Many-body Neural Network-based Force Field for Structure-based Coarse-Graining of Water

A. Moradzadeh[1] and N. R. Aluru[2*]

aluru@utexas.edu

[1]Department of Mechanical Science and Engineering, University of Illinois at Urbana−Champaign, Urbana, IL, 61801 United States, [2]Oden Institute for Computational Engineering and Sciences, Walker Department of Mechanical Engineering, The University of Texas at Austin, Austin, TX, 78712 United States

Abstract

High-fidelity results from atomistic simulations can only be obtained by using accurate force field (FF) parameters. Although empirical FFs are commonly used in the modeling of atomistic systems due to their simplicity, they have many limitations inherent in the crude approximations associated with their analytical form. Recent advances in neural network-based FFs have led to more accurate FFs by using symmetrical functions or full many-body expansions. However, this approach leads to several issues including the arbitrariness of the symmetry functions, and the intangible and uninterpretable interactions which are only known once the positions of all atoms are set. More importantly, training is another bottleneck, as high-quality force and energy information is required, which is usually not accessible from experimental data. To solve these issues within the context of structure-based coarse-graining methods, we switch in this work to a local search method to target the reference structure instead of using conventional backpropagation algorithms used to target the forces and energies of the reference structure. Our FF is decomposed into two-, three-, and higher-order terms, where each term is modeled with a separate neural network. To show the versatility of our method, we study four different systems, namely, Stillinger-Weber particles as an atomistic case and three water models, namely SPC/E, MB-pol, and *ab initio*, as coarse-graining cases. We show the successful application of our approach, by reproducing structural properties of different water models, followed by providing insight into the role of two-and three-body interactions. The results of all models indicate that the double-well isotropic pair potential, signature of waterlike behavior in an isotropic system, vanishes upon inclusion of three-body interaction, showing dominance of three-body interaction over two-body interaction in water-like behavior with single-well isotropic pair potential.

I. Introduction

Classical molecular dynamics (MD) simulation is a powerful computational tool to study various physical, chemical, and biological systems[1] by allowing researchers to obtain structural and dynamical properties of these systems with a computational cost far less than the *ab initio* molecular dynamics simulations and resolution far better than the continuum models. The accuracy of the force fields (FFs) used in MD simulations is pivotal to obtain accurate and meaningful results. To date, the most common choice for FFs in MD simulations belongs to some predefined analytical forms, known as empirical FFs, which are sometimes limited in truthful representation of real interactions.[2–4] The problem is particularly more pronounced in coarse-grained FF developments, where due to the many lost degrees of freedom, coarse-grained (CG) FF often exhibit features that are not necessarily present in the reference system.

Various strategies have been developed to improve the accuracy of empirical FFs, such as the inclusion of higher-body interactions, *e.g.,* three-body interactions. For example, a single bead water model with the Stillinger-Weber potential[5], known as the mW model[6], predicts various properties, such as the melting temperature, as accurately as more complex all-atom models. However, such FFs are usually designed with various approximations (*e.g.,* analytical form limitation) motivated by physical intuition, making their application to other systems difficult, if not impossible. In general, the selection of an analytical form and its parameterization requires considerable expertise as well as computer and human time.[7] Motivated by recent progress in the application of machine learning in atomistic simulations[8–14], researchers in FF development have embarked on employing various machine learning methods, such as kernel- and neural network-based force fields (NNFFs)[15–19]. Particularly, NNFFs solve the problem of functional form limitation using the universal approximation theorem.

Despite recent success in NNFFs, there are still several issues challenging the widespread usage of NNFFs, such as the arbitrariness of symmetry functions and state dependency of the NNFFs, as well as interoperability. Due to the implementation of NNFFs, they are also less interpretable compared to the empirical FF, as forces and energies are only known once the positions of the whole system are set. Most of the available NNFFs also require accurate force and energy data, which are not always available or measurable. Recent studies have attempted to solve this problem using more interpretable FFs and active learning schemes, but there is still a need for accurate forces and energies.[4]

Recently, coarse-grained force field development has been tackled using machine learning-based force fields such as NNFFs and kernel-based methods[20–23]. However, in many cases, the FF is not interpretable. Motivated by the need to overcome challenges with the interpretability of NNFFs and targeting structural properties in absence of accurate energies and forces, we propose a new strategy for developing NNFFs. In this study, instead of building a single NNFF, we build our NNFFs by the expansion of interactions as separate two-, three-, and n-body interactions. We limit our NNFF to a three-body expansion with the possibility of considering higher expansions in a similar manner. Our method, unlike other NNFFs[17,18,23,24], does not necessarily require forces or energies on the particles to train the FF and can be trained solely based on the structural properties of the reference system. This is achieved by replacing backpropagation for NNFF training with a local-search algorithm[25]. We start with atomistic-scale force field development using the forces and positions of the mW model, *i.e.,* force-matching of mW model without any coarse-graining. After demonstrating the success of the method within the force-matching framework, we develop a single bead coarse-grained water model targeting the structural properties of three water models, namely SPC/E, MB-pol,[26] and *ab initio* models. The choice of water model is motivated by its essential role in understanding anomalous behavior of water.[27–31] Furthermore, our study assesses the role of two- and three-body interactions in water-like model, providing significant evidence for possibly universal single-well isotropic interaction for water-like behavior in the inclusion of three-body interactions.[32] To show this point, we coarse-grain different reference water models, and compared our results with mW and extended dipole water models to provide further evidence for our findings. In the case of mW model, our finding justifies why mW model has been successful with only having a single well pairwise interaction as our NNFF also indicates a single well pairwise interaction. In the case of extended dipole water model, comparison with the model developed by Motevaselian et al.[33] shows that the presence of dipole-dipole interaction does not lead to a single-well pairwise potential, therefore, indicating significant role of three-body interactions for water-like behavior.

We show that our model can reproduce structural properties such as the radial distribution function and angular distribution of water. The interpretability of the NNFF in our study allows us to provide insights into the role of two- and three-body interactions in water-like behavior, especially as most of the past studies have focused on the two-body interaction[15,30,33,34]. Our study shows that coarse-graining of different water models with inclusion of three-body interactions leads to the suppression of double-well isotropic pair potential independent of the water model. This finding indicates superior role of anisotropic interaction in reproducing anomalous water-like behavior as well as success of Stillinger-Weber potential in modeling CG water. Compared with other machine learning-based FFs targeting similar problems, *e.g.,* kernel-based methods, our method can handle complex radial and angular dependencies in the interactions between particles. For example, a spline three-body interaction or kernel-based three-body interaction[15] capturing the complexity of our model (with fewer than 1000 free parameters) requires at least 10^6 free parameters (100 for each of the two radial components and one angular component), which is not possible to train and implement due to the computational and memory costs.

II. Methods

1. Simulations

The MD simulation of the mW model[6] is based on the Stillinger-Weber potential.[5] It is described by two- and three-body interactions over pairs and triplets, respectively, as follows:

$$u_{sw}^2(r) = A\epsilon\left[B\left(\frac{\sigma}{r}\right)^p - \left(\frac{\sigma}{r}\right)^q\right]\exp\left(\frac{\sigma}{r-a\sigma}\right) \tag{1}$$

$$u_{sw}^3(r_1, r_2, \theta) = \lambda\epsilon(\cos\theta - \cos\theta_0)^2 \exp\left(\frac{\gamma\sigma}{r_{ij}-a\sigma} + \frac{\gamma\sigma}{r_{ik}-a\sigma}\right) \tag{2}$$

where r indicates the radial distance between a pair, r_{ij}, r_{ik}, and θ are the radial distances and the angle between the central atom and two other atoms, respectively (see Figure 1c). $\sigma(= 2.3925\ \text{A°})$ and $\epsilon(= 6.189\ \text{kcal/mol})$ are the length and energy scales of the two-body potential, respectively. $\lambda\ (= 23.15)$ is the scaling parameter indicating the strength of tetrahedral interactions. $A = 7.049556277, B = 0.6022245584, p = 4$, and $q = 0$ are parameters giving rise to the form and scale of the potential, and the reduced cutoff $a = 1.8$ enforces that both forces and potentials vanish at distances larger than $a\sigma$. The angular term enforces a tetrahedral angle around $\theta_0 = 109.47°$. MD simulations of CG and SPC/E models are performed using LAMMPS[35] with a timestep of 2 fs in the NVT ensemble. The temperature is controlled using the Nosé-Hoover thermostat with a time constant of 0.2 ps. MD simulations of MB-pol are performed using OpenMM[36] with MB-pol plugin. The temperature is controlled by the Anderson thermostat[37] with the collision frequency of 1.0 ps and equations of motion are solved using the velocity Verlet algorithm with a time step of 0.2 fs. The radial and angular distribution functions of *ab initio* water are obtained using *ab initio* molecular dynamics simulations with optimized norm-conserving Gaussian pseudopotentials generated by PBE were used with a double-zeta polarized basis[38,39]. Similar to Refs[40,41], the AIMD simulations were performed using an isothermal-isochoric (NVT) ensemble where the constant temperature was controlled using the Nosé-Hoover thermostat[42]. The initial configuration of AIMD is obtained from MD simulation production run, *i.e.,* the last frame of SPC/E MD simulation. We consider cubic simulation boxes with an edge length of 20.02 Å. Following a 2 ps equilibration, a production run of 22 ps trajectories is carried out using a 0.5 fs timestep. The radial and angular distribution functions are averaged over the total simulation time with a frequency of 50 fs.

