Thermodynamics of Water and Ice from a Fast and Scalable First-Principles Neuroevolution Potential

Zekun Chen,[†] Margaret L. Berrens,[†] Kam-Tung Chan,[†] Zheyong Fan,[‡] and Davide Donadio[*,†]

[†]*Department of Chemistry, University of California, Davis, Davis, CA 95616, USA*

[‡]*College of Physical Science and Technology, Bohai University, Jinzhou 121013, P. R. China*

E-mail: ddonadio@ucdavis.edu

Abstract

Machine learning potentials enable molecular dynamics simulations to exceed the size and time scales that can be accessed by first-principles methods like density functional theory, while still maintaining the accuracy of the underlying training dataset. However, accurate machine learning potentials come with relatively high computational costs that limit their ability to predict properties requiring extensive sampling, large simulation cells, or long runs to converge. Here, we have developed and tested a neuroevolution-potential model for water trained to hybrid dispersion-corrected density functional calculations. This model exhibits accuracy and transferability comparable to state-of-the-art machine learning potentials but at a much lower computational cost. As a result, it enabled us to compute well-converged thermodynamics averages and fluctuations. This allowed us to assess the ability of our model to reproduce several thermodynamic properties of water and ice, as well as the anomalous behavior of

water density, heat capacity, and compressibility. The efficient GPU acceleration of our model and its capability to reproduce water thermodynamics in good agreement with experiments make it suitable for simulating phase transitions and slow dynamical processes in aqueous systems.

Introduction

Water is a ubiquitous substance on Earth and plays a crucial role in various fields, including energy harvesting,[1] biological processes,[2] and the evolution of life.[3–5] Despite its simple chemical formula, water is a complex system with multiple anomalies and a rich phase diagram, attributed to the nature of hydrogen bonding and the structural versatility of tetrahedral networks.[3,6–8] The intricate thermodynamics of water has been the focus of numerous molecular dynamics (MD) simulations aimed at providing molecular-level insights into water's anomalies, the melting and freezing mechanisms of ice, and the features of ice and water surfaces. These phenomena typically involve extended time and size scales, which are inaccessible to first-principles molecular dynamics (FPMD) simulations. Therefore, the thermodynamics of ice and water have been investigated employing less transferable but computationally efficient empirical force fields, such as rigid point-charge models[9] and even faster coarse-grained potentials,[10] fitted to experimental data. Higher accuracy with good computational performance can be achieved by training classical polarizable models on a suitable combination of experimental data and quantum-chemical calculations.[11] Furthermore, more complex polarizable forcefields fitted on *ab initio* quantum chemical calculations on water clusters demonstrated exceptional accuracy for molecular water systems.[12–14] However, the form of these potentials limits the range of utilization to non-reactive systems, and the complexity of their functional form hampers their performance. Force-matching and splined two-body potentials are very efficient but lack transferability over different thermodynamic conditions.[15,16]

In the past two decades, the development of machine learning potentials (MLPs) fitted

on first-principles, e.g. density functional theory (DFT), potential energy surfaces[17–25] has enabled the accurate and transferable atomistic modeling of physical phenomena that require systems up to 10^5 atoms and simulation times beyond the nanoseconds time scales previously accessible only to empirical approaches.[26,27] For example, the use of Behler-Parrinello neural network potentials (NNP) made it possible to unravel the role of van der Waals dispersion forces on the density anomaly of water,[28] the impact of nuclear quantum effects on the thermodynamic stability of ice and water,[29] the phase diagram of water at extreme conditions,[30] ice nucleation,[31] and the sub-picosecond relaxation dynamics of hydrogen bonds upon vibrational excitation.[32]

One major advantage of local MLPs over DFT is that the computational cost to compute the forces scales linearly with the size of the system. However, the computational effort to evaluate MLP energies and forces remains much larger than that of empirical forcefields, and it scales with the desired accuracy, as they both depend on the complexity of the local atomic descriptors and on the ML engine.[33] Some of the applications mentioned above, e.g. ice nucleation, are at the limit of what is affordable with NNP, and applications requiring larger models or longer timescales are still precluded, such as ice amorphization, liquid-liquid phase transitions, and the slow dynamics of interfaces. Aggressive parallelization and porting on graphic processing units (GPUs) are possible strategies to extend the reach of accurate MLPs. Neuroevolution-potentials (NEP) are a class of neural network interatomic potentials optimized to achieve efficient training and high performance on GPUs.[34,35] In this work, we construct and validate a NEP for water and ice with accuracy comparable to state-of-the-art neural network potentials and high computational efficiency comparable to empirical force fields. While the usual metric to assess the accuracy of MLPs is the error on energies and forces (either mean absolute error - MAE - or root mean square error - RMSE), here we provide an extended validation of our NEP water model against a set of thermodynamic observables for water and ice. Additionally, we test the transferability of the NEP water model to two systems that were not included in the training set, and we

assess its performance to reproduce the anomalous behavior of density, heat capacity, and isothermal compressibility of water.

In the next Section, we outline the formulation of NEP, summarize the processes of generating the training and testing database, and describe the simulation protocols used to model water thermodynamics by MD. In the Results Section, we first compare the quality of the fit of our NEP potential to that of the committee neural network potential(C-NNP)[36] over the same training-testing set. We then verify the ability of our model to reproduce the structure of liquid water with both classical MD and path-integral simulations. We then show that the NEP potential reproduces the FPMD dynamics of a small water cluster, and it can be exploited to study systematically the vibrational response as a function of the water cluster size. Later we present the calculation of a set of thermodynamic properties of water and ice and we compare it to the results obtained with the TIP4P/2005 forcefield.[9] Finally, we address the study of the anomalous behavior of water density, heat capacity, and isothermal compressibility as a function of temperature at ambient pressure, and we conclude by discussing the strengths, limitations, and prospective use of this NEP model in future studies.

Models and Methods

Neuroevolution Potentials

The NEP approach was recently proposed as a straightforward and data-efficient method to develop and fit neural network MLPs. Previous works demonstrated that NEP achieves excellent performance in the study of thermal transport in both crystalline and amorphous materials.[34,37–39] In this work, we used the NEP3[35] framework to construct our water model. Hereafter we provide a brief overview of the general features of NEP. The descriptors of NEP are made up of radial and angular components where the radial component, d_n^i, can

be expressed as:

$$d_n^i = \sum_{j \neq i} g_n\left(r_{ij}\right) \quad \text{for} \quad 0 \leq n \leq n_{\max}^{\mathrm{R}}, \tag{1}$$

where r_{ij} stands for the inter-atomic distances between atom i and atom j. $g_n\left(r_{ij}\right)$ in Eq. (1) can be expanded on the basis of Chebyshev polynomials:[40]

$$g_n\left(r_{ij}\right) = \sum_{k=0}^{N_{\mathrm{bas}}^{\mathrm{R}}} c_{nk}^{ij} f_k\left(r_{ij}\right), \quad \text{and} \tag{2}$$

$$f_k\left(r_{ij}\right) = \frac{1}{2}\left[T_k\left(2\left(\frac{r_{ij}}{r_c^R} - 1\right)^2 - 1\right) + 1\right] f_{\mathrm{cut}}\left(r_{ij}\right). \tag{3}$$

In Eq. (3), T_k stands for the k^{th} order Chebyshev polynomial and has dimension of N_{bas}^R+1. $r_{\mathrm{c}}^{\mathrm{R}}$ is the radial cutoff, and $f_{\mathrm{cut}}\left(r_{ij}\right)$ in Eq. (3) represents a Tersoff cutoff function.[41] Meanwhile, the angular component can be written using the addition theorem of the spherical harmonics:

$$d_{n\ell}^i = \sum_{m=-\ell}^{\ell} (-1)^m A_{n\ell m}^i A_{n\ell(-m)}^i. \tag{4}$$

In Eq. (4), n and ℓ fulfill the constraints such that $0 \leq n \leq n_{\max}^{\mathrm{A}}$ and $1 \leq \ell \leq \ell_{3b}^{max}$. $A_{n\ell m}^i$ can be further described as:

$$A_{n\ell m}^i = \sum_{j \neq i} g_n\left(r_{ij}\right) Y_{\ell m}\left(\theta_{ij}, \phi_{ij}\right), \tag{5}$$

where $Y_{\ell m}\left(\theta_{ij}, \phi_{ij}\right)$ are the spherical harmonics of order ℓ, m for the polar and azimuthal angles θ_{ij} and ϕ_{ij} formed by the vector $\mathbf{r_{ij}}$. The basis functions of $g_n\left(r_{ij}\right)$ in Eq.(5) are in dimension of $N_{bas}^A + 1$. After deriving the 3-body angular term $d_{n\ell}^i$, 4 and 5-body angular

terms, $d^i_{n\ell_1,\ell_2,\ell_3}$ and $d^i_{n\ell_1,\ell_2,\ell_3\ell_4}$, can be defined accordingly:

$$d^i_{n\ell_1\ell_2\ell_3} = \sum_{m_1=-\ell_1}^{\ell_1} \sum_{m_2=-\ell_2}^{\ell_2} \sum_{m_3=-\ell_3}^{\ell_3} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \times A^i_{n\ell_1 m_1} A^i_{n\ell_2 m_2} A^i_{n\ell_3 m_3}. \tag{6}$$

$$d^i_{n\ell_1\ell_2\ell_3\ell_4} = \sum_{m_1=-\ell_1}^{\ell_1} \sum_{m_2=-\ell_2}^{\ell_2} \sum_{m_3=-\ell_3}^{\ell_3} \sum_{m_4=-\ell_4}^{\ell_4} \begin{bmatrix} \ell_1 & \ell_2 & \ell_3 & \ell_4 \\ m_1 & m_2 & m_3 & m_4 \end{bmatrix} A^i_{n\ell_1 m_1} A^i_{n\ell_2 m_2} A^i_{n\ell_3 m_3} A^i_{n\ell_4 m_4}, \tag{7}$$

where $\begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix}$ represents Wigner $3j$ symbols.[42] $d^i_{n\ell}$, $d^i_{n\ell_1\ell_2\ell_3}$ and $d^i_{n\ell_1\ell_2\ell_3,\ell_4}$ are modified from the atom cluster expansion (ACE) formalism.[42–45] To develop our NEP potential, $\ell^{max}_{3b} = 4$, $\ell^{max}_{4b} = 2$ and $\ell^{max}_{5b} = 1$ are employed in computing the angular descriptors. Once radial and angular components are determined, the descriptors will be constructed as the input layer of a feed-forward neural network (FFNN).[46] The FFNN consists of a single hidden layer with N_{neu} neurons with hyperbolic tangent activation function. Among different NN potentials,[17,45,47,48] NEP is unique in that its weights are optimized using the separable natural evolution strategy (SENS).[49] Since Virial data are not available from the training set, the loss function of our NEP potential is represented by the sum of the RMSE of energy and force, suitably weighted. To avoid overfitting, both norm-1 (λ_1) and norm-2 (λ_2) regularizations are imposed upon training.[50,51]

A comprehensive list of hyperparameters for our NEP potential is reported in Table 1, which also indicates the size of the population (N_{pop}) and the number of generations (N_{gen}) used in the SENS optimization. The hyperparameters are chosen so as to obtain an ideal balance between accuracy and computational efficiency. For the radial cutoff, we have tested values starting from 6 Å, finding that we need at least $r^R_c = 8$ Å to reproduce the structure of water. At the same time, we verified that increasing the angular cutoff does not improve

the potential. Similar principles guided the choice of the order of the basis functions and of the order of interactions considered, which is up to 5-body. Specifically, we observed that increasing $N_{bas}^{R/A}$ beyond 8 Ådoes not improve significantly the error on the forces but adds to the computational cost. Finally, we adopt a single-layer feed-forward neural network with 50 neurons, which also optimizes performance maintaining transferability over several water and ice systems.

