Article

Polarizable Water Model with Ab Initio Neural Network Dynamic Charges and Spontaneous Charge Transfer

Published as part of Journal of Chemical Theory and Computation special issue "Developments of Theoretical and Computational Chemistry Methods in Asia".

Qiujiang Liang* and Jun Yang*

ACCESS | 📊 Metrics & More | 📄 Article Recommendations | SI Supporting Information

ABSTRACT: Simulating water accurately has been a challenge due to the complexity of describing polarization and intermolecular charge transfer. Quantum mechanical (QM) electronic structures provide an accurate description of polarization in response to local environments, which is nevertheless too expensive for large water systems. In this study, we have developed a polarizable water model integrating Charge Model 5 atomic charges at the level of the second-order Møller−Plesset perturbation theory, predicted by an accurate and transferable charge neural network (ChargeNN) model. The spontaneous intermolecular charge transfer has been explicitly accounted for, enabling a precise treatment of hydrogen bonds and out-of-plane polarization. Our ChargeNN water model successfully reproduces various properties of water in gas, liquid, and solid phases. For example, ChargeNN correctly captures the hydrogen-bond stretching peak and bending-libration combination band, which are absent in the spectra using fixed charges, highlighting the significance of accurate polarization and charge transfer. Finally, the molecular dynamical simulations using ChargeNN for liquid water and a large water droplet with a ~4.5 nm radius reveal that the strong interfacial electric fields are concurrently induced by the partial collapse of the hydrogen-bond network and surface-to-interior charge transfer. Our study paves the way for QM-polarizable force fields, aiming for large-scale molecular simulations with high accuracy.
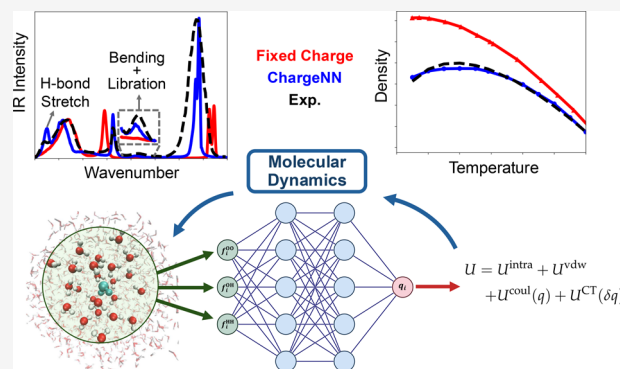
1. INTRODUCTION

As an active reagent, water plays a crucial role in regulating biomolecular functions, for instance, to stabilize their structures and facilitate dynamic interactions in proteins and nucleic acids.[1,2] Water also exhibits anomalous characteristics for nanomaterials, such as unique structural arrangements and fast diffusion rates within confined nanotubes compared to the bulk phase.[3] The polarization effect, which is the electron rearrangement in response to different local environments, has been shown to profoundly influence both static and dynamic properties of water.[4−7] For example, the induced polarization by electric field is considered to be the major source of the nonadditive cooperativity in water hydrogen bonding interactions.[8,9] Moreover, the considerable electron transfer between water molecules enhances the proton donor−acceptor orbital interactions and hence the covalency, directionality as well as relevant spectroscopic properties of water hydrogen bonds.[10] In recent years, microdroplets have been shown to accelerate chemical reactions by up to 10 million times compared to bulk solutions, widely attributed to the strong electric fields formed on their surfaces. One promising origin of the interfacial electric

fields is the large number of $H_2O^{\delta+}/H_2O^{\delta-}$ radical pairs on the microdroplet surface created by the intermolecular charge transfer.[11−14] A recent second-order Møller-Plesset perturbation theory (MP2) based molecular dynamics (MD) study indicates that substantial and inhomogeneous interfacial water charge transfer dramatically promotes the reactions of Criegee intermediates with water.[15]

Atomic partial charge is a qualitative representation of the environment-dependent electron distribution for understanding polarization effects and rationalizing chemical phenomena.[16−20] Nevertheless, partial charge is neither experimentally nor theoretically observable. In quantum chemistry, atomic charges vary significantly with different electron density partition methods, such as Mulliken,[21] Löwdin,[22] natural population

analysis (NPA),[23] quantum theory of atoms in molecules (QTAIM),[24] Hirshfeld,[25] etc. Notably, the Hirshfeld population analysis precisely describes the density redistribution during bonding.[25] Furthermore, the charge model 5 (CM5) corrects the underestimated Hirshfeld charges to achieve accurate dipole moments,[26] which has been shown to provide accurate electronic distribution and to be insensitive to the choice of basis sets and theories.[27] Calculating accurate electron densities to produce partial charges is computationally demanding using density functional theory (DFT) or higher level correlated wave function methods (e.g., MP2 and coupled cluster singles and doubles (CCSD)), making it cost-prohibitive for large molecular systems. In the past decades, efforts have been made to develop machine learning (ML) charges by mapping the atomistic local environment descriptors onto QM reference atomic charges.[28−50] However, these charge models were mainly tested on single molecules and small clusters, casting doubt on their accuracy when applied to larger molecular systems. A recent study reports a kernel ridge regression (KRR) model accurately predicts iterative Hirshfeld atomic charges[51] of water molecules in bulk liquid represented by QM/molecular mechanical (QM/MM) water clusters.[35] Remarkably, the hydrogen-bond stretch infrared (IR) peak is successfully retrieved using dynamic ML charges, whereas this weak peak disappears for fixed charges, marking the importance of polarization and intermolecular charge transfer in simulating liquid water.[35] However, this model has several limitations. First, the training and testing were conducted with an identical system (23 QM water molecules with 1977 MM water molecules), leaving the model's generalizability to larger QM regions uncertain. Second, the charge model was specifically tailored for water in bulk solution, which makes it inadequate for predicting charges for interfacial water molecules that experience more complex local environments. Finally, while ML charges were used to compute dipole moments and IR spectra using trajectories prepared with TIP4P/2005, they have not been integrated into a water model for simulating polarization in molecular dynamics.