2. Deep Learning

The NNFF developed in this study is composed of two- and three-body interactions, each described by a neural network[43,44]. Note that while we keep the FF limited to two- and three-body interactions, there is no limitation in using higher-order many-body interactions. The two-body term takes the radial distance of a pair as input, and the three-body term takes the radial distances and angle between a central atom and two other atoms in a triplet. Mathematically, the two- and three-body forces of our NNFF can be written as follows:

$$f_{nn}^{2}(\boldsymbol{r}) = \phi_2\big((r_{cf,2} - r)/r_{cf,2}; \mathbf{W}_2, \boldsymbol{b}_2\big).\, \boldsymbol{e}_r \tag{3}$$

$$f_{nn}^{3}(\boldsymbol{r}_1, \boldsymbol{r}_2, \theta) = \mathcal{E}\big(\boldsymbol{e}_{\boldsymbol{r}_1}, \boldsymbol{e}_{\boldsymbol{r}_2}\big)\, \boldsymbol{\phi}_3\big((r_{cf,3} - r_1)/r_{cf,3}, (r_{cf,3} - r_2)/r_{cf,3}, \cos\theta\,; \boldsymbol{W}_3,\, \boldsymbol{b}_3\big) \tag{4}$$

where $r_{cf,2}$ and $r_{cf,3}$ are the cutoff distances of two-body and three-body interactions, respectively. $\boldsymbol{e_r}$ is the unit vector along the radial vector between two atoms ($\boldsymbol{e}_r = \frac{\boldsymbol{r}}{|\boldsymbol{r}|}$). ϕ_2 is the output of the two-body neural network with $\phi_2 \in \mathbb{R}^1$ (see Figure 1b), and $\boldsymbol{\phi}_3$ is the output of the three-body neural network with $\boldsymbol{\phi}_3 \in \mathbb{R}^4$ (see Figure 1c). $\boldsymbol{W}$ and $\boldsymbol{b}$ are the weights and biases of the neural networks, respectively, and their indices indicate the number of expansions, namely, two- and three-body interactions. We use different neural network depths and widths to model different interactions; however, we use similar nonlinearities, swish nonlinearity ($\text{swish}(x) = \frac{x}{1+\exp(-x)}$) between the hidden layers and tanh nonlinearity for the output layer. $\mathcal{E}(\boldsymbol{e}_1, \boldsymbol{e}_2)$ is a mapping from $\mathbb{R}^4$ to Cartesian coordinates, written as follows:

$$\mathcal{E}(\boldsymbol{e}_1, \boldsymbol{e}_2) = \begin{bmatrix} e_{1,x} & e_{2,x} & e_{1,x} & e_{2,x} \\ e_{1,y} & e_{2,y} & e_{1,y} & e_{2,y} \\ e_{1,z} & e_{2,z} & e_{1,z} & e_{2,z} \end{bmatrix} \tag{5}$$

where the first two columns of $\mathcal{E}(\boldsymbol{e}_1, \boldsymbol{e}_2)$ and the first two elements of $\boldsymbol{\phi}_3$ determine the contribution of force on the central atom from the left particle of a triplet (left, central, right). Similarly, the last two columns of $\mathcal{E}(\boldsymbol{e}_1, \boldsymbol{e}_2)$ and the last two elements of $\boldsymbol{\phi}_3$ show contributions from the right atom in a triplet (left, central, right). Note that for the applications of n-body interaction within our framework, the following steps can be taken: Given the position of the central atom, the n-body interaction depends on the position of $n-1$ particles, described by $n(n-1)/2$ relative distances or angles, and it produces a force vector with a dimension of $\mathbb{R}^{(n-1)^2}$, which gets mapped to the Cartesian coordinates through the mapping matrix, $\mathcal{E}(\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_{n-1})$. The reaction forces are mapped to the contributing noncentral atom using specific columns of $\mathcal{E}(\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_{n-1})$.

Having the appropriate force calculation method for a particular pair or triplet, we use two different schemes to train our NNFF, namely, force-matching and structure-matching methods. The force-matching method is a simple method without requiring any CGMD simulations. However, it fails to reproduce the structure of the coarse-grained model, as the effective forces are not equal to instantaneous forces on the CG beads. Additionally, how we map forces on the CG beads changes the FF obtained from the force-matching method[45]. During force matching, one attempts to minimize the mean-squared error between the reference and model forces, which implies minimization of the following loss function using the backpropagation algorithm:

$$\epsilon_{\mathcal{L}}^{f}(\boldsymbol{\theta}, D) = \frac{1}{2|D|}\sum_{i\in D}\left[\boldsymbol{f}_{GT}(i) - \sum_{\forall j \in P_i} \boldsymbol{f}_{nn}^{2}(j) - \sum_{\forall j \in T_j} \boldsymbol{f}_{nn}^{3}(j)\right]^2 \tag{6}$$

where $\boldsymbol{\theta}$ indicates the free parameter of NNFF, both weights and biases of two- and three-body interactions. D is the training data set. $\boldsymbol{f}_{GT}(i)$ is the ground-truth force on particle i. P_i and T_i represent all the unique pairs and triplets of particle i, respectively. Summation over pairs and triplets proceeds over those containing atom i without duplicate terms. Note that atom i can be the left, right, or central atom in a triplet, which leads to using the first two, last two, or all the columns of $\mathcal{E}(\boldsymbol{e}_1, \boldsymbol{e}_2)$, respectively. Note that the above procedure for triplet force calculation enforces Newton's third law of action and reaction. During the training of the network, input data, *i.e.,* various pairs and triplets, along with the corresponding ground truth forces are fed into the neural networks, which are then used to train through backpropagation. Based on our studies, it is easier to use convolutional neural networks to calculate forces for the above force-matching method and model each of the two- and three-body interactions as separate filters over pairs and triplets, respectively. To eliminate the effects due to the particles outside of the cutoff distance, we multiply the output of the convolutional neural network with a rectified linear unit with linearly transformed radial distance as input ($\max(0, (r_{cf} - r)/r_{cf})$). The Adam optimizer[46] is used to optimize the weights and biases, during force-matching with L2 regularization to avoid overfitting.

Furthermore, we develop a new route to train our NNFF without requiring direct access to forces. We apply our method to reproduce the structural properties of reference systems (note that the method is general and can be applied to various properties). Our method solves the problem with force matching in the reproduction of structural properties and the fact that reference forces are not always the target of force field development or available during FF development. To do so, the backpropagation algorithm, which is used in the force-matching method, is replaced with the local search algorithm. This is similar to using evolutionary or reinforcement learning methods to train NNFF. In other words, MD acts as an environment where the policy predicts forces and forces are used for MD simulation[47,48]. The cost determined at the end of the MD simulation is the structural agreement between the CG and reference models[49,50]. Within the local-search algorithm, the training for such policy is done through two steps. First, a small fraction of the free parameters of NNFF is selected randomly, followed by a perturbation of selected parameters. The perturbation is accepted if it improves the loss function. Within our structure-based coarse-graining method, the loss function is defined as:

$$\epsilon_{\mathcal{L}}^{s}(\theta) = \frac{\sum_{n=1}^{N} |g_{ref}(r_n) - g_{cg}(r_n)|}{\sum_{n=1}^{N} |g_{ref}(r_n)|} + \alpha \sum_{l=1}^{L} \frac{\sum_{m=1}^{M} |a_{ref}^{(l)}(\theta_m) - a_{cg}^{(l)}(\theta_m)|}{L} \tag{7}$$

where g_{ref} and g_{cg} represent RDFs of the reference and CG models. a_{ref} and a_{cg} show ADFs of the reference and CG models. α indicates the importance of ADF in the structural similarity of the reference and CG models (with only 2-body interaction α is set to 0). The RDF is discretized into 360 segments from 0 to the cutoff distance of two-body interactions. Similarly, the ADF is discretized into 240 segments from 0 to 180 degrees. Furthermore, ADF is calculated for 7 different cutoff distances ranging from 0.3 *nm* to 0.45 *nm*. Additionally, note that for each iteration of local search training, a short MD simulation is performed to obtain the RDF and ADF of CG models and therefore the loss function (see reference[25] for details on the local search algorithm).

III. Results and Discussion

1. Force Matching

The objective of the force-matching method is to reproduce the ground truth forces. In this study, we use force matching to show that our scheme of training NNFF leads to retrieving the ground-truth force field of the mW model in the form of NNFF. To do so, we train our NNFF based on forces obtained at multiple temperatures. We perform an NPT simulation at 300 K and 1 bar to obtain the correct density, followed by a simulation in the NVT ensemble at T = 270, 300, 330 K to obtain the positions and forces of the mW model particles.