Table 1: Hyper-parameters for the water/ice NEP potential.

r_c^R(Å)	r_c^A(Å)	N_{bas}^R	N_{bas}^A	n_{max}^R
8.0	6.0	8	8	8
n_{max}^A	$\ell_{3b}^{max} - \ell_{4b}^{max} - \ell_{5b}^{max}$	N_{neu}	λ_e	λ_f
8	4 - 2 - 1	50	1.0	1.0
λ_1	λ_2	N_{pop}	N_{gen}	
0.05	0.05	50	2×10^5	

For training and testing our NEP model, we use the dataset created by Schran *et al.*[36] to train their C-NNP. The main body of the training set consists of structures of liquid water at 300 K from FPMD simulations, in which the dynamics of the nuclei are classical.[52–54] The training set was then expanded at each cycle of C-NNP development to gradually capture the nuclear quantum effects via path integral molecular dynamics (PIMD) simulations. Upon the final stage of training the C-NNP, 814 structures were included in the training set, which encompasses liquid water at 300, 350, and 400 K, high-pressure liquid water at 2 and 4 kbar, water slab representing the air-water interface,[55] and crystalline ice I_h and ice VIII. To validate the performance of a fully trained C-NNP, a comprehensive testing set comprising 8000 structures was chosen evenly from FPMD and PIMD simulations performed using the generation 4 C-NNP.

The electronic structure, energy, and forces in these simulations are calculated using the hybrid revPBE0 functional with D3 dispersion corrections.[56–59] This functional combines two essential features to describe water correctly: on the one hand dispersion corrections are necessary to reproduce the density anomaly of water and the density difference between water and ice,[28] and, on the other hand, the fractional exact exchange provides an improved

description of the hydrogen bonding.[60] Hence, previous studies have shown that the Behler-Parrinello neural network potentials and C-NNP fitted on DFT calculations at the level of the revPBE0-D3 hybrid functional reproduce the structural, thermodynamic, and spectroscopic properties of liquid water in good agreement with experiments.[29,36,61]

To verify the transferability of our NEP potential, we compared the dynamics of a 12-molecule water cluster at the levels of NEP and DFT-based FPMD with the revPBE0-D3 functional and the same parameters used to generate the training data.[62] DFT-based FPMD simulations were carried out using the Quickstep approach implemented in the CP2K v8.2 package.[63] Valence Kohn–Sham orbitals are expanded on a double-ζ localized basis set[64] in real space, and core states are treated implicitly using Goedecker–Teter–Hutter pseudopotentials.[65] Plane waves up to a cutoff energy of 300 Ry are used to represent the electron density. The equations of motion were integrated with 0.5 fs timestep and the temperature was kept at 260 K by stochastic velocity rescaling.[66] with a time relaxation constant of 0.1 ps. The cluster was placed in a cubic simulation box (30 Å). With this FPMD setup, we computed the vibrational densities of states (vDOS) over a 3 ps trajectory from the Fourier transform of the velocity auto-correlation function. For the 36, 108 and 324 water molecule clusters that are run using our NEP potential, we used the same protocol as the simulation runs with the 12-water cluster but these clusters were placed in a cubic simulation box of 100 Å.

Thermodynamic Properties and Water Anomalies

To assess the performance of the NEP water potential we calculated with classical MD 15 thermodynamic quantities of liquid water, ice I_h, and ice II, listed in Table 2. The table reports also the size of the simulation systems. Most of the simulations have been performed on a 522-molecule with periodic boundary conditions, which is large enough to avoid finite-size artifacts on thermodynamic averages.[67] Hereafter we describe the simulation protocol used to compute these quantities. Well-equilibrated models of water models at the desired

Table 2: List of calculated thermodynamic properties.

Notation	Units	Description	Size (N_{H_2O})
$\rho_{liquid,298.15K,1bar}$	g/cm^3	Density of liquid water at 298.15 K, 1.01325 bar	522
$\rho_{liquid,400.15K,1bar}$	g/cm^3	Density of liquid water at 400.15 K, 1.01325 bar	522
$\rho_{ice\ Ih,250K,1bar}$	g/cm^3	Density of hexagonal ice at 250 K and 1 bar	384
$\rho_{ice\ II,123K,1bar}$	g/cm^3	Density of ice II at 123 K and 1 bar	324
$\rho_{liquid,373K,10Kbar}$	g/cm^3	Density of liquid water at 373 K, 10 kbar	522
$\rho_{liquid,373K,20Kbar}$	g/cm^3	Density of liquid water at 373, 20 kbar	522
$\kappa_{T,298K,1bar}$	1/Mbar	Isothermal compressibility of water at 298 K, 1 bar	522
$T_{max\ \rho}$	K	Temperature at maximum density (TMD)	522
$ln\mathscr{D}_{278K}$		Diffusion coefficient of liquid water at 278 K	522
$ln\mathscr{D}_{298K}$		Diffusion coefficient of liquid water at 298 K	522
$ln\mathscr{D}_{318K}$		Diffusion coefficient of liquid water at 318 K	522
γ_{300K}	mN/m	Surface tension at 300 K	522, 1536
T_{melt}	K	Melting temperature	10368
ΔH_{melt}	kcal/mol	Enthalpy of melting (l/s)	522 / 384
ΔH_{vap}	kcal/mol	Enthalpy of vaporization (l/g)	522 / 1

thermodynamic conditions were obtained from MD simulations in the isobaric-isothermal canonical ensemble (NPT) for at least 10 ns and up to 40 ns.[68] At deep supercooling, below 250 K, systems were equilibrated until the last 10 ns of simulations did not exhibit any drift of the potential energy or the density. For ice models, 1 ns equilibration runs were sufficient. To probe water anomalies, we equilibrated liquid D_2O at P =1 atm (1.013 bar) and at temperatures between 180 and 360 K at intervals of 10 K, starting from the highest temperature. Thermodynamic averages were computed from production runs of at least 10 ns, also conducted in the NPT ensemble. To determine the temperature of maximum density ($T_{max\ \rho}$) of water, we fitted a quartic polynomial to the average densities computed between 260 and 360 K and we defined $T_{max\ \rho}$ as the location of the analytical maximum of the polynomial fit.

From these runs, we computed the *isothermal compressibility* (κ_T) of liquid water from the volume fluctuations of the NPT MD production runs:[69,70]

$$\kappa_T = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_{N,T} \approx \frac{\langle V^2\rangle - \langle V\rangle^2}{k_B T\langle V\rangle}, \tag{8}$$

where k_B is the Boltzmann constant and the bracket notation indicates ensemble averages.

The molar heat capacity at constant pressure $C_p = (\partial H/\partial T)_P$ was also obtained from these simulations by fitting the enthalpy as a function of the temperature in the temperature range between 260 and 320 K with a 4th-order polynomial function and taking its first derivative.

Diffusion coefficients ($\mathscr{D}$) of liquid water at three different temperatures were determined using the Green-Kubo formula from 1 ns-long microcanonical (NVE) production runs:[71]

$$\mathscr{D} = \frac{1}{3}\int_0^\infty \langle \boldsymbol{v}(t)\cdot\boldsymbol{v}(0)\rangle dt. \tag{9}$$

During these runs, we sampled the velocities every 1 fs and computed the correlation function up to 1 ps.

To compute the surface tension of water at 300 K (γ_{300K}), two bulk water models containing 522 and 1536 molecules, equilibrated at 300 K at atmospheric pressure, were transformed into slabs by introducing a vacuum region of 80Å in the z direction. The slabs were further equilibrated in the canonical NVT ensemble at 300 K. The surface tension (γ_T) is calculated from the average of the pressure normal and parallel to the surface computed in a 10 ns NVT production run as:[72,73]

$$\gamma = \frac{L_z}{2}\left\langle P_{zz} - \frac{(P_{xx}+P_{yy})}{2}\right\rangle, \tag{10}$$

where L_z indicates the length of the simulation cell along the direction perpendicular to the surface. P_{xx} and P_{yy} are the diagonal components of the pressure parallel to the surface and P_{zz} is the perpendicular component.

The melting temperature (T_{melt}) is calculated using the two-phase method. In this approach, a large system (10^4 water molecules) is prepared by joining an ice I_h model and a liquid water model. After a short equilibration at a low temperature that allows the water-ice interface to relax, the system is simulated in the NPT ensemble at different temperatures.

Monitoring the dynamics of the solid-liquid interface for a sufficiently long time allows one to identify the melting temperature: if $T < T_{melt}$ then the solid phase grows, otherwise the liquid phase prevails. The crystal growth/melting is monitored using a local structural order parameter, such as the Steinhardt local q_6.[74–76]

After determining T_{melt}, independent simulations of hexagonal ice and bulk water were performed to calculate ΔH_{melt}:

$$\Delta H_{melt} = H_l - H_s = \left(\hat{U}_l - \hat{U}_s\right) + \left(P_l\hat{V}_l - P_s\hat{V}_s\right), \tag{11}$$

where the hat notation in Eq. 11 indicates a molar basis for the respective quantities. Both solid (s) and liquid (l) phase runs were conducted with 200 ps NPT equilibration, followed by 1 ns NVT productions. Similarly, ΔH_{vap} can be calculated by running independent simulations of water in gas (g) and liquid (l) phases. ΔH_{vap} is defined as:

$$\Delta H_{vap} = H_g - H_l = \left(\hat{U}_g - \hat{U}_l\right) + RT \tag{12}$$

where R is the ideal gas constant.