Simulating water has been challenging for decades.[52] Since the introduction of the rigid nonpolarizable water models with fixed charges,[53−55] they have been substantially improved by refining parameters,[56−60] introducing intramolecular flexibility,[61−63] and integrating implicit polarizability.[64] Moreover, tremendous efforts have been made to recover the explicit polarization effects. Some polarizable models incorporate induced multipoles in response to the local electric fields using atomic polarizability tensors[28,29,65−77] or Drude oscillators.[78−82] In contrast, the charge-flow models enable geometry-dependent charge redistribution.[83−98] Although charge-flow polarizable models appear enticing due to their explicit incorporation of charge exchange between atoms, they suffer from several deficiencies, including high computational costs associated with iteratively updating charges, nonlinear computational scaling of the polarizability with system sizes, inadequate description of intermolecular charge transfer, and a lack of out-of-plane polarization.[99,100] Furthermore, despite extensive fine-tuning of the charge functional parameters to fit the DFT electrostatic potentials, the absence of spontaneous quantum mechanical information during the solution of dynamical charges may result in questionable electron density arrangements. Beyond the empirical force fields, energy potentials from first-principles for large water systems can be built by employing many-body expansion (MBE) for potentials of small water

fragments obtained through fitting with functionals or machine learning.[101−117] Nevertheless, constructing an accurate potential requires a tremendous amount of training data to cover the vast chemical space. The applicability of potentials is severely limited due to their system-specific nature that does not generalize well to unseen molecules. Furthermore, important electronic properties such as electron density and multipole moments are absent in the potentials. Notably, the QM-polarizable force fields that incorporate instantaneous QM-level polarizability enabled by machine-learning, such as FFLUX using machine learned QTAIM atomic multipoles, have been successfully applied to various chemical systems.[28,29,74−77,118]

In this study, we present a novel QM-polarizable water model that incorporates dynamical atomic charges and charge transfer exclusively trained on QM data acquired at the ab initio MP2 level of theory. The charges are produced using a deep neural network charge model called ChargeNN. A set of Interaction Classified Functions (ICFs) was employed as the ChargeNN features, explicitly capturing the types of atomic interactions to accurately depict local environments. Trained on diverse $(H_2O)_{25}$ structures sampled from MD trajectories, the ChargeNN is capable of accurately predicting charges for much larger water clusters, for example, $(H_2O)_{190}$, as compared to computed MP2 charges. The ChargeNN has also been integrated into a water model for dynamical polarization simulations. The intermolecular charge transfer has been explicitly accounted for using machine-learned pairwise water charge transfer, which effectively captures the out-of-plane polarization. We have implemented analytical energy gradients in the periodic boundary condition (PBC) within the shared-memory message passing interface (MPI) framework for efficient MD simulations. Our results validate the capability of ChargeNN-derived water model to reproduce a variety of water properties in excellent agreement with experimental data, for instance, the model generalizes well to a few thousand water molecules in unit cell for predicting the liquid water properties. Finally, the statistical analysis of liquid water and large water microdroplets using ChargeNN reveals that the partial breakage of hydrogen-bond networks, along with water charge transfer, concurrently promotes the layer electric fields at the air/water interface. The ChargeNN water model is thus of broad applicability in studying condensed ice, liquid water, nanophase, large microdroplets, and solvent interactions with materials and biomolecules.

2. METHODS

2.1. Charge Model 5. Charge Model 5 (CM5)[26] based on Hirshfeld population analysis (HPA)[25] was implemented for third-order many-body-expansion orbital-specific-virtual MP2 (MBE(3)-OSV-MP2) to investigate the water polarization and charge transfer in response to the variation in the local environment. The Hirshfeld atomic charges (q_i^{HPA}) are obtained by integrating the electron density within the region surrounding each atom,

$$q_i^{HPA} = Z_i - \int d\mathbf{r} \frac{\rho_i^0(\mathbf{r})}{\sum_j \rho_j^0(\mathbf{r})} \rho(\mathbf{r})$$

(1)

where Z_i is the nuclear charge. $\rho^0(\mathbf{r})$ and $\rho(\mathbf{r})$ denote the electron densities of the promolecule and real molecule at position $\mathbf{r}$, respectively. In contrast to the hard-boundary partitioning approaches like QTAIM,[119,120] where electron density at each grid point is attributed to a single atom, the
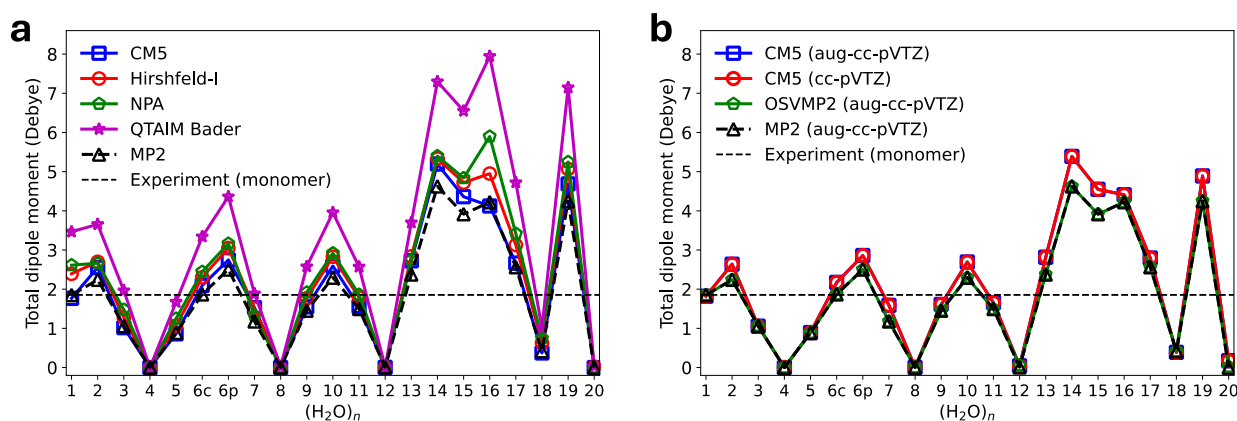
Figure 1. Water dipole moments derived from charges (a) Water dipole moments obtained from CM5, Hirshfeld-I,[51] NPA[121] and QTAIM charges, using electron density of MP2/aug-cc-pVTZ computed with Gaussian 16.[122] CM5 and NPA charges were obtained with Gaussian 16, while Multiwfn[123,124] was used to compute Hirshfeld-I and QTAIM charges. (b) Water dipole moments obtained from CM5 charges, computed with MBE(3)-OSV-MP2 using aug-cc-pVTZ and cc-pVTZ basis sets. Experimental structure of water monomer and structures from Cambridge water clusters[125] were used for this test. $(H_2O)_{6c}$ and $(H_2O)_{6p}$ refer to cage and prism conformers of water hexamers, respectively. Experimental dipole moment of the water monomer is from ref 126.

boundary-less Hirshfeld population analysis allows electron density to contribute to multiple atoms. It provides a clear partitioning of the electron density and accurately describes the density redistribution from promolecule to real molecule during bonding. Nevertheless, Hirshfeld charges are significantly underestimated due to the Hirshfeld weighting factor, which causes the atomic population to closely resemble that of the isolated atom.[51]

