter, H.; Dayma, B.; Karnachev, A.; Kulkarni, A.; Komatsu, S.; Martin.B; Schiratti, J.-B.; Mary, H.; Byrne, D.; Eyzaguirre, C.; cinjon; Bakhtin, A. *PyTorchLightning/pytorch-lightning: 0.7.6 release*, 2020. DOI: 10.5281/zenodo.3828935 (accessed 2022-12-24).
(48) Metz, M. P.; Szalewicz, K. A statistically guided grid generation method and its application to intermolecular potential energy surfaces. *J. Chem. Phys.* **2020**, *152*, 134111.
(49) Hou, D.; Ma, Y.-T.; Zhang, X.-L.; Li, H. The origins of intra- and inter-molecular vibrational couplings: A case study of H_2O-Ar on full and reduced-dimensional potential energy surface. *J. Chem. Phys.* **2016**, *144*, 014301.
(50) Tang, K. T.; Toennies, J. P. An improved simple model for the van der Waals potential based on universal damping functions for the dispersion coefficients. *J. Chem. Phys.* **1984**, *80*, 3726−3741.
(51) Kreek, H.; Meath, W. J. Charge-overlap effects. dispersion and induction forces. *J. Chem. Phys.* **1969**, *50*, 2289−2302.