Molecular Dynamics Simulations

Different properties have been calculated following different MD protocols. Structural, and dynamical (e.g. vibrational density of states, self-diffusion coefficient) properties of water were computed from 1 ns production runs in the canonical ensemble at constant-volume (NVT) using stochastic velocity rescaling[66] with a coupling constant $\tau = 0.05$ ps. These runs are preceded by a 200 ps equilibration in the NPT ensemble with $\tau_T = 0.05$ ps and $\tau_P = 0.5$ ps coupling constants for the thermostat and barostat.[68] The water model used for these runs consists of 522 H_2O molecules in a cubic box with periodic boundary conditions in the three dimensions. The equations of motion are integrated with a timestep of 0.5 fs, which guarantees energy conservation with numerical fluctuations of 10^{-4} eV/molecule and

no appreciable drift.

To assess the accuracy of water NEP for quantum simulations, the radial distribution function (RDF) of water at room temperature was also computed at the quantum mechanical level using path-integral MD accelerated with a generalized Langevin equation thermostat (PIGLET), to allow for convergence on the number of beads, such that only 8 beads were used for the simulation.[77] This simulation is performed on a 64-water molecule model in a cubic box ($l = 12.45$ Å, $\rho = 0.99$ g/cm^3) for 100 ps with the same timestep as the classical simulations. We use the i-PI driver[78] with LAMMPS.[79]

To compute the density, isothermal compressibility, and heat capacity of water as a function of temperature, from 180 to 360 K, we equilibrate the 522-molecule model for 10 ns in the NPT ensemble, and we collect data for at least 10 ns after equilibration. For these simulations, we use D_2O as the larger mass of deuterium enables energy conservation with a 1 fs timestep. The temperature and pressure coupling times used to enforce NPT conditions are $\tau_T = 1$ ps and $\tau_p = 10$ ps during both equilibration and production. For the other thermodynamic properties reported in Table 2, we adopted the same procedure described for the structural properties. The uncertainty on the thermodynamic properties is calculated as the standard error over five blocks of the production runs unless specified otherwise in the results section.[80]

The two-phase simulations to determine the melting temperature were performed on a system of 10,368 water molecules in a cell with dimensions $\sim 5.4 \times 4.7 \times 12.7$ nm^3. The ice/liquid interface is in the plane perpendicular to the z axis and approximately half of the molecules are in the proton-disordered ice I_h structure while the other half is in the liquid state. The system is run for at least 1 ns in the NPT ensemble with a timestep of 0.5 fs and thermostat/barostat relaxation times $\tau_T = 0.05$ ps and $\tau_p = 5$ ps. Crystal growth or melting is monitored using the local q_6 order parameter with the same parameters used in recent work on ice melting.[81] The NEP potential development and all MD simulations were performed using GPUMD-3.6.[35]

Runtime Profiling

MD simulations protocols to obtain well-converged thermodynamic properties of water are enabled by the efficient GPU implementation of the NEP potential.[35] Figure 1 shows the run-time performance of our NEP water model on parallel CPUs (AMD Ryzen Threadripper 3970X, 4.0 GHz) and on GPUs (GeForce RTX 2080Ti), compared to TIP4P/2005. Benchmarks were performed on a 32,400-molecule (97,200 atoms) water model in an orthorhombic box with size 22.5×7.0×6.6 nm^3. MD simulations are run with a 1 fs timestep at an initial temperature of 290 K. Computational speed is given in ns/day. In its LAMMPS CPU implementation, NEP is about 20 times slower than TIP4P/2005 but the parallel scaling remains linear up to 32 CPUs. The GPU implementation of NEP in GPUMD-3.6 exhibits excellent scaling up to four parallel GPUs. The runtime speed on one GPU is comparable to TIP4P/2005 on 8 CPUs, and on 4 GPUs it exceeds that of TIP4P/2005 on 32 CPUs.

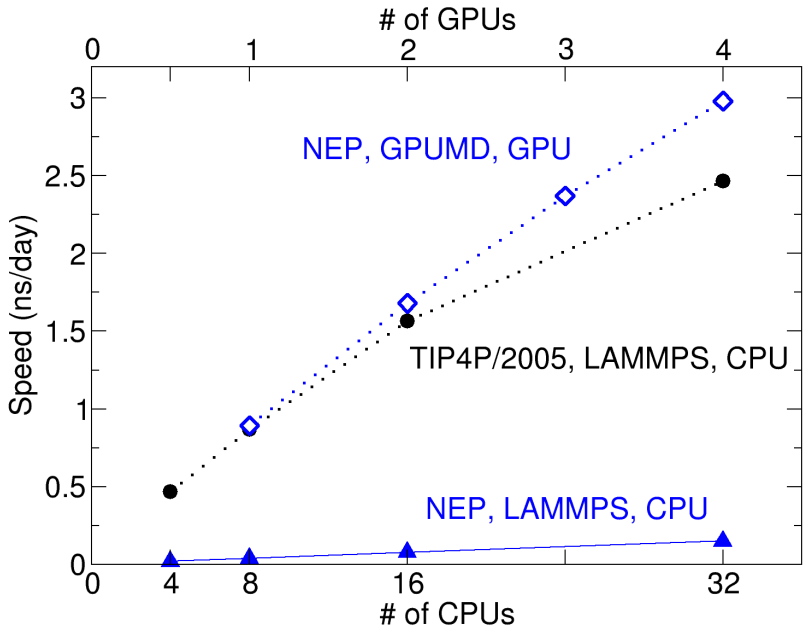

Figure 1: Computational efficiency of NEP run on parallel CPUs (LAMMPS implementation) and on parallel GPUs (GPUMD implementation), compared to TIP4P/2005 (LAMMPS implementation).

Results & Discussions

Training of the Neuroevolution Potential: comparison with Committee-Neural Network Potential

Here we compare the ability of the NEP potential and the C-NNP to predict the forces of the testing database. Figure 2 shows the force distribution of the testing set used to validate the NEP potential. We have chosen 500 configurations for each system from Newtonian and PIMD. The force components extend up to 34 eV/Å, with a markedly broader distribution in PIMD simulations. The training forces RMSE of NEP is 82 meV/Å, slightly larger than 58 meV/Å of C-NNP. These errors are overall similar or better than those reported in previous works in which NNPs for water were employed, which are in the 70–120 meV/Å range.[28,29,47]

Figure 3 shows that the error on the forces is fairly homogeneous across the systems included in the testing data set for both C-NNP and NEP, the former being systematically more accurate. However, different from the C-NNP whose testing performance is seemingly biased towards two phases of ices,[36] our NEP potential preserves a more evenly distributed force RMSE across the whole set of systems. Both machine learning potentials perform better on configurations sampled by classical mechanics and, at this level, they exhibit a slightly larger RMSE for the water slab, featuring liquid/vapor interfaces, which are more complex than bulk systems.[55] For configurations extracted from PIMD runs the error is consistent across the whole system. The larger error obtained for quantum configurations is likely due to the larger range of forces spanned by quantum dynamics. Overall, both potentials yield good transferability across a broad range of systems.[84]

The slightly superior training performance of C-NNP can be attributed to two factors: C-NNP is constructed using a high-dimensional neural network (HDNN) and contains multiple hidden layers, which capture finer details within the training set. Additionally, the force predictions from the C-NNP are achieved by ensemble-averaging the force predictions from each of the eight committee member NNPs, thus resulting in lower RMSE compared to the

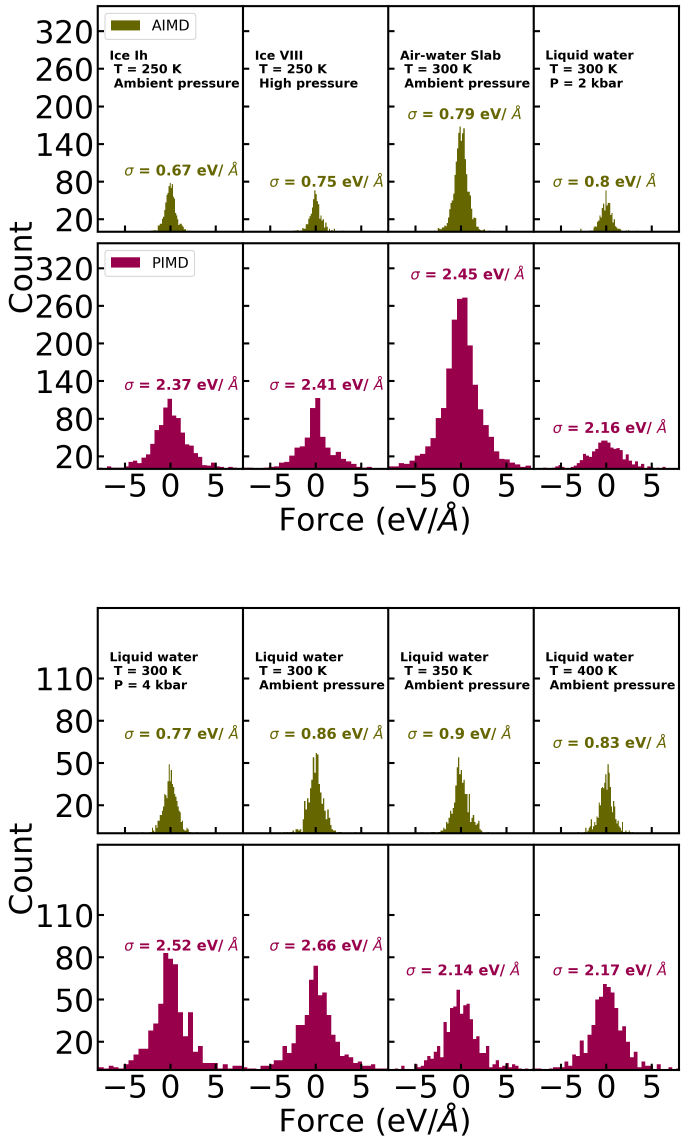

Figure 2: Force distributions of ice Ih, ice VIII, air-water slab, and liquid water at 2k bar (top graph), liquid water at 4k bar and at 300 K, 350 K, and 400 K (bottom graph). σ represents the standard deviation of the force distributions. Forces from classical dynamics (green) and PIMD (red) are shown in the top and bottom rows of each panel. It is worth noting that the exact target pressure used to sample the ice VIII configurations are not specified in ref.[36] Further details of the MD settings and the thermodynamic conditions used in sampling configurations for those force distributions can be found in ref.[36,82,83]

individual models.[36,85–87] Hence, the accuracy of C-NNP comes at a higher computational cost. Despite a slightly higher error on the forces, the simpler formulation of NEP makes it suitable to address problems where large-scale models and long simulation times are necessary