Once the data were obtained, we trained our NNFF with a cutoff distance of 0.4310 *nm*, equal to the mW model cutoff distance. The two- and three-body interactions in NNFF form are modeled with 3 and 4 hidden layers, respectively. All the hidden layers except the output layer have swish nonlinearity, and the output layer has tanh nonlinearity. We trained the network for approximately $2\text{x}10^6$ iterations with a batch size of 64. In Figure 2a, we compare the RDF between the mW and NNFF models, followed by a comparison of the two-body forces of the ground truth mW and NNFF models at different radial distances, as shown in Figure 2b. Similarly, we compare the ADF of the mW and NNFF models with various cutoff distances in Figure 3a-b. We also compare the norm of the three-body forces from the mW and NNFF models at different angles for the equidistant configurations of neighboring atoms in Figure 3c. The method shows good agreement in terms of RDF and ADF reproducibility, as well as recovering the ground truth forces.

2. Structure-based Coarse-Graining

The force-matching scheme is not able to reproduce the structural properties of the reference system. To solve this problem, we employ the local-search method over NNFFs pretrained by force-matching or iterative Boltzmann inversion methods. The local-search algorithm starts by randomly grouping small fractions of NNFF parameters together, followed by adding a random perturbation, usually a uniform perturbation, to the selected parameters. The new candidate NNFF is then used for running CGMD simulation. If the results are improved, then the new candidate becomes the best candidate, and additional perturbations are applied to it; otherwise, another fraction of NNFF parameters is selected. After all groups of NNFF parameters are explored a new grouping of NNFF parameters occurs. The procedure continues until the desired accuracy is obtained or iterations budget is exhausted.[25] The size of fractions and perturbation size are hyper-parameters of local search method.

In our study, we model water both using classical MD, (SPC/E and MB-pol models) and AIMD. For each reference system, we develop two different NNFFs, one with only two-body interactions (NN_2^t, where *t* represents the reference model, *i.e.* $t \in \{SPCE, MBPol, AIMD\}$) and one with both two- and three-body interactions (NN_3^t). In addition to the interpretability of both the $NN_2^t$and NN_3^t models, the models allow us to understand the role of two- and three-body interactions in water-like behavior. Both NN_2^t and NN_3^t model have two-body interactions with a cutoff distance of 0.8 nm, and NN_3^t model has the three-body interaction with a cutoff of 0.45 nm. The NN_2^t model with only two-body interaction has 3 hidden layers with widths of 4, 16, and 4. In the NN_3^t model, the two-body interaction has 3 hidden layers with widths of 6, 12, and 6, and the three-body interaction has 4 hidden layers with widths of 3, 6, 12, and 8. All the hidden layers except the output layer have swish nonlinearity, and the output layer has tanh nonlinearity. Note that both RDF and ADF are calculated between the oxygen atoms, *i.e.,* the oxygen atom of water molecules is the mapped CG model. Our NNFF is implemented as a user package inside the LAMMPS package.

We start by discussing results of SPC/E model, followed by discussions of MB-pol and AIMD water models and finally discussing the results of three models together. Both $NN_2^{SPCE}$and NN_3^{SPCE} models reproduce the RDF of oxygen-oxygen with high accuracy with an error of less than 0.01 (see Figure 4a). In Figure 4b, we compare the two-body interaction between the NN_2^{SPCE} and NN_3^{SPCE} models. As seen in Figure 4b, the NN_3^{SPCE} model has only a single well, while NN_2^{SPCE} model has a double-well. The double-well interaction observed in the NN_2^{SPCE} model is consistent with previous studies that used pairwise isotropic interaction to represent CG model of SPC/E water model[51,52]. Investigation into three-body correlation shows that only NN_3^{SPCE} captures the ADF of SPC/E water model as shown in Figure 5a-c. The difference is especially pronounced at short distance (Figure 5a), where the hydrogen-bonding and tetrahedral structure of water play a significant role. This observation agrees with the directionality of water interactions at short range, which is missing when the CG model lacks three-body interactions. In Figure 5c, we show the norm of the three-body forces of the NN_3^{SPCE} model at different angles for the equidistant configurations of neighboring atoms. Figure 5c indicates that the three-body forces of NN_3^{SPCE} model have far more complex behavior than those of the mW model, which is limited by its analytical form. Therefore, we observe importance of three-body and directional interactions for accurate modeling of CG water model consistent with the previous studies. [53,54]

We also investigate the transferability and representability of the NN_2^{SPCE} and NN_3^{SPCE} models. Coarse-graining often comes at the cost of not reproducing all the properties of the system, however, consistent qualitative behavior in terms of dynamics is observed between structure-based coarse-graining and reference models.[55] To assess representability of model, we calculate the isothermal compressibility, thermal expansion coefficient, density, and diffusion coefficient. The isothermal compressibility (k_T) and thermal expansion coefficient (α_p) are evaluated based on the following equations,

$$k_T = -\frac{1}{V(p,T)}\frac{V(p+\epsilon,T)-V(p-\epsilon,T)}{2\epsilon} \tag{8}$$

$$\alpha_p = \frac{1}{V(p,T)}\frac{V(p,T+\epsilon)-V(p,T-\epsilon)}{2\epsilon} \tag{9}$$

where p and V are the pressure and volume of the system, respectively. ϵ is the perturbation around the reference system. For CG models, the reference pressure is chosen based on the NVT simulation results. The results of both isothermal compressibility and expansion coefficients are shown in Table I, where we observe a better reproduction of both properties by the NN_3^{SPCE} model compared to the NN_2^{SPCE} at T 300 K, again indicating the role of three-body interactions in terms of reproducibility.

The transferability is assessed based on the change in structure, density, and diffusion coefficients with temperature change. The structural change is quantified in terms of change in the RDF and ADF of CG models at 270 K and 330 K compared with the SPC/E model RDFs and ADFs (shown in Figure 6). Both models RDF and ADF errors remain in a similar order of magnitude as their error in the target 300 K temperature. However, structural errors of both models show more sensitivity to increase in temperature, mainly due to the dominant role of entropy at higher temperature. The density is obtained from the NPT simulation with reference pressure obtained from the NVT simulation at 300 K. The results (shown in Table II) indicate that both methods reproduce density with less than 0.5 % error, but both methods fail to capture the anomalous behavior of density, again mostly due to the increasing role of entropy. To assess dynamical properties, the diffusion coefficients are compared between the CG and reference models. The diffusion coefficients are computed based on the Einstein relationship from the mean squared displacement ($MSD(\tau) = \langle|\boldsymbol{r}(\tau) - \boldsymbol{r}(0)|^2\rangle = 6D\tau$, as lag time (τ) goes to infinity), see Table II for the values of diffusion coefficient. In general, we observe higher diffusion coefficients in the NN_2^{SPCE} models compared with the NN_3^{SPCE}model, as NN_3^{SPCE} model represents the free energy landscape of the reference system better, but it still overestimates the diffusion coefficient by approximately an order of magnitude.

RDF obtained using MB-pol water model reproduces experimental RDF of water better than SPC/E model. Motivated by the drastic dependency of CG models on the structural properties of the reference systems[56] and understanding general trend in the two- and three-body interactions of CG water, we develop $NN_2^{MBPol}$and NN_3^{MBPol} models. Both models reproduce RDF of MB-pol water model with high accuracy and RDF error of around 0.03 as shown in Figure 7a. In Figure 7b, we compare the two-body interaction between the NN_2^{MBPol} and NN_3^{MBPol} models. As seen in Figure 7b, the NN_3^{MBPol} model has only a single well, while NN_2^{MBPol} model has a double-well, consistent with the observations in SPC/E water model. The ADF results of NN_2^{MBPol} and NN_3^{MBPol} models show that only NN_3^{MBPol} succeeds in reproducing ADF of reference MB-pol water model. Again, the difference is pronounced at short distance (Figure 8a). In Figure 8c, we show the norm of the three-body forces of the NN_3^{MBPol} model at different angles for the equidistant configurations of neighboring atoms. Figures 4a and 7a and Figures 5c and 8c indicate that both two- and three-body forces present in SPC/E and MB-pol waters are different from each other, but they follow a similar qualitative behavior in terms of having a single well two-body interaction in the presence of three-body interactions.

Finally, we investigate behavior of NN_2^{AIMD} and NN_3^{AIMD} models. In Figure 9a, we compare RDF of both models with the reference model and observe a good agreement between them with a small RDF error (~0.03). Comparison between the two-body interactions of NN_2^{AIMD} and NN_3^{AIMD} models shows that the NN_3^{AIMD} model again has a single well form. In Figure 10a-b, we compare ADF of NN_2^{AIMD} and NN_3^{AIMD} models and observe better performance of NN_3^{AIMD} model in reproducing ADF of reference AIMD water model, especially at short distances. In Figure 10c, we show the norm of the three-body forces of the NN_3^{AIMD} model at different angles for the equidistant configurations of neighboring atoms. We observe both two- and three-body forces are different compared to both SPC/E and MB-pol water models, especially as the reference AIMD water model is more structured compared to both classical models.