By introducing a unified set of parameters optimized by fitting to the experimental or DFT dipole moments of 614 molecules, the CM5 model maps the underestimated Hirshfeld charges onto a new set of charges, providing a more accurate description of the electrostatic potential:

$$q_i^{CM5} = q_i^{HPA} + \sum_{j \neq i} A_{ij} B_{ij}$$

$$B_{ij} = \exp[-\alpha(r_{ij} - R_{Z_i} - R_{Z_j})] \tag{2}$$

where A_{ij} and α are the CM5 model coefficients to be optimized. Constant R_Z denotes the atomic covalent radius. CM5 has been shown to deliver precise electronic charge distributions and be insensitive to different theoretical levels and basis sets.[26,27] As demonstrated in Figure 1a, CM5 charges outperform iterative Hirshfeld (Hirshfeld-I), NPA, and QTAIM charges in reproducing the dipole moments of water clusters containing 1−20 molecules. Furthermore, CM5 exhibits numerically consistent dipole moments between cc-pVTZ and aug-cc-pVTZ basis sets, as shown in Figure 1b. The CM5 charges in this work were obtained from the MBE(3)-OSV-MP2/cc-pVTZ relaxed density matrix.

2.2. Deep Neural Networks for Atomic Charges.
Tremendous efforts have been made to design descriptors of local atomic environments for building machine learning models of energy potentials and atomic properties.[112−115,127,128] For instance, the well-known radial (G_i^{rad}) and angular (G_i^{ang}) atom-centered symmetry functions (ACSFs)[127] write

$$G_i^{rad} = \sum_j e^{-\eta(r_{ij}-r_s)^2} s_c(r_{ij})$$

$$G_i^{ang} = 2^{1-\zeta} \sum_{j,k \neq i} (1 + \lambda\cos\theta_{ijk})^\zeta \cdot e^{-\eta(r_{ij}^2 + r_{ik}^2 + r_{jk}^2)}$$

$$\cdot s_c(r_{ij}) s_c(r_{ik}) s_c(r_{jk}) \tag{3}$$

where η, r_s, ζ and λ are tunable parameters. s_c denotes the cutoff function that enforces a smooth transition of atoms j entering and exiting the selection region for the atom i

$$s_c(r_{ij}) = \begin{cases} 0.5\left[\cos\left(\dfrac{\pi r_{ij}}{r_c}\right) + 1\right], & r_{ij} \leq r_c \\ 0, & r_{ij} > r_c \end{cases} \tag{4}$$

However, ACSF descriptors omit the dependence of the atomic interactions on element type, which can significantly impact training performance and data efficiency.

To incorporate the information on atomic interaction types, we propose an interaction classified function (ICF) to describe the interaction-specific ($\alpha\beta$) contributions to the charge of atom i of the atom type α, which satisfies translational, rotational and permutational invariance but requires no angular ACSF parametrization for angularly independent atomic charges:

$$f_i^{\alpha\beta} = \sum_{j \in \beta} e^{-r_{ij}^2/k_f} s_c(r_{ij}), \quad \text{where } i \in \alpha \tag{5}$$

In the above formula, a single parameter k_f controls the interaction decay with respect to the atomic distance, effectively alleviating the heavy process of the parameter tuning. Despite their simplicity, the radial ICFs have been found to sufficiently describe local interactions needed for reproducing accurate atomic charges with excellent data efficiency (see Section 3.1). To determine the atom selection cutoff (r_c) in eq 4, we tested the charge convergence for both bulk water and air/water interface, and found that 24 surrounding water molecules sufficiently form the local environment to produce charges comparable to those surrounded by 127 molecules, as demonstrated in Figure S1. The selected cutoff r_c was accordingly set to be 4.4 Å, which is

the minimum radius of multiple $(H_2O)_{25}$ droplets extracted from well-equilibrated liquid water boxes.

The ICF feature set for the charge of atom i in the molecule I reads

$$\mathbf{F}_{I_I} = \left\{ f_{i_I}^{\alpha_I \beta_I}, f_{i_I}^{\alpha_I \beta_J}, f_{j_I}^{\alpha_I \beta_J} \right\} \tag{6}$$

where the intra- ($f_{i_I}^{\alpha_I \beta_I}$) and intermolecular ICFs ($f_{i_I}^{\alpha_I \beta_J}$) are separately placed for detailing interaction types. In addition, the intermolecular ICFs of other atoms in the same molecule ($f_{j_I}^{\alpha_I \beta_J}$) are also appended to the feature set to account for the perturbation arising from the interactions for the neighboring atoms, which notably improves the prediction accuracy. Three decay parameters k_f were employed for each ICF in eq 6 to simulate varying interactions with the atomic distance, as illustrated in Figure S2. To better handle the distinct charge ranges and distributions, two separate neural networks were specifically trained for the charges of oxygen and hydrogen.

Although the electrostatic interaction, modeled by Coulomb potential with flexible atomic charges, implicitly accounts for environment-dependent charge flows, it is substantially underestimated in the absence of the intermolecular charge transfer (CT) contribution. For recovering the missing interactions, an extra neural network is used to predict intermolecular CT between water pairs (IJ) by supplying a feature set composed of intermolecular ICFs:

$$\mathbf{F}_{IJ} = \{ f^{\alpha_I \beta_J} \}$$

$$f^{\alpha_I \beta_J} = \sum_{\substack{i \in \alpha_I \\ j \in \beta_J}} e^{-r_{ij}^2 / k_f} s_c^{CT}(r_{ij}) \tag{7}$$

We examined the decay of the CT with the atomic distance, as the decay of the CT energy in eq 9 specifically depends on CT. Figure S3 shows that CT deteriorates to only $\sim 10^{-5}$ e at the hydrogen bond distance of 5.5 Å for a water dimer. Considering that CT would also be hindered by the water molecules filled between distant water pairs, 5.5 Å is sufficiently large for the CT cutoff.

The CM5 atomic charges of $(H_2O)_{25}$ were utilized to train the neural networks. Reference charges were computed using the low-scaling MBE(3)-OSV-MP2 method[129] with the cc-pVTZ basis set, as MP2 has been shown to accurately predict the energetic and electronic properties of water.[130−132] To benchmark the charge model, charges were prepared for 20000 $(H_2O)_{25}$, 200 $(H_2O)_{32}$, 50 $(H_2O)_{64}$, 10 $(H_2O)_{64}$ and 1 $(H_2O)_{190}$ cluster conformations that were randomly sampled from the equilibrated MD trajectories using SWM4-NDP with harmonic bonds and angles on OpenMM.[133,134] The NVT trajectories were simulated in the gas phase at 298.15 K, with a time step of 1 fs.