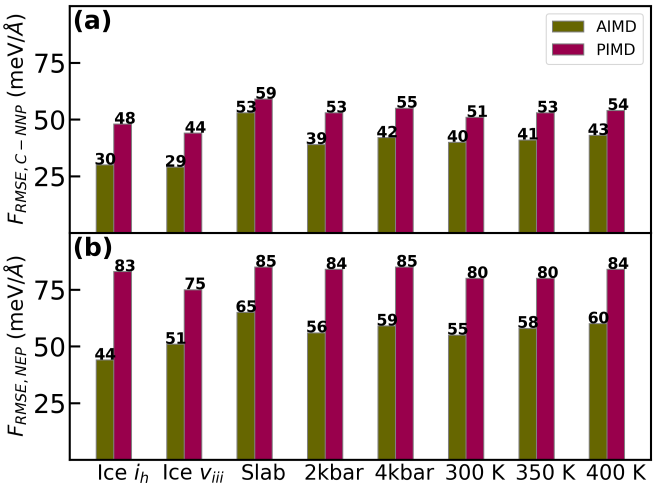

Figure 3: Testing performance of force from C-NNP (a) and NEP potentials (b). For each system, 500 structures are sampled from the FPMD and PIMD simulations using the generation-4 CNNP.[36]

to get well-converged results.

Structural Properties of Liquid Water

Since MLPs are designed to replace the expensive force calculations during DFT-based MD simulations, the quality assessment of the MLPs should go beyond error metrics on energies and forces.[43,90–93] Here, we assess the performance of NEP in reproducing the RDF of liquid water at 300 K ($g_{OO}(r), g_{OH}(r), g_{HH}(r)$). We compare the NEP RDFs from classical MD to those obtained using C-NNP, which was formerly validated by Schran et al. against FPMD.[36] Figure 4 shows that three RDFs are indistinguishable between the two MLPs. The classical $g_{OO}(r)$ is also very similar to the experimental one obtained by X-ray diffraction.[88] In turn, to compare the OH and the HH RDFs to experiments, it is necessary to consider nuclear quantum effects.[16,28,61,77,89,94–97] Figure 4 shows that PIMD with the NEP potential accurately reproduces the broadening of the first peak of both $g_{OH}(r)$ and $g_{HH}(r)$, as well as the variations in the structure of water due to nuclear quantum effects, leading to an excel-

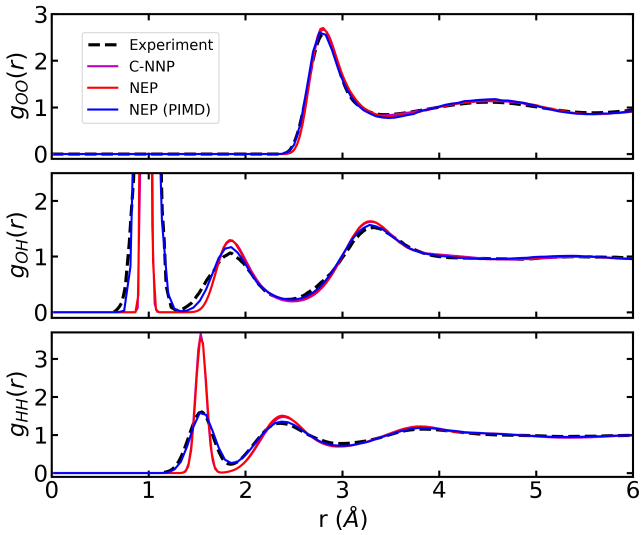

Figure 4: Radial distribution functions for liquid water at 300 K. Black-dotted RDFs represent experimental reference.[88,89] Blue RDFs are computed from path-integral simulations with the PIGLET algorithm.[77]

lent agreement with the RDFs inferred from X-ray diffraction. Overall, the NEP potential accurately reproduces the experimental structure of water at the quantum mechanical level and it is in good agreement with former MD results with either MLPs or FPMD.

Dynamics of Water Clusters

In addition to replicating the properties of known systems, proficient MLPs should predict the properties of systems not present in the training dataset, as long as they exhibit sufficient similarity to those included in the training data. Here we probe the transferability of our NEP potential, by testing the dynamics of a small water cluster against FPMD. Figure 5 (a) shows the vDOS of a 12-molecule water cluster at 260 K obtained from FPMD and an MD run with NEP. It can be seen that the positions and overall shapes of the main peaks are consistent between the two approaches. In particular, NEP can reproduce the merging of the two low-frequency bands (which can usually be distinguished in bulk water) into a single broad band between 0 and 1000 cm^{-1} that corresponds to hindered translations and

librations. The overall features of this band, as well as those of the high-frequency stretching band, including the free-OH peak at 4000 cm^{-1}, are well reproduced by NEP. The differences in the peak intensities could be due not only to differences in the potential energy surface but also to the different lengths of the simulations. Additionally, in the NEP simulation, the bending peak at $\sim$ 1550 cm^{-1} appears to split, and its main component shifts to slightly lower frequencies.

The good agreement between FPMD and NEP vDOS allows us to assess the effect of varying the water cluster size on the vibrational spectrum. Using the same MD parameters as for the 12-molecule water cluster, we simulated the hydrogen-vDOS with water-cluster systems that are three, nine, and twenty-seven times larger in size with NEP (Fig. 5(b)). The recognizable differences in the vDOS for the varying-sized water clusters can be attributed to finite-size effects. The most obvious change is that the surface-to-volume ratio decreases as the size of the cluster increases. This changes the relative intensity of the free-OH stretching peak at 4000 cm^{-1}. We notice that the shoulder at a higher frequency than the free-OH peak, which is probably an artifact of the NEP model for very small clusters, disappears in the spectrum of the 36-water cluster. Concurrently, the relative intensity of the hydrogen-bonded OH stretching band increases and the band becomes smoother. No significant changes occur in either the intensity or the position of the bending peak. The low-frequency broad band which includes molecular librations and hindered translations evolves toward resembling the vibrational DOS of bulk water. In the 36-water cluster, this band still exhibits separate features, which are completely erased at larger sizes.

Thermodynamic and Response Properties

While the quality of MLPs is often evaluated by estimating the errors on the energies and forces of a test set and testing structural and vibrational properties,[62] the low computational cost of NEP provides us with easy access to equilibrium thermodynamics and response functions that can be computed by MD. To attain statistically well-converged thermodynamic

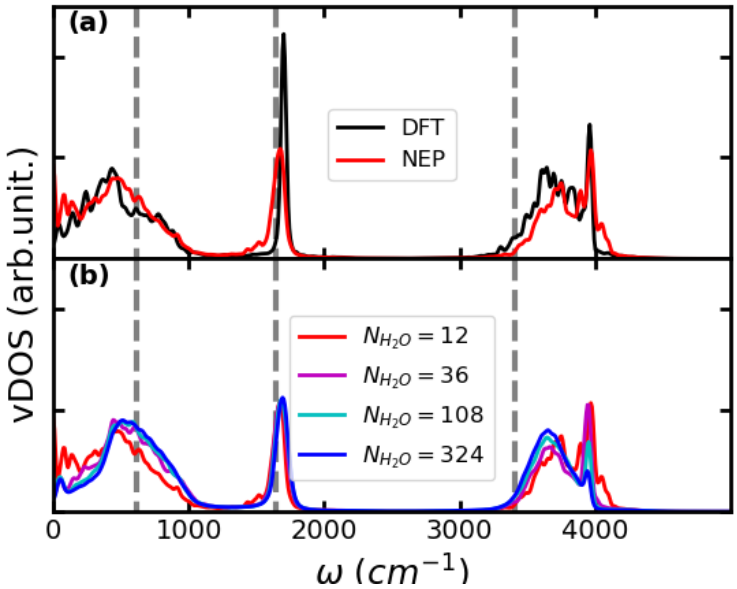

Figure 5: Hydrogen-vDOS for water clusters modeled with the NEP potential in comparison to vDOS modeled with DFT at the level of revPEB0-D3 (a) and water clusters of varying sizes modeled with the NEP potential (b). Three vertical dashed lines correspond to the OH stretching (3406 cm^{-1}), bending (1646 cm^{-1}), and librational modes (615 cm^{-1}).[98]

properties of liquids, long MD simulations of large systems are often required.[9,67,69,100] Even though these simulations can be feasible with accurate MLPs, such as C-NNP or the Gaussian approximation potential (GAP), they involve very high computational costs which may be a serious limitation for research groups who do not have access to high-performance computing resources. Owing to the optimized implementation on the modern GPUs,[35,101] NEP enables the calculations of thermodynamic and response properties at a computational cost comparable to classical force fields. Here, we apply our NEP potential to model 15 thermodynamic and response quantities, among those previously used to score rigid fixed-charge models.[9]

In Table 3, the results obtained from NEP simulations are compared to the corresponding experimental values, and to those obtained with TIP4P/2005, which is the overall best-scoring forcefield in the class of rigid fixed-charge models. Before comparing the results, however, it is important to note that the TIP4P/2005 parameters were fitted to reproduce

Table 3: Thermodynamic quantities and their uncertainty (σ) modeled with the NEP potential, compared to experiments and to the TIP4P/2005 empirical forcefield.[9,99] σ corresponds to the 63% confidence interval. Uncertainties on experimental data and TIP4P/2005 are not available. ⋆For TIP4P/2005 the errors on T_{melt} and $T_{max\ \rho}$ are referred to the corresponding values for H_2O, 273.15 and 277.1 K, respectively.