We briefly recap the key characteristics of NN_2^t and NN_3^t models. We observe that both the NN_2^t and NN_3^t models can capture the RDF of the reference water; however, only the NN_3^t model captures the ADF behavior of water. We observe a double-well interaction in NN_2^t CG models; however, NN_3^t only has a single well independent of the reference model. The fact that in the presence of a three-body interaction, the double-well structure of the two-body interaction vanishes and ADF is better reproduced shows that water structural properties are best described by a higher-order expansion of interactions rather than a simple pairwise interaction with a double-well. This is also consistent with the directional dependency of hydrogen bonds[57] in water, which is best described by a three-body interaction rather than a double-well two-body interaction.

IV. Conclusions

In this study, we train a neural network-based force field with two- and three-body interactions, which makes the force field more interpretable. Within our framework, the requirement for accurate forces and energies is eliminated by using the local search algorithm instead of backpropagation. Therefore, our method is suitable for application in structure-based coarse-graining. To show the capability of our method, we successfully develop coarse-grained models of different water models. We also investigate the dependency of the coarse-grained force field of water on the number of expansions, which shows that the double-well interaction, known as a signature of water-like behavior among spherically symmetric pairwise interactions, vanishes with the inclusion of three-body interactions independent of the reference water model. This indicates a consistent and possibly universal dominance of three-body interaction in water-like behavior among different water models. We also notice that the two-body interaction fails to reproduce the angular distribution of water, especially over a short range. Based on our findings, we conclude that water-like behavior is better captured using the three-body interaction, which is consistent with the directional dependency of interactions in water.

Acknowledgments

A.M. would like to thank Anshul Saxena for *ab initio* data on radial and angular distribution of water. The work on *ab initio* water models was supported by the Center for Enhanced Nanofluidic Transport (CENT), an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences (Award # DE-SC0019112). All other aspects of this work were supported by the National Science Foundation under Grant 2140225. The authors acknowledge the use of Blue Waters supercomputing resources at the University of Illinois at Urbana-Champaign. Furthermore, this work partially used the Extreme Science and Engineering Discovery Environment (XSEDE) Stampede2 at the Texas Advanced Computing Center through allocation TG-CDA100010. This work also utilizes resources supported by the National Science Foundation’s Major Research Instrumentation program, grant #1725729, as well as the University of Illinois at Urbana-Champaign.

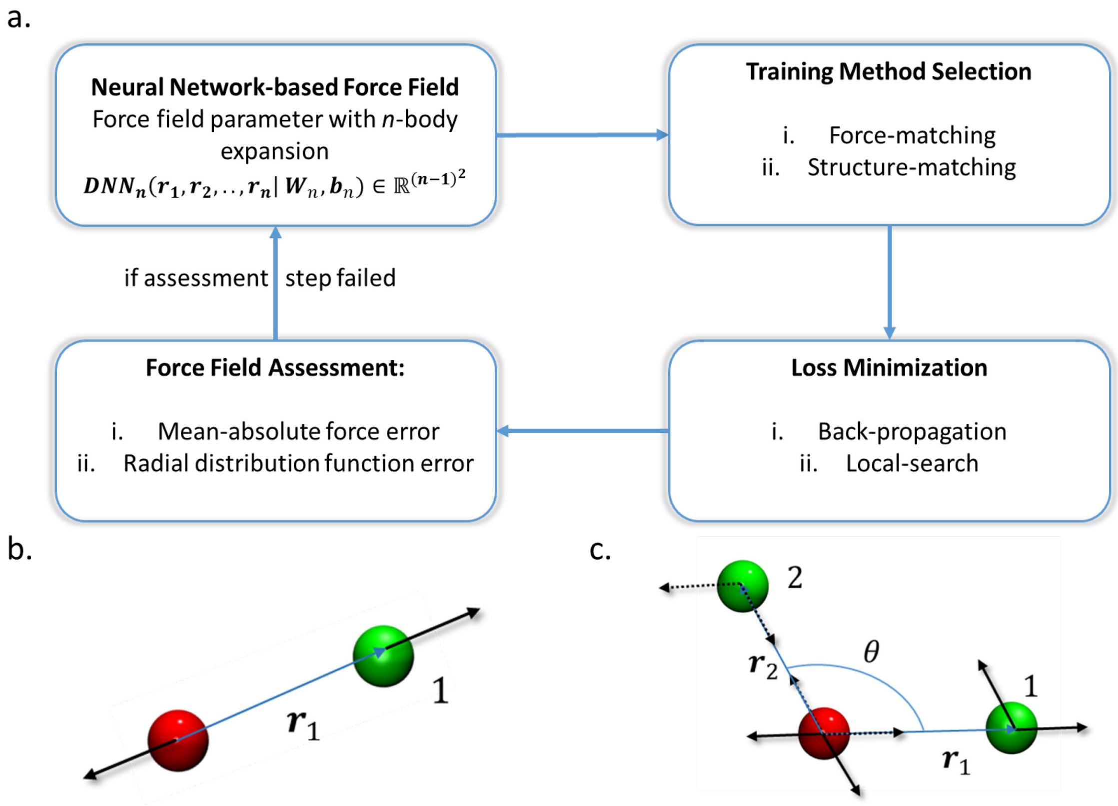

Figure 1. Neural network-based force-field training. a) Force-field development begins by selection of n-body expansion and initialization of neural network architecture, weights, and biases. Based on the objective of force field development and its differentiability, either back-propagation or local-search methods are employed to minimize the loss function. Once the loss function is minimized, its generalizability is assessed. If the neural network force field did not pass the criteria for its assessment, the cycle repeats. The n-body term in the neural-network force-field depends on the position of $n-1$ particles, described by $n(n+1)/2$ relative distances or angles, and it produces a force vector with a dimension of $\mathbb{R}^{(n-1)^2}$, which gets mapped to the Cartesian coordinates through the mapping matrix, $\mathcal{E}(\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_{n-1})$. b) The two-body term for a tagged particle (red particle) only depends on the radial distance ($|\boldsymbol{r}_1|$) of its neighboring particle (green particle) and the mapping is the unit vector between two particles ($\boldsymbol{e}_1 = \boldsymbol{r}_1/|\boldsymbol{r}_1|$). c) The three-body term depends on the position of three-particles best described by two radial distances ($|\boldsymbol{r}_1|$ and $|\boldsymbol{r}_2|$) and the angle between them (θ). The neural network predicts four values, acting on the central red particle, where the first two values, shown with the solid black lines, correspond to action-and-reaction from the first neighboring particle and the second two, shown with dashed black lines, correspond to action-and-reaction from the second neighboring particle. The mapping matrix for a three-body term is described in Eq. 5.

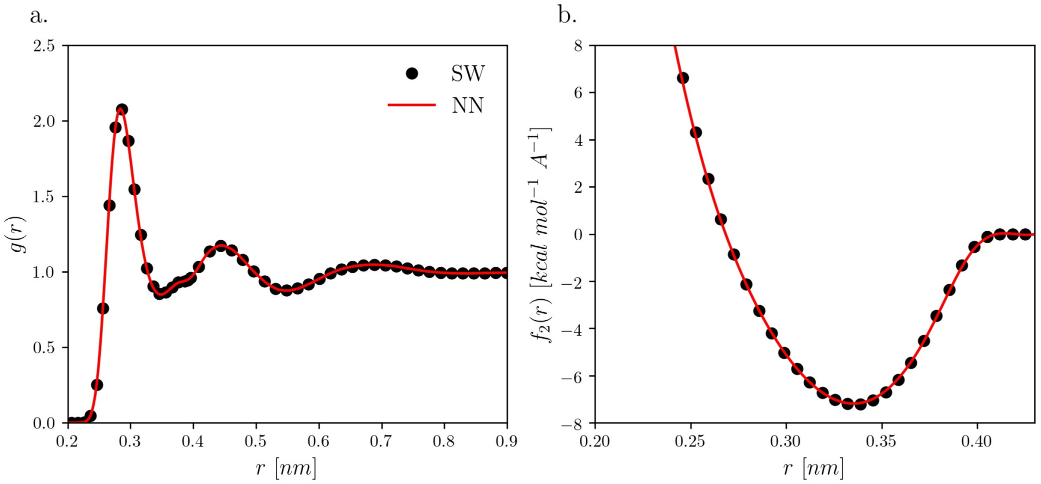

Figure 2. Radial distribution function and two-body force comparison between neural network force field and ground truth mW model. a. Comparison between RDFs of mW and NN-based models b. Comparison between two-body force of mW and NN-based models. Black circles show the mW model results and red lines show NN-based force field results.