Four hidden layers are employed to effectively capture intricate relationships between features and atomic charges, which also mitigates overfitting issues. The Swish function[135] is adopted as the activation function:

$$A(x) = \frac{x}{1 + e^{-\beta x}} \tag{8}$$

The Swish function is differentiable and ensures smooth charge functions for new molecular configuration, which avoids discontinuity on the potential energy surface. Following multiple trials, we discovered that utilizing $\beta = 3$ allows $A(x)$

to retain good linearity around $x = 0$, enhancing the learning performance while circumventing erratic charge functions caused by overfitting. The mean squared error (MSE) was chosen as the loss function. Among the 20,000 configurations of $(H_2O)_{25}$, 8000 were randomly selected for training, with 50 used as the validation data set. The remaining data points were used for evaluating the prediction performance. After extensive testing, we applied the "early stopping" technique with 200 epochs to ensure satisfactory loss convergence and to prevent overfitting.

2.3. Water Model with ChargeNN Charges. The potential energy functional of the 3-site flexible water model incorporating ChargeNN charges and charge transfer contributions is expressed as a sum of the intramolecular energy (U^{intra}) and the intermolecular energy (U^{inter}),

$$U^{intra} = \sum_I \left[\sum_i \frac{k_b}{2} \left(r_{O_I H_{i_I}} - r_{OH}^0 \right)^2 + \frac{k_a}{2} (\theta_I - \theta^0)^2 \right]$$

$$U^{inter} = \sum_{i<j} \epsilon_{ij} \left[7 \left(\frac{\sigma_{ij}}{r_{ij}} \right)^9 - 5 \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right]$$

$$+ \sum_{i<j} k_e \frac{q_i q_j}{r_{ij}} - \sum_{I<J} k_c \frac{\delta q_{IJ}^2}{\delta q_{IJ}^2 + 3(\delta q^0)^2} \tag{9}$$

In the formula for intramolecular energy, k_b and k_a represent the coefficients of bond stretching and bending, respectively. $r_{O_I H_{i_I}}$ denotes the O−H bond length, while θ_I refers to the H−O−H angle of the molecule I, with r_{OH}^0 and θ^0 being the equilibrium values. For the intermolecular energy, ϵ_{ij} and σ_{ij} are standard 12−6 Lennard-Jones (LJ) parameters. k_e and k_c pertain to coulomb constant and charge-transfer constant, respectively. q_i denotes the charge of atom i. δq_{IJ} represents the charge transferred between molecules I and J, with the parameter δq^0 as the reference CT that yields the strongest CT force.

We observed that atomic charges exhibit significant sensitivity to stretching bonds and bending angles, necessitating the addition of flexibility in the intramolecular forces. Moreover, this flexibility has been reported to improve predicted water properties, and also permits recovering intramolecular vibrational spectra.[61,80] Consequently, we adopted harmonic bonds and angles for better description of the dynamical behavior of the water molecules. The equilibrium bond length r_{OH}^0 and angle θ^0 were tuned to ensure that the average values in the liquid phase align with experimental measurements at 298.15 K and 1 atm. The coefficients of bond length (k_b) and angle (k_a) were adjusted to produce accurate positions of the bond stretching and bending peaks in the IR spectrum.

The electrostatic energy is calculated using ChargeNN charges. To facilitate simulations under periodic boundary conditions (PBC), Ewald summation is employed, utilizing the minimum image convention to construct the local environment for charge generation. It is important to note that the total charge of a cluster or unit cell could be nonzero due to the accumulative error of atomic charges. To ensure a zero net charge of the whole system, all atomic charges must be subtracted by a small constant ($\Delta q = \sum_i q_i / N$, where N is the number of atoms) of approximately ±0.0001 e. Additionally, nonhomogeneous correction can be accomplished by considering the weights of atomic charges[34] or integrating an explicit
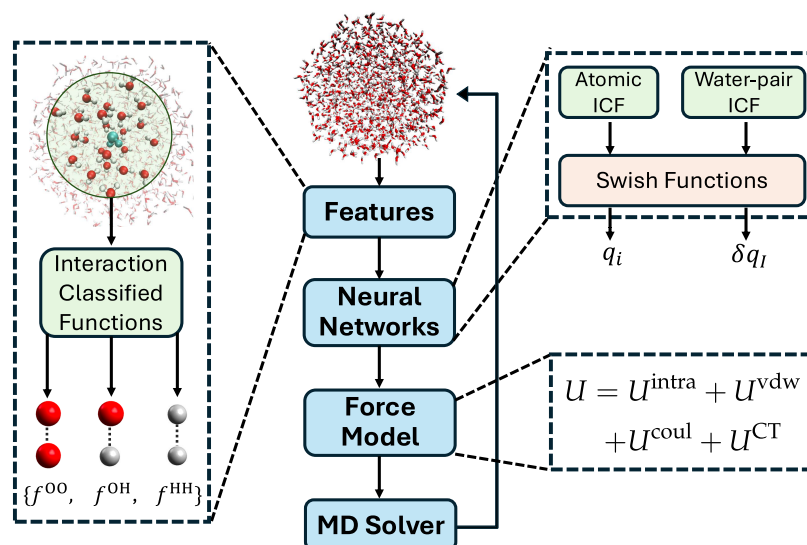
Figure 2. Implementation flowchart. Schematic illustration for the implementation of the polarizable water model using the charges produced by the ChargeNN model.
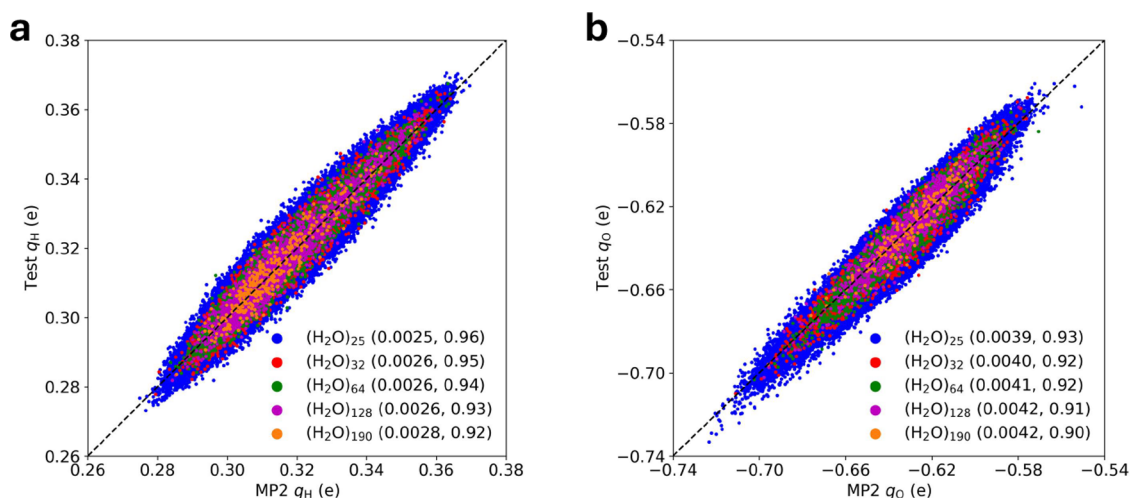
Figure 3. Benchmark results for the ChargeNN model. Comparisons of ChargeNN charges and MP2 charges for (a) hydrogen and (b) oxygen atoms of testing water clusters containing 25, 32, 64, 128, and 190 molecules. The charge models were trained with $(H_2O)_{25}$. The MAEs in elementary charge (e) and coefficients of determination (R^2) are given in the right brackets.