Properties	Experiment	NEP	σ	Error (%)	TIP4P/2005	Error (%)
$\rho_{liquid,298.15K,1bar}(g/cm^3)$	0.999	1.001	0.002	0.2	0.993	0.601
$\rho_{liquid,400.15K,1bar}(g/cm^3)$	0.9375	0.967	0.005	3.2	0.93	0.8
$\rho_{ice\ Ih,250K,1bar}(g/cm^3)$	0.92	0.915	0.002	0.5	0.921	0.11
$\rho_{ice\ II,123K,1bar}(g/cm^3)$	1.19	1.13	0.003	5.0	1.199	0.76
$\rho_{liquid,373K,10Kbar}(g/cm^3)$	1.201	1.165	0.004	3.0	1.204	0.25
$\rho_{liquid,373K,20Kbar}(g/cm^3)$	1.322	1.277	0.004	3.4	1.321	0.076
$T_{max\ \rho}\ (K)$	284.75 (D_2O)	292.2	0.4	2.6	278	0.325⋆
$T_{melt}\ (K)$	276.97 (D_2O)	290.5	0.5	4.0	252	9.01⋆
$\Delta H_{melt}(meV/H_2O)$	62.44	57.92	1.74	6.9	50.30	24.1
$\Delta H_{vap}(meV/H_2O)$	456.2	517.7	1.92	13.5	519.9	14.0
$\gamma_{300K}(mN/m)$	71.73	53.4	1.6	25.5	69.3	3.38
$\kappa_{T,298K,1bar}(Mbar^{-1})$	45.3	28.6	0.4	36.8	41	9.5
$ln\mathscr{D}_{278K}$	-11.24	-11.04	0.16	1.8	-11.27	0.267
$ln\mathscr{D}_{298K}$	-10.68	-10.66	0.13	0.2	-10.79	1.03
$ln\mathscr{D}_{318K}$	-10.24	-10.28	0.085	0.4	-10.39	1.46

a set of thermodynamic properties of water and ice: the temperature of maximum density ($T_{max\ \rho}$) of water the enthalpy of vaporization, the densities of water at ambient conditions, of ice II and ice V, and the range of temperature stability of ice III at a pressure of 3 kbar.[99] However, the TIP4P family of models has very few parameters, thus some properties cannot be fitted concurrently: for example a good $T_{max\ \rho}$ comes at the expense of the ice I_h melting temperature (T_{melt}). The NEP scheme offers more parameters and flexibility, but it is not fitted to experiments. Hence, the origin of the deviations from measured thermodynamic properties lies, in the underlying DFT functional, in the chosen hyperparameters, and in the fitting procedure. Additionally, some of the properties are affected by nuclear quantum effects, which are not accounted for in our MD simulations, whereas, since TIP4P/2005 is fitted to experiments, it may be considered an effective quantum model.

Table 3 shows that the equilibrium densities of water and ice I_h obtained by NEP are within 2% of the experimental references at temperatures below 300 K and within $\sim$ 3.5%

at $T = 373$ K and $p = 2$ kbar. Our model reproduces very accurately the density difference between ice I_h and water. This is an important property of the underlying hybrid van der Waals corrected DFT functional to which NEP was trained.[29,36] However, we note that the equilibrium density of water at room temperature of the NEP model is higher and closer to the experimental density than that reported in Ref.,[29] possibly because we train the NEP model on a richer dataset. The lower accuracy of NEP at high temperatures and high pressures as well as for ice II may stem from the fact that no data at these conditions are included in the training set. In particular, the error on the equilibrium density of ice II, a system that does not appear in the training set, is within 5%, which is usually considered acceptable for DFT. This is further evidence of the transferability of neural network potentials across different ice phases.[102]

Rescaling the molecular masses, we can compare the equilibrium density at 298 K to that of D_2O where we get 1.113 g/cm^3 with NEP, which is within 0.7% of the experimental value of 1.105 g/cm^3. Since we estimate the melting temperature from classical MD simulations, for properties that are severely affected by nuclear quantum effects, it is more compelling to compare our results to D_2O data. This is the case, for example, of the temperature of maximum density and the melting temperature. The $T_{max\rho}$ of heavy water is well reproduced by the NEP model, yet not as close as TIP4P/2005 reproduces the $T_{max\rho}$ of H_2O as it was fitted to this property. The NEP T_{melt} is within 4% of that of liquid D_2O: this is a major improvement over the large systematic underestimate of TIP4P models. At the same time, the enthalpy of melting and vaporization are modeled with good accuracy. The overall performance of the NEP model on the thermodynamic properties of water and ice suggests that it is suitable for investigating phase transitions at the molecular level, including ice melting and crystallization, ice nucleation, and the condensation of droplets from the vapor phase. However, the surface energy of the liquid/vapor interface may raise some concerns, as it is underestimated by about 25%. This may be due to the lack of explicit long-range electrostatics in the NEP model.[103] Additionally, the NEP model fails to reproduce accu-

rately the isothermal compressibility of water, which is underestimated by more than 30%. This is not entirely unexpected, given the reduced accuracy of our model at high pressures. There may be several different reasons for this shortcoming: First of all, acceptable errors in the thermodynamic properties turn into large errors in response functions. Secondarily, the training dataset may not contain enough high-pressure data. Another reason, as in the case of the surface tension, could be the absence of long-range tails in the NEP model. Finally, it may be a shortcoming of the underlying density functional framework. As for the dynamical response of water, the NEP model reproduces very well diffusivity over the range of temperatures between 278 and 318 K.

Water Anomalies

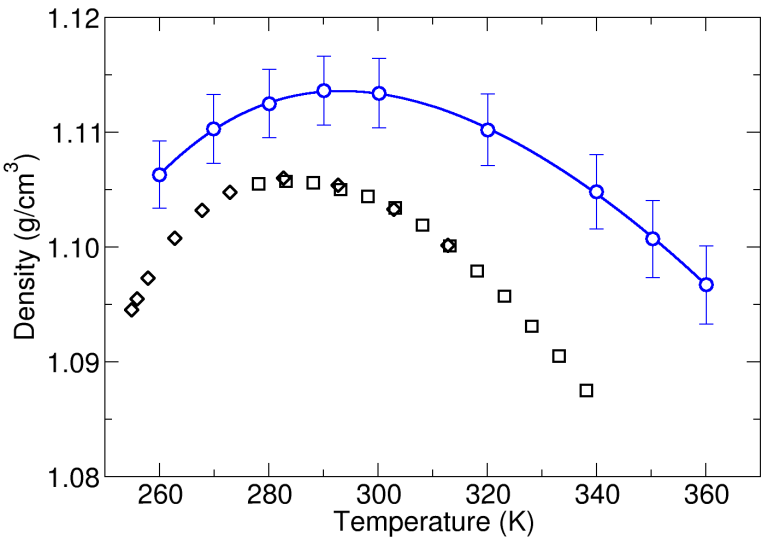

Figure 6: Density of heavy water as a function of temperature computed using NEP (blue dots) and compared to two sets of experimental data: Ref.[104] (black squares) and Ref.[105] (black diamonds). Error bars correspond to a 63% confidence interval. Experimental uncertainties are smaller than the symbols size. The blue line corresponds to a quartic polynomial fit of the NEP data, used to determine the temperature of maximum density.

Here we further assess the accuracy of the NEP model to reproduce water thermody-

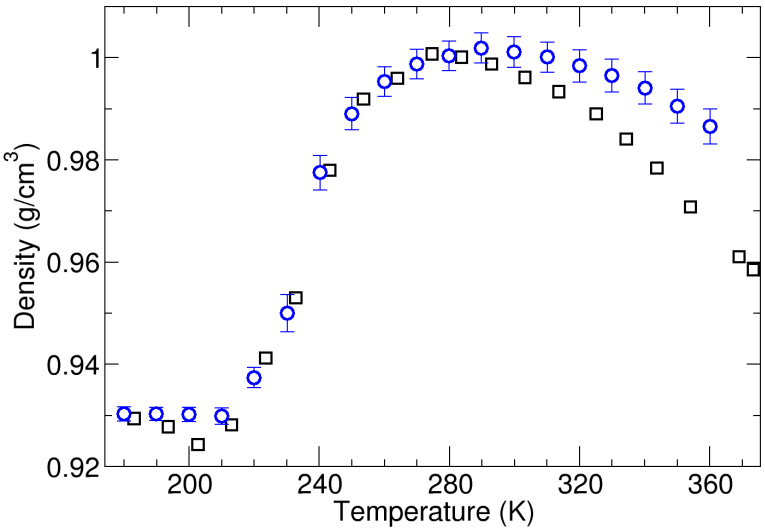

Figure 7: Density of water as a function of temperature computed using NEP (blue dots) and compared to experiments on confined water extending to the supercooled region Ref.[106] (black squares) Error bars correspond to a 63% confidence interval. Experimental uncertainties are smaller than the symbols size.

namics by probing the anomalous behavior of the density, the isothermal compressibility, and the molar constant-pressure heat capacity of D_2O. These properties were studied for the TIP4P/2005 model, which was proven to capture the anomalous behavior of water over a broad range of pressure and temperature.[67] We mostly compare to heavy water experiments because our MD simulations are classical and nuclear quantum effects are mitigated in D_2O. All the simulations are carried out in the NPT ensemble with $p = 1$ atm on a 522-molecule periodic water model. Figure 6 shows the density of NEP water between 250 and 360 K compared to available measurements.[104,105] The simulation data are systematically higher but they exhibit the same anomalous trend as the experiments, with a maximum density at 292 K. This is 7.5 K higher than the experimental reference, but we should remember that the melting point of the model is also overestimated, and the two quantities are interrelated.[99] Below 290 K, the theoretical and experimental data exhibit parallel trends, but the simulated densities deviate further at higher temperatures, indicating that the positive

thermal expansion coefficient of water is underestimated. Figure 7 extends the comparison between simulated and experimental density to the supercooled regime. At deep supercooling, we do not have experimental data for D_2O, hence we rescaled the molecular mass of our simulations to that of H_2O. Experiments revealed that upon deep supercooling water, density exhibits a minimum, occurring at 202 K.[106] This trend is well reproduced by the NEP model, for which the minimum density is observed at $\sim$ 210 K. However, the minimum density in the simulations is shallower than in experiments and the system does not expand significantly upon lowering the temperature. This may be due to difficulties in equilibrating water models at low temperatures as the diffusivity is extremely low. A similar trend was observed in simulations using the TIP4P/2005 forcefield.[67,69]