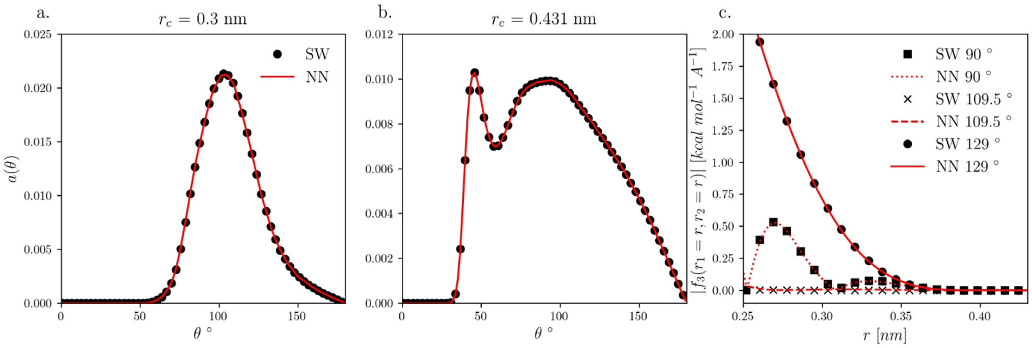

Figure 3. Angular distribution functions and three-body force comparison between neural network force field and ground truth mW model. a. Comparison between ADF mW and NN-based models with cutoff of 0.3 nm b. Comparison between ADF mW and NN-based models with cutoff of 0.43 nm c. Comparison between the norm of three-body force on the central particle of mW and NN-based models at different angles in the equidistant radial configuration. Black points show the mW model results and red lines show NN-based model results.

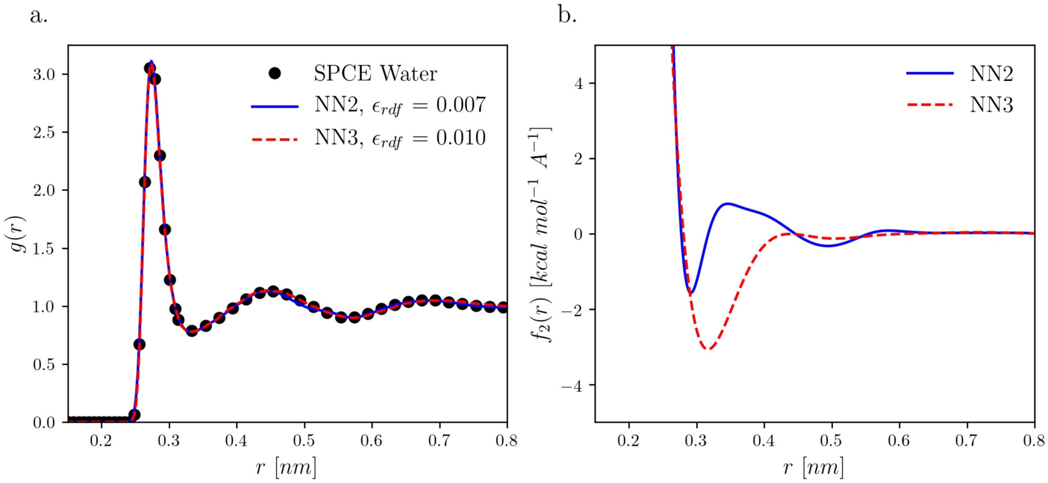

Figure 4. Radial distribution function and two-body force comparison between different neural network force fields and SPC/E model. Both NN_2^{SPCE} and NN_3^{SPCE} models have a two-body force field, however, NN_3^{SPCE} has additional NN modeling its 3-body interaction. Both models produce the RDF properties of SPC/E models. a. Comparison between RDFs of SPC/E model and NN-based models b. Comparison between two-body force of NN_2^{SPCE} and NN_3^{SPCE}; NN_2^{SPCE} shows the double-well form, while NN_3^{SPCE} has a single well, and its interaction range is shorter than the NN_2^{SPCE} model. Black circles show the SPC/E model results and blue solid and red dashed lines show the results of NN_2^{SPCE} and NN_3^{SPCE} CG models, respectively.

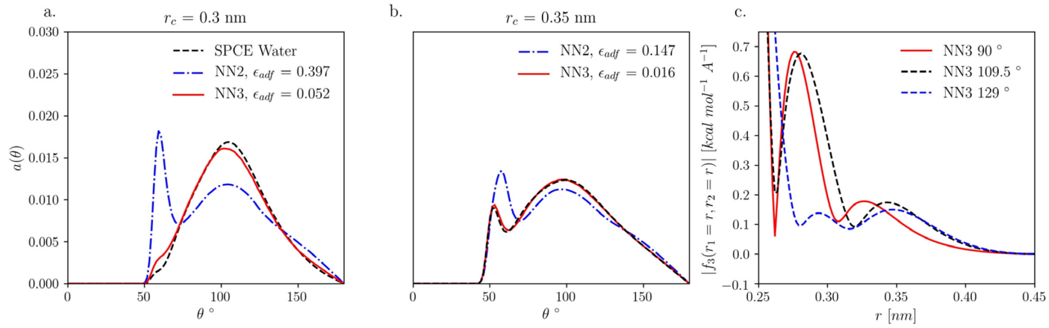

Figure 5. Angular distribution functions and three-body force comparison between different neural network force fields and SPC/E model. NN_3^{SPCE}models with three-body interaction capture ADF behavior of water far better than the NN_2^{SPCE}model. a. Comparison between ADFs of SPC/E, NN_2^{SPCE}, and NN_3^{SPCE}models within cutoff of 0.3 nm b. Comparison between ADFs of SPC/E, NN_2^{SPCE}, and NN_3^{SPCE}models within cutoff of 0.35 nm c. Norm of three-body force on the central particle of NN_3^{SPCE}models at different angles in the equidistant radial configuration. In a and b, black dashed lines show the SPC/E model results and dashed blue and solid red lines show the results of $NN_2^{SPCE}$and NN_3^{SPCE}CG models. All lines in part c show the norm of the force of 3-body interaction.

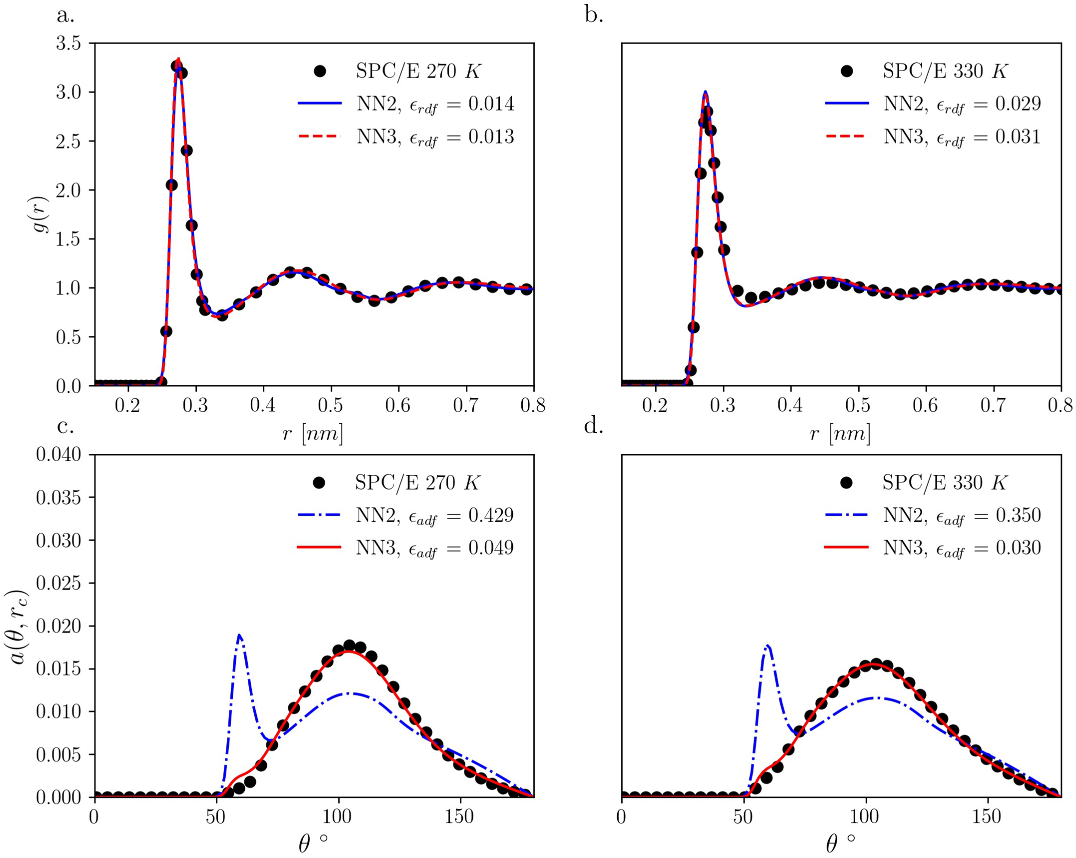

Figure 6. Transferability analysis for the structural properties at 270 K and 330 K between the SPC/E and NN_2^{SPCE} and NN_3^{SPCE} CG models. a. RDFs of SPC/E, NN_2^{SPCE}, and NN_3^{SPCE} models at 270 K. b. RDF of NN_2^{SPCE} and NN_3^{SPCE} at 330 K. c. comparison between ADFs of SPC/E, NN_2^{SPCE}, and NN_3^{SPCE}models within cutoff of 0.3 nm at 270 K. d. comparison between ADFs of SPC/E, NN_2^{SPCE}, and NN_3^{SPCE}models within cutoff of 0.3 nm at 330 K. In a and b, black circles show the SPC/E model results and dashed blue and solid red lines show the results of $NN_2^{SPCE}$and NN_3^{SPCE}CG models.