penalizing term to enforce electroneutrality into the loss function. There have been attempts to include charge transfer interactions with empirical charge transfer.[136−139] For our model that uses CT obtained quantum mechanically, the charge transfer term formulates a smooth ratio function of squared CT. The formula offers several merits, for instance, it is symmetric and the CT energy converges to zero in the absence of CT, preventing abrupt changes in potential energy around the CT cutoff. Additionally, the formula effectively captures the rapid intensification of CT interactions when initial charges are small, as well as the gradual attenuation due to the electron migration from equilibrium. Moreover, since the interwater CT occurs away from the nuclear sites, the CT term naturally accounts for out-of-plane polarization, resulting in more accurate water conformations.

As the inclusion of the CT term strengthens the short-range interactions, the traditional 12−6 LJ potential for van der Waals (VDW) forces yielding a steep repulsive wall constrains interwater distances to a compact range, which leads to an excessively high first peak in the oxygen−oxygen radial distribution function (RDF). Instead, we employ a 9−6 functional form with minor revisions to the original LJ formula, retaining the 12−6 LJ parameters ε and σ for compatibility with the existing force fields. In the process of the parameter tuning, the VDW equation was written as $U_{ij}^{\mathrm{vdw}} = A_{ij}/r_{ij}^9 - B_{ij}/r_{ij}^6$ for convenience. The intermolecular parameters k_c, k_d, A_{ij} and B_{ij} were optimized for the accurate heat of vaporization, density and radial distribution functions (RDF) for liquid water at the ambient conditions. All parameters can be found in Table S1.

2.4. Implementation. Figure 2 illustrates the implementation of the ChargeNN water model for molecular dynamics (MD) simulations. In each MD step, atomic coordinates adjusted by the MD solver are utilized to compute the ML features interaction-classified functions **F** and their analytical derivatives with respect to nuclear positions d**F**/d**R**. The ML features are then fed into the pretrained neural networks for the predictions of atomic charges and charge transfer. By employing autodifferentiation through back-propagation, the gradients of

the charges concerning the features $dq/d\mathbf{F}$ can be conveniently evaluated. Subsequently, the predicted atomic charges and charge transfer are used to determine the potential energy as described in eq 9, where total gradients are obtained using the chain rule:

$$\frac{dU}{d\mathbf{R}} = \frac{\partial \overline{U}}{\partial \mathbf{R}} + \frac{\partial U}{\partial q} \cdot \frac{dq}{d\mathbf{F}} \cdot \frac{d\mathbf{F}}{d\mathbf{R}} \qquad (10)$$

where $\partial \overline{U}/\partial \mathbf{R}$ denotes the derivative of all terms that are explicitly dependent on the nuclear coordinates $\mathbf{R}$. Finally, the energy and gradients are passed to the MD solver to update the positions.

The ChargeNN program was written in Python with double precision, integrated with Tensorflow for handling neural networks. The algorithm has been parallelized based on Message Passing Interface (MPI) standard of version 3. The shared memory windows built in MPI-3's remote memory access module enable significantly lower-latency data communications compared to traditional point-to-point communications.

3. RESULTS AND DISCUSSION

3.1. Performance of the Charge Model. The ChargeNN model demonstrates exceptional data efficiency, requiring only 8000 training $(H_2O)_{25}$ clusters to achieve converged mean absolute errors (MAEs) for the remaining testing clusters, as shown in Figure S4. According to Figure 3, when trained on 8000 $(H_2O)_{25}$, the charge model yields testing MAEs of 0.0025 e for hydrogen and 0.0039 e for oxygen within the same system. These values are markedly lower than the respective MAEs of 0.0036 and 0.0045 e for $(H_2O)_{23}$ predicted by a recent KRR water charge model trained with the same number of data points.[35] In addition, our model shows exceptionally high prediction reliability with the coefficients of determination (R^2) of 0.96 for hydrogen and 0.93 for oxygen, outperforming the 0.88 reported for the KRR model.[35] Notably, Figure 3 shows that the ChargeNN model considerably improves the transferability, only with a slow deterioration in accuracy as the water cluster size increases. For example, the models trained with $(H_2O)_{25}$ make accurate prediction of the charges for much larger $(H_2O)_{190}$, yielding MAEs of 0.0028 e for hydrogen and 0.0042 e for oxygen, both with R^2 values above 0.9. Figure S5 shows that predicting molecular charges by summing the ChargeNN atomic charges within each molecule is less satisfactory, with $R^2 > 0.92$ and marginally larger MAEs of ~0.007 e due to the accumulative errors. This error can be improved in future by imposing constraints on the loss function to penalize the predicted molecular charges. For assessing ChargeNN's performance in predicting charge transfer, we prepared charge transfer of 40,000 water pairs extracted from the MD trajectory of $(H_2O)_{25}$. The ChargeNN trained on 20,000 water pairs achieves an MAE of only 0.001 e and R^2 of 1.00 for the remaining testing pairs.