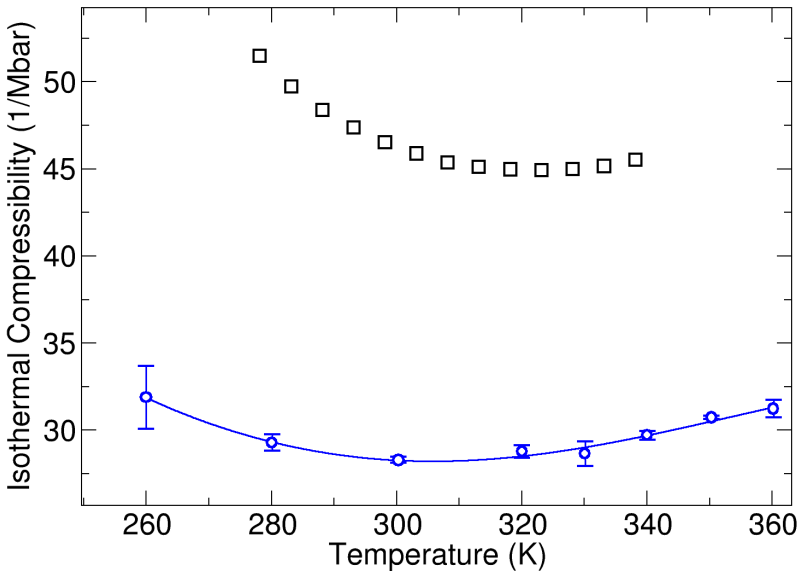

Figure 8: Isothermal compressibility of D_2O as a function of the temperature: comparison between MD simulations using NEP (blue dots) and experimental measurements (black squares).[107] Error bars are calculated as the semi-difference between the results of two 5 ns runs. Uncertainties of experimental measurements were not reported in Ref.[107]

In addition to the density anomalies, water displays an anomalous behavior in its isothermal compressibility, which in experiments has a minimum at 319.65 K for H_2O and at 322.5

K for D_2O. As discussed in the previous section, the NEP model underestimates κ_T, but Figure 8 shows that it reproduces qualitatively the anomalous trend. The simulated compressibility curve is shallower than experiments and its minimum occurs at a lower temperature (307±5 K). It is nevertheless remarkable that the NEP model captures the correct physical behavior of water compressibility, even though, in general, non-polarizable point-charge models, such as TIP4P and SPC/E, provide κ_T closer in value to experiments.

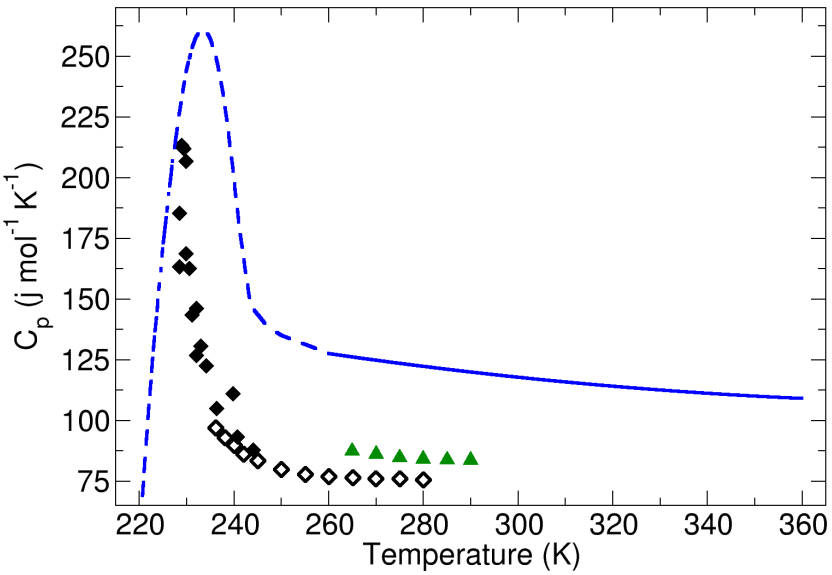

Figure 9: Isobaric heat capacity of D_2O computed from the derivative of a polynomial fit of the average enthalpy in the low- and high-temperature range (dashed and solid blue lines), compared to measurements D_2O (green triangles),[108] and supercooled H_2O (black diamonds).[109,110]

The isobaric heat capacity of water presents an anomalous non-monotonic behavior with both a maximum at deep supercooling $\sim$ 230 K, and a shallow minimum at 309 K for H_2O at ambient pressure. For D_2O the minimum shifts to $\sim$ 370 K, close to the boiling point, which is out of the temperature range of the simulations in this work. For this reason, we focus on the maximum. The existence of a low-temperature maximum of C_p is related to the second critical point of water and the liquid-liquid phase transition, and it has been

recently observed by calorimetric measurements.[109–111] Figure 9 shows C_p computed as the derivative of the average enthalpy from the NPT simulations at different temperatures. To obtain the derivative across the whole temperature range, we have divided the data into low-temperature (T<260 K) and high-temperature (T≥260 K) and fitted the two separate data sets to different third-order polynomials. At temperatures above 260 K, the NEP model largely overestimates C_p, which is more than 60% higher than the experimental references. However, as the temperature increases, the slope of C_p decreases compatibly with the existence of a minimum at $T > 360$ K. More importantly, our simulations capture the presence of a maximum of C_p in the supercooled region. The location of the maximum at 233 ± 5 K is compatible with the latest measurements (230 K), and the maximum value of C_p is in the same ballpark. However, the C_p peak in the experiment is very narrow, possibly indicating a singularity related to the phase behavior of supercooled water. If this is the case, finite-size effects may add up to the coarse temperature sampling of our simulations leading to an apparently broader C_p maximum.

Conclusions

In summary, we have assessed the quality of a neuroevolution potential trained on a database of water and ice models treated at the level of hybrid DFT with van der Waals dispersion corrections. The performance of the NEP model is comparable to that of state-of-the-art committee neural network potentials both in terms of the ability to fit the forces of a diverse training/testing dataset and to reproduce the structure of water at both classical and quantum levels. The efficient implementation of NEP on GPUs allows us to compute and assess several thermodynamic properties of water and ice systems, which turn out to be reasonably close to the corresponding experimental measurements. In particular, the NEP model gives a reasonably good estimate of the melting point of ice I_h at ambient pressure, and reproduces very well the density of both water and ice near the phase transition.

Finally, by examining the water anomalies at ambient pressure, we obtain density trends in remarkably good agreement with experiments, while properties related to the thermodynamic fluctuations display the correct physical trends, but can be off in absolute value. We cannot fully ascertain whether the observed discrepancies with experiments depend on the chosen DFT flavor or on how the NEP model was constructed and parameterized. A possible shortcoming of our NEP model is that it is local in nature, with an 8 Å cutoff for two-body and 6 Å for three-body interactions, thus lacking the contribution of long-range electrostatics. As opposed to the molecular structure and the vibrational response, which are controlled by hydrogen bonding, volume fluctuations, compressibility, and properties that depend critically on the contribution of the tail of the potential to the Virial, are considerably affected by neglecting long-range interactions.

Since our NEP potential reproduces accurately the thermodynamics of water and ice, and its computational efficiency on GPUs is comparable to that of simple all-atom potentials, we envisage its use to address problems that require long simulations of large systems, from thousands to several tens of thousand molecules, such as homogeneous and heterogeneous ice nucleation at mild supercooling and in confined geometries, and crystal growth. Additionally, this model can serve as a starting point to study more complex solutions or mineral/water interfaces, as more elements can be added to the fit through transfer learning.

Data Availability

The training data set, GPUMD input files of MD simulations, i-PI input and parameter files for the PIMD simulations, as well as the post-processing scripts that support our study are publicly accessible at https://github.com/nanotheorygroup/water_ice _nep.

Acknowledgments

We are grateful to Tianshu Li, Shunda Chen, Wanyu Zhao, and Xiaochen Jin for useful discussions. This work is supported by the National Science Foundation under Grant No. 2053235.

References

(1) Shukla, A.; Buddhi, D.; Sawhney, R. Solar water heaters with phase change material thermal energy storage medium: A review. *Renew. Sust. Energ. Rev.* **2009**, *13*, 2119–2125.

(2) Bagchi, B. *Water in biological and chemical processes: from structure and dynamics to function*; Cambridge University Press: Cambridge, 2013.

(3) Ball, P. *Life's matrix: a biography of water*; Univ of California Press: California, 2001.

(4) Cottin, H.; Kotler, J. M.; Bartik, K.; Cleaves, H. J.; Cockell, C. S.; De Vera, J.-P. P.; Ehrenfreund, P.; Leuko, S.; Ten Kate, I. L.; Martins, Z., et al. Astrobiology and the Possibility of Life on Earth and Elsewhere.... *Space Sci. Rev.* **2017**, *209*, 1–42.

(5) Isson, T. T.; Planavsky, N. J.; Coogan, L.; Stewart, E.; Ague, J.; Bolton, E.; Zhang, S.; McKenzie, N.; Kump, L. Evolution of the global carbon cycle and climate regulation on earth. *Global Biogeochem Cycles* **2020**, *34*, e2018GB006061.

(6) Finney, J. L. The water molecule and its interactions: the interaction between theory, modelling, and experiment. *J. Mol. Liq.* **2001**, *90*, 303–312.

(7) Finney, J. L. Bernal and the structure of water. J Phys Conf Ser. 2007; p 004.

(8) Shi, R.; Tanaka, H. The anomalies and criticality of liquid water. *Proc. Natl. Acad. Sci. U.S.A.* **2020**, *117*, 26591–26599.

(9) Vega, C.; Abascal, J. L. Simulating water with rigid non-polarizable models: a general perspective. *Phys. Chem. Chem. Phys.* **2011**, *13*, 19663–19688.

(10) Molinero, V.; Moore, E. B. Water modeled as an intermediate element between carbon and silicon. *J. Phys. Chem. B* **2009**, *113*, 4008–4016.

(11) Wang, L.-P.; Head-Gordon, T.; Ponder, J. W.; Ren, P.; Chodera, J. D.; Eastman, P. K.; Martinez, T. J.; Pande, V. S. Systematic Improvement of a Classical Molecular Model of Water. *J. Phys. Chem. B* **2013**, *117*, 9956–9972.

(12) Bukowski, R.; Szalewicz, K.; Groenenboom, G. C.; van der Avoird, A. Predictions of the Properties of Water from First Principles. *Science* **2007**, *315*, 1249–1252.

(13) Babin, V.; Leforestier, C.; Paesani, F. Development of a "First Principles" Water Potential with Flexible Monomers: Dimer Potential Energy Surface, VRT Spectrum, and Second Virial Coefficient. *J. Chem. Theory Comput.* **2013**, *9*, 5395–5403.

(14) Reddy, S. K.; Straight, S. C.; Bajaj, P.; Huy Pham, C.; 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*, 194504.