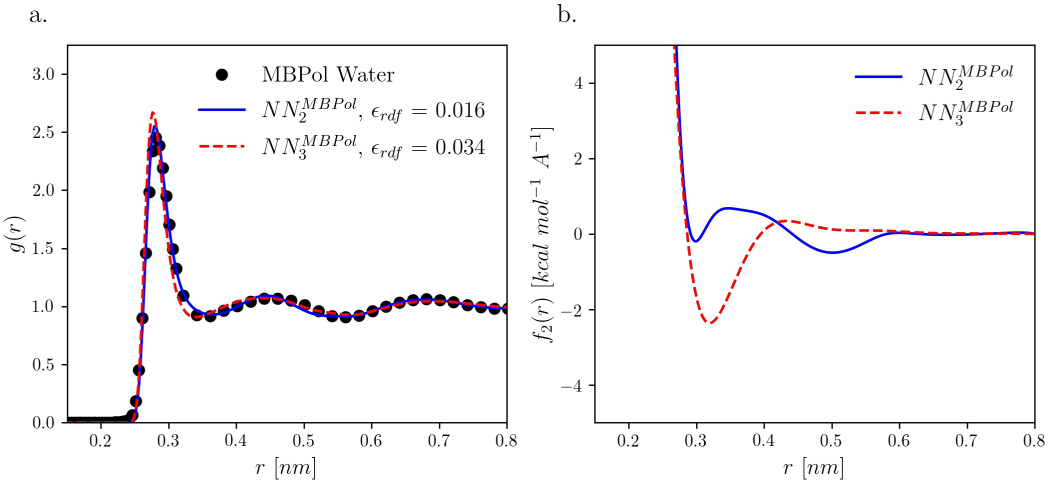

Figure 7. Radial distribution function and two-body force comparison between different neural network force fields and MB-pol model. Both $NN_2^{MBPol}$and NN_3^{MBPol}models have a two-body force field, however, NN_3^{SPCE}has an additional 3-body interaction. Both models reproduce the RDF properties of MB-pol models. a. Comparison between RDFs of MB-pol model and NN-based models b. Comparison between two-body force of $NN_2^{MBPol}$and NN_3^{MBPol}; NN_2^{MBPol}shows the double-well form, while NN_3^{MBPol}has a single well, and its interaction range is shorter than the NN_2^{MBPol}model. Black circles show the MB-pol model results and blue solid and red dashed lines show the results of $NN_2^{MBPol}$and NN_3^{MBPol} CG models, respectively.

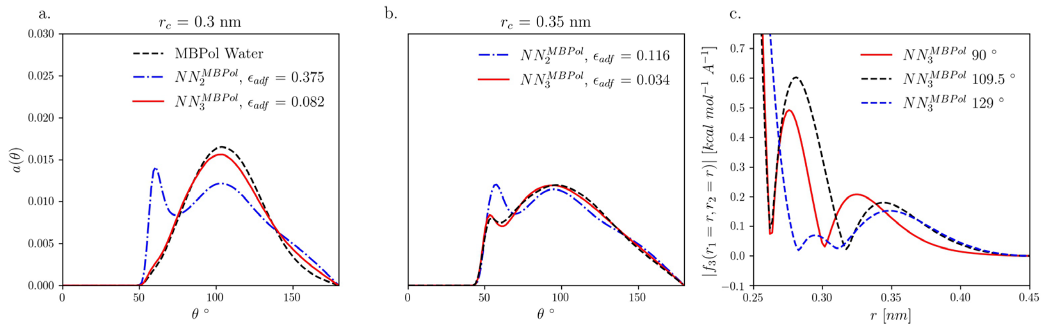

Figure 8. Angular distribution functions and three-body force comparison between different neural network force fields and MB-pol model. NN_3^{MBPol} models with three-body interaction capture ADF behavior of water far better than the NN_2^{MBPol} model. a. Comparison between ADFs of MB-pol, NN_2^{MBPol}, and NN_3^{MBPol}models within cutoff of 0.3 nm b. Comparison between ADFs of MB-pol, NN_2^{MBPol}, and NN_3^{MBPol}models within cutoff of 0.35 nm c. Norm of three-body force on the central particle of NN_3^{MBPol}models at different angles in the equidistant radial configuration. In a and b, black dashed lines show the MB-pol model results and dashed blue and solid red lines show the results of $NN_2^{MBPol}$and NN_3^{MBPol}CG models, respectively. All lines in part c show the norm of the force of 3-body interaction.

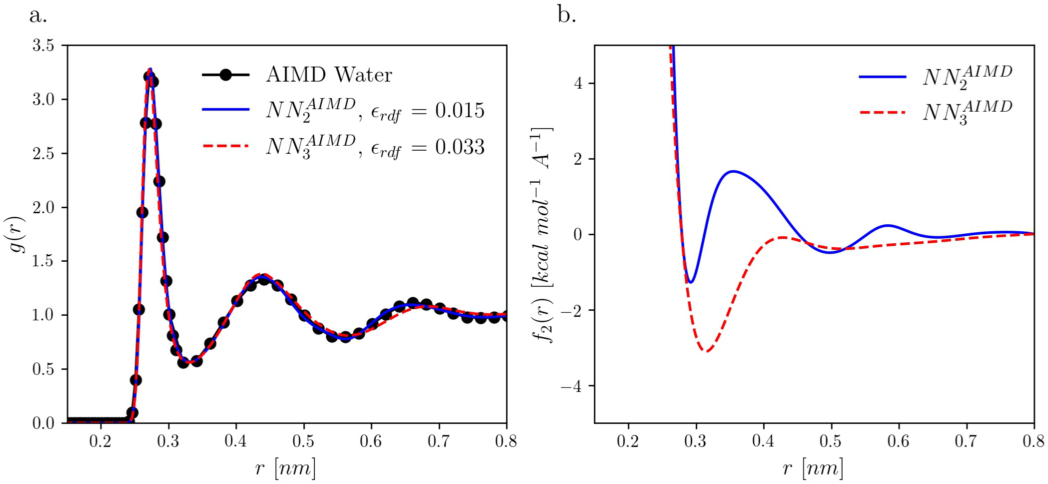

Figure 9. Radial distribution function and two-body force comparison between different neural network force fields and AIMD model. Both NN_2^{AIMD} and NN_3^{AIMD} models have a two-body force field, however, NN_3^{AIMD} has an additional 3-body interaction. Both models reproduce the RDF properties of AIMD models. a. Comparison between RDFs of AIMD model and NN-based models b. Comparison between two-body force of NN_2^{AIMD} and NN_3^{AIMD} ; NN_2^{AIMD} shows the double-well form, while NN_3^{AIMD} has a single well, and its interaction range is far shorter than the NN_2^{AIMD} model. Black circles show the AIMD model results and blue solid and red dashed lines show the results of NN_2^{AIMD} and NN_3^{AIMD} CG models, respectively.

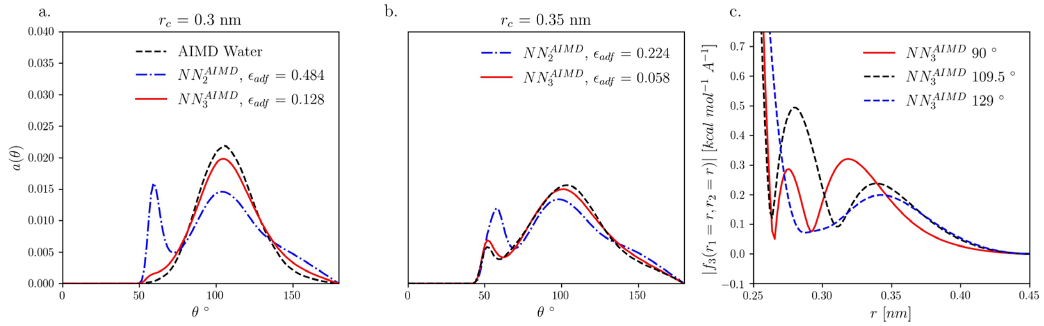

Figure 10. Angular distribution functions and three-body force comparison between different neural network force fields and AIMD model. NN_3^{AIMD} models with three-body interaction capture ADF behavior of water better than the NN_2^{AIMD} model. a. Comparison between ADFs of AIMD, NN_2^{AIMD}, and NN_3^{AIMD} models within cutoff of 0.3 nm b. Comparison between ADFs of AIMD, NN_2^{AIMD}, and NN_3^{AIMD} models within cutoff of 0.35 nm c. Norm of three-body force on the central particle of NN_3^{AIMD} models at different angles in the equidistant radial configuration. In a and b, black dashed lines show the AIMD model results and dashed blue and solid red lines show the results of NN_2^{AIMD} and NN_3^{AIMD} CG models, respectively. All lines in part c show the norm of the force of 3-body interaction.