3.2. Performance of the Water Model. To assess the performance of the ChargeNN water model, we calculated various properties of water in gas, liquid, and solid states. Computational details are provided in the SI. We compare our results with those of SPC/FW,[61] which is the fixed charge counterpart of the 3-site ChargeNN model, and those with a widely used polarizable model SWM4-NDP.[79] As shown in Table 1, the ChargeNN's predictions of water monomer and dimer properties agree well with experimental results, demonstrating its capability to accurately simulate gas-phase

Table 1. Benchmark Results for the ChargeNN Water Model, Compared to SPC/FW, SWM4-NDP, and Experiments

|	units	ChargeNN	SPC/FW[a]	SWM4-NDP[b]	expt.[c]
Monomer					
μ_{total}	Debye	1.79	2.19	1.85	1.85
Dimer					
μ_{total}	Debye	2.397	3.594	2.062	2.643
r_{OO}	Å	2.81	2.73	2.83	2.98
θ_A	°	65	22	71	58
E_{int}	kcal/mol	−4.92	−7.14	−5.15	−5.4
Liquid					
$<r_{OO}>$	Å	0.97	1.03	0.96	0.97
$<\angle_{HOH}>$	°	106.1	107.7	104.52	106.5
ρ	g/cm³	0.996	1.012	0.998	0.997
E_{int}	kcal/mol	−9.917	−11.926	−9.927	−9.92
H_{vap}	kcal/mol	10.50	10.72	10.52	10.52
D	10^{-5} cm² s⁻¹	2.08	2.32	2.33	2.3
$<\mu_{mol}>$	Debye	1.76[d]	2.39	2.46	2.9
ϵ		50.9[e]	79.6	79.0	78.4
Ice					
T_{melt}	K	288	190	185	273

[a]The gas-phase properties, interaction energy, and melting temperature for SPC/FW were computed in this work, while other data were from refs 140−145. [b]SWM4-NDP results from refs 79,80. [c]Experimental results from refs 140−145 [d]The origin was set to be the center of the nuclear charge of each molecule. [e]Molecules were wrapped to the unit cell.

water clusters. Figure 4a illustrates that the angle θ_A of the water dimer obtained with ChargeNN (65°) closely matches the experimental value of 58°, in contrast to the unreasonable 22° by SPC/FW, highlighting the significant impact of polarization and charge transfer on the equilibrium geometries. Despite of the inclusion of intermolecular charge-transfer interactions in our model, enhanced electrostatics using high-order polarization may further improve the structure predictions.[146] For instance, the O−O distance of the water dimer optimized by the multipolar AMOEBA (2.91 Å)[67] aligns more closely with experimental measurement (2.98 Å) compared to ChargeNN (2.81 Å).

Table 1 also demonstrates that the ChargeNN water model accurately reproduces bulk water properties at room temperature and standard pressure, such as average bond length and angle, density, interaction energy, vaporization heat, and diffusion constants. As shown in Figure 4b, ChargeNN produces an RDF for liquid oxygen−oxygen distances in good agreement with the MP2[147] and early experimental data.[148] Nevertheless, the latest X-ray diffraction experiment[149] has measured a less pronounced first peak of approximately 2.57 than the predicted one of 2.98, indicating the need for further parameter refinement to mitigate the "over-structured" hydrogen-bond network. The predicted infrared spectra in Figure 4c were obtained through a Fourier transformation of the autocorrelation function of the total dipole moments, in which the one produced by ChargeNN matches the experimental spectrum closely in both band positions and intensities. Notably, ChargeNN successfully captures the hydrogen-bond stretching peak at ~200 cm⁻¹, which is absent in the spectra by SPC/FW and SWM4-NDP. Our results and a previous study[35] indicate that accurately describing hydrogen-bond stretching requires precise polarization and intermolecular CT. Additionally, ChargeNN retrieves the bending-libration combination band at ~2150
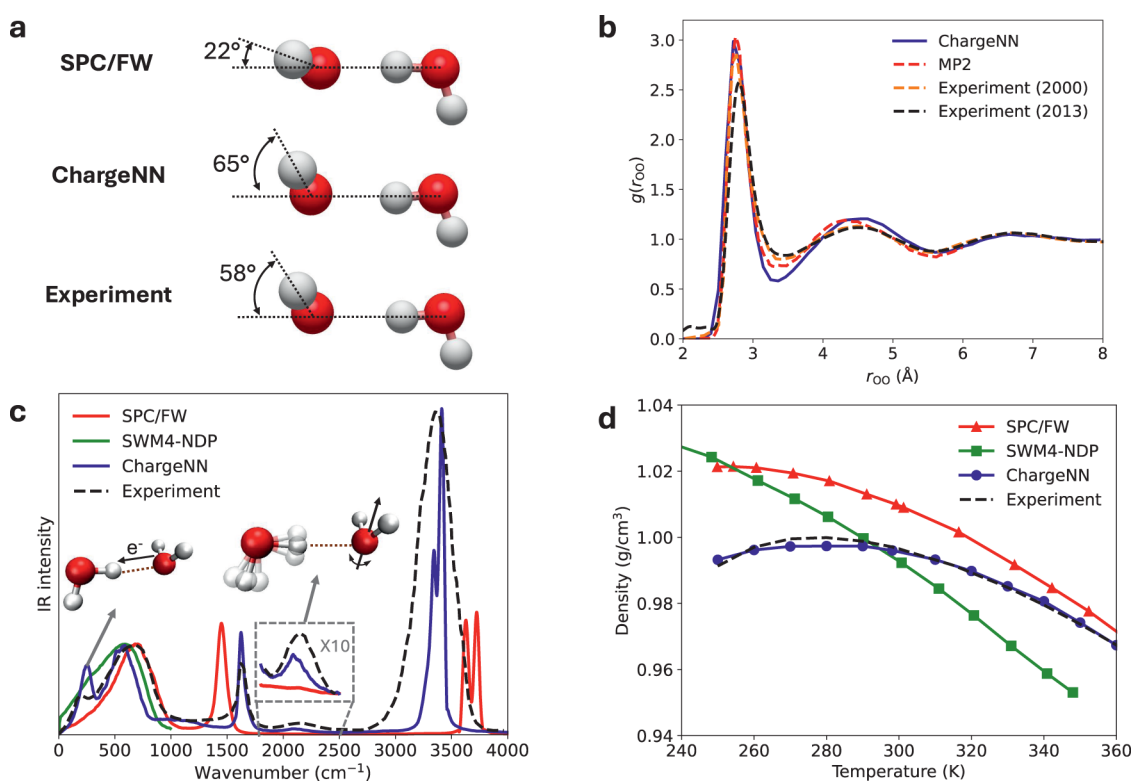
Figure 4. Assessment of the ChargeNN water model. (a) Geometries of water dimers optimized by SPC/FW and ChargeNN, compared to the experimental structure.[141] (b) The liquid radial distribution functions of the oxygen−oxygen distance obtained with ChargeNN, MP2[147] and experiments[148,149] under ambient conditions. (c) Liquid infrared spectra generated by SPC/FW, SWM4-NDP,[35] ChargeNN and experiment[152] under ambient conditions. (d) Temperature dependent densities from SPC/FW,[153] SWM4-NDP,[81] ChargeNN and experiment[154] at 1 atm.
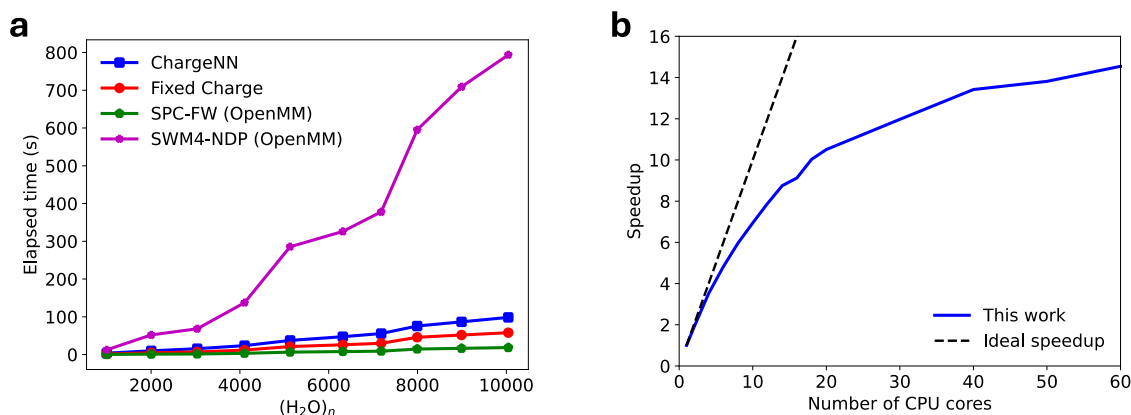