(15) Izvekov, S.; Parrinello, M.; Burnham, C. J.; Voth, G. A. Effective force fields for condensed phase systems from *ab initio* molecular dynamics simulation: A new method for force-matching. *J. Chem. Phys.* **2004**, *120*, 10896–10913.

(16) Fritsch, S.; Potestio, R.; Donadio, D.; Kremer, K. Nuclear Quantum Effects in Water: A Multiscale Study. *J. Chem. Theory Comput.* **2014**, *10*, 816–824.

(17) Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Phys. Rev. Lett.* **2007**, *98*, 146401.

(18) Bartók, A. P.; Payne, M. C.; Kondor, R.; Csányi, G. Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons. *Phys. Rev. Lett.* **2010**, *104*, 136403.

(19) Shapeev, A. V. Moment Tensor Potentials: A Class of Systematically Improvable Interatomic Potentials. *Multiscale Modeling & Simulation* **2016**, *14*, 1153–1173.

(20) Wood, M. A.; Thompson, A. P. Extending the accuracy of the SNAP interatomic potential form. *J. Chem. Phys.* **2018**, *148*, 241721.

(21) Zhang, L.; Han, J.; Wang, H. Deep Potential Molecular Dynamics: A Scalable Model with the Accuracy of Quantum Mechanics. *Phys. Rev. Lett.* **2018**, *120*, 143001.

(22) Zhang, L.; Lin, D.-Y.; Wang, H.; Car, R.; E, W. Active learning of uniformly accurate interatomic potentials for materials simulation. *Phys. Rev. Mater.* **2019**, *3*, 023804, Publisher: American Physical Society.

(23) Gubaev, K.; Podryabinkin, E. V.; Hart, G. L. W.; Shapeev, A. V. Accelerating high-throughput searches for new alloys with active learning of interatomic potentials. *Comput. Mater. Sci.* **2019**, *156*, 148–156.

(24) Batzner, S.; Musaelian, A.; Sun, L.; Geiger, M.; Mailoa, J. P.; Kornbluth, M.; Molinari, N.; Smidt, T. E.; Kozinsky, B. E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nat. Commun.* **2022**, *13*, 2453.

(25) Musaelian, A.; Batzner, S.; Johansson, A.; Sun, L.; Owen, C. J.; Kornbluth, M.; Kozinsky, B. Learning local equivariant representations for large-scale atomistic dynamics. *Nat. Commun.* **2023**, *14*, 579.

(26) Khaliullin, R. Z.; Eshet, H.; Kühne, T. D.; Behler, J.; Parrinello, M. Nucleation mechanism for the direct graphite-to-diamond phase transition. *Nat. Mater.* **2011**, *10*, 693–697.

(27) Deringer, V. L.; Bernstein, N.; Csányi, G.; Ben Mahmoud, C.; Ceriotti, M.; Wilson, M.; Drabold, D. A.; Elliott, S. R. Origins of structural and electronic transitions in disordered silicon. *Nature* **2021**, *589*, 59–64.

(28) Morawietz, T.; Singraber, A.; Dellago, C.; Behler, J. How van der Waals interactions determine the unique properties of water. *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113*, 8368–8373.

(29) Cheng, B.; Engel, E. A.; Behler, J.; Dellago, C.; Ceriotti, M. Ab initio thermodynamics of liquid and solid water. *Proc. Natl. Acad. Sci. U.S.A.* **2019**, *116*, 1110–1115.

(30) Zhang, L.; Wang, H.; Car, R.; E, W. Phase Diagram of a Deep Potential Water Model. *Phys. Rev. Lett.* **2021**, *126*, 236001.

(31) Piaggi, P. M.; Weis, J.; Panagiotopoulos, A. Z.; Debenedetti, P. G.; Car, R. Homogeneous ice nucleation in an ab initio machine-learning model of water. *Proc. Natl. Acad. Sci. U.S.A.* **2022**, *119*, e2207294119.

(32) Yang, J.; Dettori, R.; Nunes, J. P. F.; List, N. H.; Biasin, E.; Centurion, M.; Chen, Z.; Cordones, A. A.; Deponte, D. P.; Heinz, T. F., et al. Direct observation of ultrafast hydrogen bond strengthening in liquid water. *Nature* **2021**, *596*, 531–535.

(33) Zuo, Y.; Chen, C.; Li, X.; Deng, Z.; Chen, Y.; Behler, J.; Csányi, G.; Shapeev, A. V.; Thompson, A. P.; Wood, M. A.; Ong, S. P. Performance and Cost Assessment of Machine Learning Interatomic Potentials. *J. Phys. Chem. A* **2020**, *124*, 731–745.

(34) Fan, Z.; Zeng, Z.; Zhang, C.; Wang, Y.; Song, K.; Dong, H.; Chen, Y.; Ala-Nissila, T. Neuroevolution machine learning potentials: Combining high accuracy and low cost in atomistic simulations and application to heat transport. *Phys. Rev. B* **2021**, *104*, 104309.

(35) Fan, Z.; Wang, Y.; Ying, P.; Song, K.; Wang, J.; Wang, Y.; Zeng, Z.; Xu, K.; Lindgren, E.; Rahm, J. M., et al. GPUMD: A package for constructing accurate machine-learned potentials and performing highly efficient atomistic simulations. *J. Chem. Phys.* **2022**, *157*.

(36) Schran, C.; Brezina, K.; Marsalek, O. Committee neural network potentials control generalization errors and enable active learning. *J. Chem. Phys.* **2020**, *153*, 104105.

(37) Fan, Z. Improving the accuracy of the neuroevolution machine learning potential for multi-component systems. *J. Phys. Condens. Matter* **2022**, *34*, 125902.

(38) Dong, H.; Cao, C.; Ying, P.; Fan, Z.; Qian, P.; Su, Y. Anisotropic and high thermal conductivity in monolayer quasi-hexagonal fullerene: A comparative study against bulk phase fullerene. *Int. J. Heat Mass Transf.* **2023**, *206*, 123943.

(39) Wang, Y.; Fan, Z.; Qian, P.; Caro, M. A.; Ala-Nissila, T. Quantum-corrected thickness-dependent thermal conductivity in amorphous silicon predicted by machine learning molecular dynamics simulations. *Phys. Rev. B* **2023**, *107*, 054303.

(40) Rivlin, T. J. *Chebyshev polynomials*; Courier Dover Publications, 2020.

(41) Dodson, B. W. Development of a many-body Tersoff-type potential for silicon. *Phys. Rev. B* **1987**, *35*, 2795.

(42) Drautz, R. Atomic cluster expansion for accurate and transferable interatomic potentials. *Phys. Rev. B* **2019**, *99*, 014104.

(43) Kovács, D. P.; Oord, C. v. d.; Kucera, J.; Allen, A. E.; Cole, D. J.; Ortner, C.; Csányi, G. Linear atomic cluster expansion force fields for organic molecules: beyond rmse. *J. Chem. Theory Comput.* **2021**, *17*, 7696–7711.

(44) Dusson, G.; Bachmayr, M.; Csányi, G.; Drautz, R.; Etter, S.; van der Oord, C.; Ortner, C. Atomic cluster expansion: Completeness, efficiency and stability. *J. Comput. Phys.* **2022**, *454*, 110946.

(45) Batzner, S.; Musaelian, A.; Sun, L.; Geiger, M.; Mailoa, J. P.; Kornbluth, M.; Molinari, N.; Smidt, T. E.; Kozinsky, B. E (3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nat. Commun.* **2022**, *13*, 1–11.

(46) Svozil, D.; Kvasnicka, V.; Pospichal, J. Introduction to multi-layer feed-forward neural networks. *Chemometr Intell Lab Syst* **1997**, *39*, 43–62.

(47) Imbalzano, G.; Anelli, A.; Giofré, D.; Klees, S.; Behler, J.; Ceriotti, M. Automatic selection of atomic fingerprints and reference configurations for machine-learning potentials. *J. Chem. Phys.* **2018**, *148*, 241730.

(48) Xie, T.; Grossman, J. C. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Phys. Rev. Lett.* **2018**, *120*, 145301.

(49) Schaul, T.; Glasmachers, T.; Schmidhuber, J. High dimensions and heavy tails for natural evolution strategies. Proceedings of the 13th annual conference on Genetic and evolutionary computation. 2011; pp 845–852.

(50) Hoerl, A. E.; Kennard, R. W. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* **1970**, *12*, 55–67.

(51) Friedman, J.; Hastie, T.; Tibshirani, R. Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* **2010**, *33*, 1.

(52) Iftimie, R.; Minary, P.; Tuckerman, M. E. Ab initio molecular dynamics: Concepts, recent developments, and future trends. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6654–6659.

(53) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Comput. Phys. Commun* **2005**, *167*, 103–128.

(54) Iannuzzi, M.; Laio, A.; Parrinello, M. Efficient exploration of reactive potential energy surfaces using Car-Parrinello molecular dynamics. *Phys. Rev. Lett.* **2003**, *90*, 238302.

(55) Baer, M. D.; Mundy, C. J.; McGrath, M. J.; Kuo, I.-F. W.; Siepmann, J. I.; Tobias, D. J. Re-examining the properties of the aqueous vapor–liquid interface using dispersion corrected density functional theory. *J. Chem. Phys.* **2011**, *135*, 124712.

(56) Perdew, J. P.; Ernzerhof, M. Rationale for mixing exact exchange with density functional approximations. *J. Chem. Phys.* **1996**, *105*.

(57) Zhang, Y.; Yang, W. Comment on "Generalized gradient approximation made simple". *Phys. Rev. Lett.* **1998**, *80*, 890.

(58) Adamo, C. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*.

(59) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.

(60) Zhang, C.; Donadio, D.; Gygi, F.; Galli, G. First Principles Simulations of the Infrared Spectrum of Liquid Water Using Hybrid Density Functionals. *J. Chem. Theory Comput.* **2011**, *7*, 1443–1449.

(61) Marsalek, O.; Markland, T. E. Quantum dynamics and spectroscopy of ab initio liquid water: The interplay of nuclear and electronic quantum effects. *J. Phys. Chem. Lett.* **2017**, *8*, 1545–1551.

(62) Schran, C.; Thiemann, F. L.; Rowe, P.; Müller, E. A.; Marsalek, O.; Michaelides, A. Machine learning potentials for complex aqueous systems made simple. *Proc. Natl. Acad. Sci. U.S.A.* **2021**, *118*, e2110077118.