Table I. Comparison between the isothermal compressibility and thermal expansion coefficient of CG and reference models at 300 *K*.

	$\boldsymbol{SPC/E}$	$\boldsymbol{NN_2^{SPC/E}}$	$\boldsymbol{NN_3^{SPC/E}}$
$\boldsymbol{k_T \times 10^6\ [bar^{-1}]}$	$\boldsymbol{45.68 \pm 0.69}$	$\boldsymbol{69.97 \pm 1.27}$	$\boldsymbol{27.98 \pm 0.81}$
$\boldsymbol{\alpha_p \times 10^6\ [K^{-1}]}$	$\boldsymbol{493.8 \pm 4.0}$	$\boldsymbol{1243.7 \pm 7.6}$	$\boldsymbol{420.5 \pm 8.0}$

Table II. Transferability of CG models in terms of density change and diffusion coefficient.

$T\ [K]$		SPC/E	$NN_2^{SPC/E}$	$NN_3^{SPC/E}$
270	$\rho\ [\#/nm^3]$	33.71	33.74	33.80
	$D\ [10^{-5}\ cm^2/s\ ]$	1.36	12.64	9.32
300	$\rho\ [\#/nm^3]$	33.38	33.40	33.40
	$D\ [10^{-5}\ cm^2/s\ ]$	2.68	14.73	12.09
330	$\rho\ [\#/nm^3]$	33.76	33.20	32.98
	$D\ [10^{-5}\ cm^2/s\ ]$	4.60	16.58	14.81

References

(1) Hospital, A.; Goñi, J. R.; Orozco, M.; Gelpí, J. L. Molecular Dynamics Simulations: Advances and Applications. *Adv. Appl. Bioinform. Chem.* **2015**, *8*, 37–47. https://doi.org/10.2147/AABC.S70333.

(2) 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. https://doi.org/10.1103/PhysRevLett.120.143001.

(3) Nejat Pishkenari, H.; Yousefi, F. S.; Taghibakhshi, A. Determination of Surface Properties and Elastic Constants of FCC Metals: A Comparison among Different EAM Potentials in Thin Film and Bulk Scale. *Mater. Res. Express* **2019**, *6* (1), 015020. https://doi.org/10.1088/2053-1591/aae49b.

(4) Zhang, L.; Lin, D.-Y.; Wang, H.; Car, R. Active Learning of Uniformly Accurate Interatomic Potentials for Materials Simulation. *Phys. Rev. Mater.* **2019**, *3*, 23804. https://doi.org/10.1103/PhysRevMaterials.3.023804.

(5) Stillinger, F. H.; Weber, T. A. Computer Simulation of Local Order in Condensed Phases of Silicon. *Phys. Rev. B* **1985**, *31* (8), 5262–5271. https://doi.org/10.1103/PhysRevB.31.5262.

(6) Molinero, V.; Moore, E. B. Water Modeled as an Intermediate Element between Carbon and Silicon. *J. Phys. Chem. B* **2009**, *113* (13), 4008–4016. https://doi.org/10.1021/jp805227c.

(7) Mayne, C. G.; Saam, J.; Schulten, K.; Tajkhorshid, E.; Gumbart, J. C. Rapid Parameterization of Small Molecules Using the Force Field Toolkit. *J. Comput. Chem.* **2013**, *34* (32), 2757–2770. https://doi.org/10.1002/jcc.23422.

(8) Putin, E.; Asadulaev, A.; Vanhaelen, Q.; Ivanenkov, Y.; Aladinskaya, A. V.; Aliper, A.; Zhavoronkov, A. Adversarial Threshold Neural Computer for Molecular *de Novo* Design. *Mol.*

Pharm. **2018**, acs.molpharmaceut.7b01137. https://doi.org/10.1021/acs.molpharmaceut.7b01137.

(9) Moradzadeh, A.; Aluru, N. R. Molecular Dynamics Properties without the Full Trajectory: A Denoising Autoencoder Network for Properties of Simple Liquids. *J. Phys. Chem. Lett.* **2019**, 7568–7576. https://doi.org/10.1021/acs.jpclett.9b02820.

(10) Moradzadeh, A.; Aluru, N. R. Transfer-Learning-Based Coarse-Graining Method for Simple Fluids: Toward Deep Inverse Liquid-State Theory. *J. Phys. Chem. Lett.* **2019**, *10* (6), 1242–1250. https://doi.org/10.1021/acs.jpclett.8b03872.

(11) Moradzadeh, A.; Aluru, N. R. Understanding Simple Liquids through Statistical and Deep Learning Approaches. *J. Chem. Phys.* **2021**, *154* (20), 204503. https://doi.org/10.1063/5.0046226.

(12) Kadupitiya, J.; Fox, G. C.; Jadhao, V. Deep Learning Based Integrators for Solving Newton's Equations with Large Timesteps. **2020**.

(13) Sharma, P.; Jadhao, V. Molecular Dynamics Simulations on Cloud Computing and Machine Learning Platforms. *2021 IEEE 14th Int. Conf. Cloud Comput.* **2021**, 751–753. https://doi.org/10.1109/CLOUD53861.2021.00101.

(14) Jeong, J.; Moradzadeh, A.; Aluru, N. R. Extended DeepILST for Various Thermodynamic States and Applications in Coarse-Graining. *J. Phys. Chem. A* **2022**, acs.jpca.1c10865. https://doi.org/10.1021/ACS.JPCA.1C10865.

(15) Scherer, C.; Scheid, R.; Andrienko, D.; Bereau, T. Kernel-Based Machine Learning for Efficient Simulations of Molecular Liquids. *J. Chem. Theory Comput.* **2020**, *16* (5), 3194–3204. https://doi.org/10.1021/acs.jctc.9b01256.

(16) Glielmo, A.; Sollich, P.; De Vita, A. Accurate Interatomic Force Fields via Machine Learning with Covariant Kernels. *Phys. Rev. B* **2017**, *95*. https://doi.org/10.1103/PhysRevB.95.214302.

(17) Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High-Dimensional

Potential-Energy Surfaces. *Phys. Rev. Lett.* **2007**, *98* (14), 146401. https://doi.org/10.1103/PhysRevLett.98.146401.

(18) Nguyen, T. T.; Székely, E.; Imbalzano, G.; Behler, J.; Csányi, G.; Ceriotti, M.; Götz, A. W.; Paesani, F. Comparison of Permutationally Invariant Polynomials, Neural Networks, and Gaussian Approximation Potentials in Representing Water Interactions through Many-Body Expansions. *J. Chem. Phys.* **2018**, *148* (24), 241725. https://doi.org/10.1063/1.5024577.

(19) Hernandez, A.; Balasubramanian, A.; Yuan, F.; Mason, S. A. M.; Mueller, T. Fast, Accurate, and Transferable Many-Body Interatomic Potentials by Symbolic Regression. *npj Comput. Mater.* **2019**, *5* (1), 1–11. https://doi.org/10.1038/s41524-019-0249-1.

(20) Chan, H.; Cherukara, M. J.; Narayanan, B.; Loeffler, T. D.; Benmore, C.; Gray, S. K.; Sankaranarayanan, S. K. R. S. Machine Learning Coarse Grained Models for Water. *Nat. Commun.* **2019**, *10* (1), 379. https://doi.org/10.1038/s41467-018-08222-6.

(21) Wang, J.; Olsson, S.; Wehmeyer, C.; Pérez, A.; Charron, N. E.; De Fabritiis, G.; Noé, F.; Clementi, C. Machine Learning of Coarse-Grained Molecular Dynamics Force Fields. *ACS Cent. Sci.* **2019**, *5* (5), 755–767. https://doi.org/10.1021/acscentsci.8b00913.

(22) Wang, W.; Gómez-Bombarelli, R. Coarse-Graining Auto-Encoders for Molecular Dynamics. *npj Comput. Mater.* **2019**, *5* (1), 1–9. https://doi.org/10.1038/s41524-019-0261-5.

(23) Zhang, L.; Han, J.; Wang, H.; Car, R.; Weinan, W. E. DeePCG: Constructing Coarse-Grained Models via Deep Neural Networks. *J. Chem. Phys.* **2018**, *149* (3), 034101. https://doi.org/10.1063/1.5027645.

(24) Wang, H.; Zhang, L.; Han, J.; E, W. DeePMD-Kit: A Deep Learning Package for Many-Body Potential Energy Representation and Molecular Dynamics. *Comput. Phys. Commun.* **2018**, *228*, 178–184. https://doi.org/10.1016/j.cpc.2018.03.016.

(25) Aly, A.; Guadagni, G.; Dugan, J. B. Derivative-Free Optimization of Neural Networks Using Local Search. In *2019 IEEE 10th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference, UEMCON 2019*; Institute of Electrical and Electronics Engineers Inc., 2019; pp 0293–0299. https://doi.org/10.1109/UEMCON47517.2019.8993007.