Figure 5. Timing performance of ChargeNN. (a) Elapsed time for force calculations on varying water box sizes using a single CPU core (AMD EPYC 7H12, 2.6 GHz), comparing in-house ChargeNN and fixed point charges models with OpenMM's SPC-FW and SWM4-NDP. (b) Speedup of ChargeNN calculations of periodic $(H_2O)_{10,044}$ with respect to the number of the CPU cores.

cm^{-1}, which is missing in the SPC/FW spectrum, further underscoring the importance of explicit polarization in predicting vibrational properties. Nevertheless, we observe that the O−H stretching band at around 3400 cm^{-1} is considerably narrower than the reference, which may be improved by replacing the harmonic intramolecular forces with anharmonic forces[66] or reactive force fields.[150,151]

We computed the molecular dipole moments by setting the origin as the nuclear charge center of each partially charged water molecule, following the reported protocol.[35,132] While ChargeNN accurately reproduces dipole moments for water monomer and dimer, as well as the liquid IR spectrum, it yields an average molecular dipole moment of approximately 1.76 D,

which is notably lower than the experimental estimation of 2.9 D. While the choice of the origins may have a significant impact on the dipole moments of charged molecules, there could be parallel problems in the population analysis methods. A recent comprehensive study discovered that none of the selected charge models (Mulliken, Hirshfeld, QTAIM, RESP, ChelpG, Hirshfeld-I, and NPA) is capable of reproducing both total dipole moments and average molecular dipole moments of water clusters.[132] Similarly, CM5 charges agree well with the reference values for total dipole moments, but fall in short to provide a reasonable average molecular dipole moment. Additionally, our water model underestimates the dielectric constant obtained from the total liquid dipole moments, which is
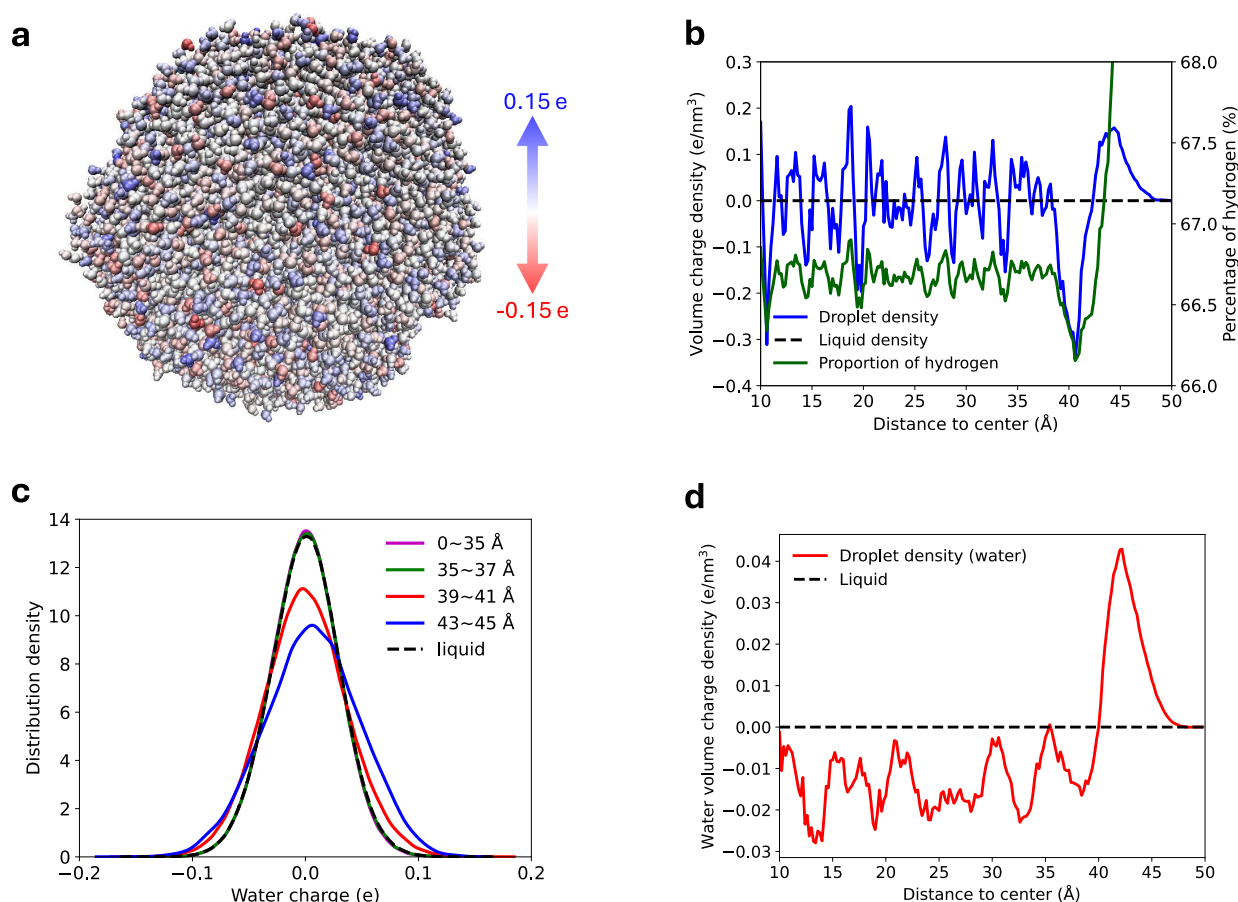
Figure 6. Statistical results for the illustrative application with a large droplet. (a) Geometry of the studied droplet colored according to molecular charges. (b) Correlation between VCD and PHA as a function of the layer distance to the droplet center. (c) Charge distributions of water molecules (d) Volume charge density using net molecular charges. The center of mass of each water molecule was used to compute the layer distance to the center.