(63) Kühne, T. D. et al. CP2K: An electronic structure and molecular dynamics software package - Quickstep: Efficient and accurate electronic structure calculations. *J. Chem. Phys.* **2020**, *152*, 194103.

(64) VandeVondele, J.; Hutter, J. Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases. *J. Chem. Phys.* **2007**, *127*, 114105.

(65) Goedecker, S.; Teter, M.; Hutter, J. Separable dual-space Gaussian pseudopotentials. *Phys. Rev. B* **1996**, *54*, 1703.

(66) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.

(67) González, M. A.; Valeriani, C.; Caupin, F.; Abascal, J. L. F. A comprehensive scenario of the thermodynamic anomalies of water using the TIP4P/2005 model. *J. Chem. Phys.* **2016**, *145*, 054505.

(68) Bernetti, M.; Bussi, G. Pressure control using stochastic cell rescaling. *J. Chem. Phys.* **2020**, *153*, 114107.

(69) Pi, H. L.; Aragones, J. L.; Vega, C.; Noya, E. G.; Abascal, J. L.; Gonzalez, M. A.; McBride, C. Anomalies in water as obtained from computer simulations of the TIP4P/2005 model: density maxima, and density, isothermal compressibility and heat capacity minima. *Molecular Physics* **2009**, *107*, 365–374.

(70) Allen, M. P.; Tildesley, D. J. In *Computer Simulation of Liquids*; Press, O. C., Ed.; 1987.

(71) Guevara-Carrion, G.; Vrabec, J.; Hasse, H. Prediction of self-diffusion coefficient and shear viscosity of water and its binary mixtures with methanol and ethanol by molecular simulation. *J. Chem. Phys.* **2011**, *134*, 074508.

(72) Alejandre, J.; Tildesley, D. J.; Chapela, G. A. Molecular dynamics simulation of the orthobaric densities and surface tension of water. *J. Chem. Phys.* **1995**, *102*, 4574–4583.

(73) Vega, C.; de Miguel, E. Surface tension of the most popular models of water by using the test-area simulation method. *J. Chem. Phys.* **2007**, *126*, 154707.

(74) Steinhardt, P. J.; Nelson, D. R.; Ronchetti, M. Bond-orientational order in liquids and glasses. *Phys. Rev. B* **1983**, *28*, 784–805.

(75) Moroni, D.; ten Wolde, P. R.; Bolhuis, P. G. Interplay between Structure and Size in a Critical Crystal Nucleus. *Phys. Rev. Lett.* **2005**, *94*, 235703.

(76) Li, T.; Donadio, D.; Galli, G. Nucleation of tetrahedral solids: A molecular dynamics study of supercooled liquid silicon. *J. Chem. Phys.* **2009**, *131*, 224519.

(77) Ceriotti, M.; Manolopoulos, D. E.; Parrinello, M. Accelerating the convergence of path integral dynamics with a generalized Langevin equation. *J. Chem. Phys.* **2011**, *134*, 084104.

(78) Kapil, V.; Rossi, M.; Ondrej, M.; Petraglia, R.; Litman, Y. T.; Spura, T.; Cheng, B.; Cuzzocrea, A.; Meißner, R. H.; Wilkins, D. M.; Helfrecht, B. A.; Juda, P.; Bienvenue, S. P.; et. al, i-PI 2.0: A universal force engine for advanced molecular simulations. *Comput Phys Commun* **2019**, *236*, 214–223.

(79) Thompson, A. P.; Aktulga, H. M.; Berger, R.; Bolintineanu, D. S.; Brown, W. M.; Crozier, P. S.; in 't Veld, P. J.; Kohlmeyer, A.; Moore, S. G.; Nguyen, T. D.; Shan, R.;

Stevens, M. J.; Tranchida, J.; Trott, C.; Plimpton, S. J. LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Comp. Phys. Comm.* **2022**, *271*, 108171.

(80) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd ed.; Computational Science Series; Academic Press: San Diego, 2002; Vol. 1.

(81) Berrens, M. L.; Bononi, F. C.; Donadio, D. Effect of sodium chloride adsorption on the surface premelting of ice. *Phys. Chem. Chem. Phys.* **2022**, *24*, 20932–20940.

(82) Vacha, R.; Marsalek, O.; Willard, A. P.; Bonthuis, D. J.; Netz, R. R.; Jungwirth, P. Charge transfer between water molecules as the possible origin of the observed charging at the surface of pure water. *J. Phys. Chem. Lett.* **2012**, *3*, 107–111.

(83) Schran, C.; Marx, D. Quantum nature of the hydrogen bond from ambient conditions down to ultra-low temperatures. *Phys. Chem. Chem. Phys.* **2019**, *21*, 24967–24975.

(84) Miksch, A. M.; Morawietz, T.; Kästner, J.; Urban, A.; Artrith, N. Strategies for the construction of machine-learning potentials for accurate and efficient atomic-scale simulations. *Mach. learn.: sci. technol.* **2021**, *2*, 031001.

(85) Hashem, S.; Schmeiser, B. Improving model accuracy using optimal linear combinations of trained neural networks. *IEEE trans. neural netw.* **1995**, *6*, 792–794.

(86) Breiman, L. Bagging predictors. *Mach Learn* **1996**, *24*, 123–140.

(87) Hashem, S. Optimal linear combinations of neural networks. *Neural networks* **1997**, *10*, 599–614.

(88) Skinner, L. B.; Benmore, C.; Neuefeind, J. C.; Parise, J. B. The structure of water around the compressibility minimum. *J. Chem. Phys.* **2014**, *141*, 214507.

(89) Chen, W.; Ambrosio, F.; Miceli, G.; Pasquarello, A. Ab initio electronic structure of liquid water. *Phys. Rev. Lett.* **2016**, *117*, 186401.

(90) Young, T. A.; Johnston-Wood, T.; Deringer, V. L.; Duarte, F. A transferable active-learning strategy for reactive molecular force fields. *Chem. Sci.* **2021**, *12*, 10944–10955.

(91) Kovács, D. P.; McCorkindale, W.; Lee, A. A. Quantitative interpretation explains machine learning models for chemical reaction prediction and uncovers bias. *Nat. Commun.* **2021**, *12*, 1–9.

(92) Fonseca, G.; Poltavsky, I.; Vassilev-Galindo, V.; Tkatchenko, A. Improving molecular force fields across configurational space by combining supervised and unsupervised machine learning. *J. Chem. Phys.* **2021**, *154*, 124102.

(93) Fu, X.; Wu, Z.; Wang, W.; Xie, T.; Keten, S.; Gomez-Bombarelli, R.; Jaakkola, T. S. Forces are not Enough: Benchmark and Critical Evaluation for Machine Learning Force Fields with Molecular Simulations. *TMLR* **2023**, *2835-8856*.

(94) Mantz, Y. A.; Chen, B.; Martyna, G. J. Structural correlations and motifs in liquid water at selected temperatures: Ab initio and empirical model predictions. *J. Phys. Chem. B* **2006**, *110*, 3540–3554.

(95) Morrone, J. A.; Car, R. Nuclear quantum effects in water. *Phys. Rev. Lett.* **2008**, *101*, 017801.

(96) Wang, L.-P.; Martinez, T. J.; Pande, V. S. Building force fields: An automatic, systematic, and reproducible approach. *J. Phys. Chem. Lett.* **2014**, *5*, 1885–1891.

(97) Ruiz Pestana, L.; Marsalek, O.; Markland, T. E.; Head-Gordon, T. The quest for accurate liquid water properties from first principles. *J. Phys. Chem. Lett.* **2018**, *9*, 5009–5016.

(98) Bertie, J. E.; Lan, Z. Infrared intensities of liquids XX: The intensity of the OH stretching band of liquid water revisited, and the best current values of the optical constants of H2O (l) at 25 C between 15,000 and 1 cm- 1. *Applied Spectroscopy* **1996**, *50*, 1047–1057.

(99) Abascal, J. L.; Vega, C. A general purpose model for the condensed phases of water: TIP4P/2005. *J. Chem. Phys.* **2005**, *123*, 234505.

(100) Malosso, C.; Zhang, L.; Car, R.; Baroni, S.; Tisi, D. Viscosity in water from first-principles and deep-neural-network simulations. *Npj Comput. Mater.* **2022**, *8*, 139.

(101) Fan, Z.; Pereira, L. F. C.; Wang, H.-Q.; Zheng, J.-C.; Donadio, D.; Harju, A. Force and heat current formulas for many-body potentials in molecular dynamics simulations with applications to thermal conductivity calculations. *Phys. Rev. B* **2015**, *92*, 094301.

(102) Monserrat, B.; Brandenburg, J. G.; Engel, E. A.; Cheng, B. Liquid water contains the building blocks of diverse ice phases. *Nat. Commun.* **2020**, *11*, 1–8.

(103) Gao, A.; Remsing, R. C. Self-consistent determination of long-range electrostatics in neural network potentials. *Nat. Commun.* **2022**, *13*, 1572.

(104) Lide, D. R. *CRC handbook of chemistry and physics*; CRC press, 2004; Vol. 85.

(105) Hare, D. E.; Sorensen, C. M. Densities of supercooled H2O and D2O in 25 μm glass capillaries. *J. Chem. Phys.* **1986**, *84*, 5085–5089.

(106) Mallamace, F.; Branca, C.; Broccio, M.; Corsaro, C.; Mou, C.-Y.; Chen, S.-H. The anomalous behavior of the density of water in the range 30 K $<$ T $<$ 373 K. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 18387–18391.

(107) Millero, F. J.; Lepple, F. K. Isothermal Compressibility of Deuterium Oxide at Various Temperatures. *J. Chem. Phys.* **1971**, *54*, 946–949.

(108) Angell, C. A.; Sichina, W. J.; Oguni, M. Heat capacity of water at extremes of supercooling and superheating. *J. Phys. Chem.* **1982**, *86*, 998–1002.

(109) Murphy, D. M.; Koop, T. Review of the vapour pressures of ice and supercooled water for atmospheric applications. *Q. J. R. Meteorol. Soc.* **2005**, *131*, 1539–1565.

(110) Pathak, H. et al. Enhancement and maximum in the isobaric specific-heat capacity measurements of deeply supercooled water using ultrafast calorimetry. *Proc. Natl. Acad. Sci. U.S.A.* **2021**, *118*, e2018379118.

(111) Angell, C. A. Supercooled Water. *Annu Rev Phys Chem* **1983**, *34*, 593–630.