(26) Reddy, S. K.; Straight, S. C.; Bajaj, P.; Pham, C. H.; Riera, M.; Moberg, D. R.; Morales, M. A.; Knight, C.; Götz, A. W.; Paesani, F. On the Accuracy of the MB-Pol Many-Body Potential for Water: Interaction Energies, Vibrational Frequencies, and Classical Thermodynamic and Dynamical Properties from Clusters to Liquid Water and Ice. *J. Chem. Phys.* **2016**, *145* (19), 194504. https://doi.org/10.1063/1.4967719.

(27) Nilsson, A.; Pettersson, L. G. M. The Structural Origin of Anomalous Properties of Liquid Water. *Nat. Commun. 2015 61* **2015**, *6* (1), 1–11. https://doi.org/10.1038/ncomms9998.

(28) Monroe, J. I.; Shell, M. S. Decoding Signatures of Structure, Bulk Thermodynamics, and Solvation in Three-Body Angle Distributions of Rigid Water Models. *J. Chem. Phys.* **2019**, *151* (9), 094501. https://doi.org/10.1063/1.5111545.

(29) Hamm, P. Markov State Model of the Two-State Behaviour of Water. *J. Chem. Phys.* **2016**, *145* (13), 134501. https://doi.org/10.1063/1.4963305.

(30) Chaimovich, A.; Shell, M. S. Anomalous Waterlike Behavior in Spherically-Symmetric Water Models Optimized with the Relative Entropy. *Phys. Chem. Chem. Phys.* **2009**, *11* (12), 1901. https://doi.org/10.1039/b818512c.

(31) Errington, J. R.; Debenedetti, P. G. Relationship between Structural Order and the Anomalies of Liquid Water. *Nature* **2001**, *409* (6818), 318–321. https://doi.org/10.1038/35053024.

(32) Medders, G. R.; Babin, V.; Paesani, F. A Critical Assessment of Two-Body and Three-Body Interactions in Water. *J. Chem. Theory Comput.* **2013**, *9* (2), 1103–1114.

https://doi.org/10.1021/CT300913G.

(33) Motevaselian, M. H.; Mashayak, S. Y.; Aluru, N. R. Extended Coarse-Grained Dipole Model for Polar Liquids: Application to Bulk and Confined Water. *Phys. Rev. E* **2018**, *98* (5), 052135. https://doi.org/10.1103/PhysRevE.98.052135.

(34) Shell, M. S. The Relative Entropy Is Fundamental to Multiscale and Inverse Thermodynamic Problems. *J. Chem. Phys.* **2008**, *129* (14), 1–7. https://doi.org/10.1063/1.2992060.

(35) Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **1995**, *117* (1), 1–19. https://doi.org/10.1006/JCPH.1995.1039.

(36) Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L.-P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D.; Wiewiora, R. P.; Brooks, B. R.; Pande, V. S. OpenMM 7: Rapid Development of High Performance Algorithms for Molecular Dynamics. *PLOS Comput. Biol.* **2017**, *13* (7), e1005659. https://doi.org/10.1371/JOURNAL.PCBI.1005659.

(37) Tanaka, H.; Nakanishi, K.; Watanabe, N. Constant Temperature Molecular Dynamics Calculation on Lennard-Jones Fluid and Its Application to Watera). *J. Chem. Phys.* **1998**, *78* (5), 2626. https://doi.org/10.1063/1.445020.

(38) VandeVondele, J.; Hutter, J. Gaussian Basis Sets for Accurate Calculations on Molecular Systems in Gas and Condensed Phases. *J. Chem. Phys.* **2007**, *127* (11), 114105. https://doi.org/10.1063/1.2770708.

(39) Krack, M. Pseudopotentials for H to Kr Optimized for Gradient-Corrected Exchange-Correlation Functionals. *Theor. Chem. Acc.* **2005**, *114* (1–3), 145–152. https://doi.org/10.1007/s00214-005-0655-y.

(40) Chen, M.; Ko, H. Y.; Remsing, R. C.; Calegari Andrade, M. F.; Santra, B.; Sun, Z.; Selloni, A.; Car, R.; Klein, M. L.; Perdew, J. P.; Wu, X. Ab Initio Theory and Modeling of Water. *Proc. Natl.*

Acad. Sci. U. S. A. **2017**, *114* (41), 10846–10851. https://doi.org/10.1073/pnas.1712499114.

(41) Gartner, T. E.; Zhang, L.; Piaggi, P. M.; Car, R.; Panagiotopoulo, A. Z.; Debenedet, P. G. Signatures of a Liquid-Liquid Transition in an Ab Initio Deep Neural Network Model for Water. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117* (42), 26040–26046. https://doi.org/10.1073/pnas.2015440117.

(42) Nosé, S. A Unified Formulation of the Constant Temperature Molecular Dynamics Methods. *J. Chem. Phys.* **1984**, *81* (1), 511. https://doi.org/10.1063/1.447334.

(43) LeCun, Y.; Bengio, Y.; Hinton, G. Deep Learning. *Nature* **2015**, *521* (7553), 436–444. https://doi.org/10.1038/nature14539.

(44) Goh, G. B.; Hodas, N. O.; Vishnu, A. Deep Learning for Computational Chemistry. *J. Comput. Chem.* **2017**, *38* (16), 1291–1307. https://doi.org/10.1002/jcc.24764.

(45) Wang, Y. T.; Izvekov, S.; Yan, T. Y.; Voth, G. A. Multiscale Coarse-Graining of Ionic Liquids. *J. Phys. Chem. B* **2006**, *110* (8), 3564–3575. https://doi.org/Doi 10.1021/Jp0548220.

(46) Kingma, D. P.; Ba, J. Adam: A Method for Stochastic Optimization. **2014**.

(47) Al-Sahaf, H.; Bi, Y.; Chen, Q.; Lensen, A.; Mei, Y.; Sun, Y.; Tran, B.; Xue, B.; Zhang, M. A Survey on Evolutionary Machine Learning. *Journal of the Royal Society of New Zealand*. Taylor and Francis Asia Pacific April 3, 2019, pp 205–228. https://doi.org/10.1080/03036758.2019.1609052.

(48) Drugan, M. M. Reinforcement Learning versus Evolutionary Computation: A Survey on Hybrid Algorithms. *Swarm Evol. Comput.* **2019**, *44*, 228–246. https://doi.org/10.1016/j.swevo.2018.03.011.

(49) Moradzadeh, A.; Motevaselian, M. H.; Mashayak, S. Y.; Aluru, N. R. Coarse-Grained Force Field for Imidazolium-Based Ionic Liquids. *J. Chem. Theory Comput.* **2018**, *14* (6), 3252–3261.

https://doi.org/10.1021/acs.jctc.7b01293.

(50) Mashayak, S. Y.; Jochum, M. N.; Koschke, K.; Aluru, N. R.; Rühle, V.; Junghans, C. Relative Entropy and Optimization-Driven Coarse-Graining Methods in VOTCA. *PLoS One* **2015**, *10* (7). https://doi.org/10.1371/journal.pone.0131754.

(51) Cho, C. H.; Singh, S.; Robinson, G. W. Liquid Water and Biological Systems: The Most Important Problem in Science That Hardly Anyone Wants to See Solved. *Faraday Discuss.* **1996**, *103* (0), 19–27. https://doi.org/10.1039/fd9960300019.

(52) Barros De Oliveira, A.; Neves, E. B.; Gavazzoni, C.; Paukowski, J. Z.; Netz, P. A.; Barbosa, M. C. Liquid Crystal Phase and Waterlike Anomalies in a Core-Softened Shoulder-Dumbbells System. *J. Chem. Phys.* **2010**, *132* (16), 164505. https://doi.org/10.1063/1.3386384.

(53) Duboué-, E.; Dijon, D.-; Laage, D. Characterization of the Local Structure in Liquid Water by Various Order Parameters. *J. Phys. Chem. B* **2015**, *119*, 26. https://doi.org/10.1021/acs.jpcb.5b02936.

(54) Chaimovich, A.; Shell, M. S. Tetrahedrality and Structural Order for Hydrophobic Interactions in a Coarse-Grained Water Model. *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.* **2014**, *89* (2), 1–13. https://doi.org/10.1103/PhysRevE.89.022140.

(55) Bereau, T.; Rudzinski, J. F. Accurate Structure-Based Coarse Graining Leads to Consistent Barrier-Crossing Dynamics. *Phys. Rev. Lett.* **2018**, *121* (25), 256002. https://doi.org/10.1103/PhysRevLett.121.256002.

(56) Wang, H.; Stillinger, F. H.; Torquato, S. Sensitivity of Pair Statistics on Pair Potentials in Many-Body Systems. *J. Chem. Phys.* **2020**, *153* (12), 124106. https://doi.org/10.1063/5.0021475.

(57) Chen, B.; Ivanov, I.; Klein, M. L.; Parrinello, M. Hydrogen Bonding in Water. *Phys. Rev. Lett.* **2003**, *91* (21), 215503. https://doi.org/10.1103/PhysRevLett.91.215503.

TOC