largely due to the nonuniqueness of the dipole moment of an extended system according to the modern theory of polarization.[155]

The accurate characterization across a range of temperatures is important to the performance of a water model. Figure 4d demonstrates that ChargeNN successfully predicts the temperature-dependent density of liquid water, with the temperature of maximum density around 280 K, very close to the experimental value of 277 K. In contrast, this prediction remains a challenge for the fixed charge model SPC/FW and other well-tuned 3-site models,[60] emphasizing the substantial improvement resulting from ChargeNN polarization. The polarizable SWM4-NDP also falls short in predicting the density variation with temperature. Additionally, Figure S8 presents that the computed temperature-dependence of vaporization enthalpy by ChargeNN agrees well with the reference. For the simulation of ice, we calculated the orientational tetrahedral order parameters of the equilibrium configurations across various temperatures, starting from a perfect hexagonal (I_h) ice structure. The ChargeNN yields a melting temperature of 288 K, close to the experimental value of 273 K. While this result is significantly better than SPC/FW at 190 K and SWM4-NDP at 185 K,[80] there is still room to fine-tune the parameters to achieve ice properties comparable to TIP4P/Ice.[57]

As demonstrated in Figure 5a, ChargeNN exhibits exceptional computational efficiency, making it highly promising for large-scale water simulations. For instance, using a single CPU core for

a PBC box containing 10,044 water molecules, ChargeNN's force calculation time is only one-eighth of that for the economical polarizable SWM4-NDP and merely five times that of the 3-site fixed charge model SPC-FW, both implemented in OpenMM for CPU platforms. Moreover, while the incorporation of machine-learned charges involves additional steps such as feature preparation, charge prediction, and charge derivatives, these ChargeNN-related steps only account for approximately 40% of the total time. Furthermore, the non-ChargeNN components (equivalent to fixed charge computations) are three times slower than OpenMM's SPC-FW, indicating room for algorithmic improvement. In addition, Figure 5b reveals suboptimal parallel efficiency of the ChargeNN program, necessitating further optimization. We intend to transfer the algorithm to C++ for enhanced performance, with plans to adapt it to CUDA or OpenCL for extensive parallelization on graphics processing units (GPUs) in future implementations.

3.3. Demonstrative Application. As an illustrative application, we probe the origin of the emerging on-water reactivity on the air/water interface using ChargeNN. The electronic process for creating the strong electric fields on the water microdroplet surface remains elusive, and has been attributed to the abundance of dangling OH bonds at the interface,[156,157] the presence of OH^-/H^+ due to proton transfer[158−160] and electron transfer.[14,15,161] To understand the source of the interfacial electric fields, we carried out

statistical analysis to identify differences in geometries and charge distributions between the droplet and bulk liquid. We extracted the last 100 structural snapshots at 100 fs intervals from a 1 ns (ns) MD trajectory of a large water droplet with a radius ~45 Å containing 10327 water molecules. For the liquid simulation, we analyzed the last 500 snapshots from a 1 ns MD trajectory of a large unit cell containing 2123 water molecules.

Our results reveal that the primary factor influencing the layer charge density is the hydrogen-to-oxygen ratio. The volume charge density (VCD) has been extensively used as an indicator for the strength of the layer electric fields,[161] and is highly correlated with the proportion of the hydrogen atoms (PHA) (see Figure 6b). For deep layers inside the droplet at distances less than 37 Å from the center, the PHA fluctuates around an equilibrium percentage of 67% indicating an homogeneous distribution, with the VCD oscillating near the liquid VCD of 0 e/nm^3. The PHA on the intermediate layer at the distance of 41 Å from the center drops significantly to 66.2%, and coincides with the most negatively charged layer with a VCD of -0.35 e/nm^3. Beyond this layer, the PHA grows continuously to break the theoretical limit of 667% (that is, 2 H atoms and 1 O atom per water) and even exceed 68% at the air/water interface, revealing an inhomogeneous distribution of H-bond network near the surface, where the VCD increases consistently and reaches a maximum of 0.18 e/nm^3 at 45 Å. Our results align with the "dangling OH theory"[156,157] that more OH bonds dangle on the droplet surface due to the partial collapse of the hydrogen-bond network near the surface, leading to a higher proportion of hydrogen atoms and a positively charged surface layer. Consequently, the inward layer near the interface is abundant in oxygen atoms, and overall negatively charged.

In addition, the considerable charge transfer also significantly contributes to the layer charge density. As shown in Figure 6c, in the inner shells at distances shorter than 37 Å from the center, the molecular charges are closely clustered around zero, similar to the distribution in the liquid phase. In contrast, the vast interfacial charge separation results in a substantially increasing population of the charged water molecules. For understanding the explicit CT effect on the charge density, we obtained the water VCDs using net molecular charges of the water molecules whose centers of mass reside within the layers, as demonstrated in Figure 6d. The interior of the droplet is overall negatively charged with a water VCD of ~ -0.015 e/nm^3, indicating electron migration from the droplet surface to the inter layers. This is drastically different from the uniform charge density distribution in liquid phase. Furthermore, the water VCD of the positive droplet surface is approximately 0.045 e/nm^3, constituting 25% of the total VCD at the air/water interface.

4. APPLICATION OUTLOOK

Beyond the simulation of water, we can envisage broader applications with ChargeNN. First, the ChargeNN water model can be seamlessly combined with the existing force fields to simulate interactions between proteins and water solvents. Such interactions markedly affect the structures and dynamic behavior of proteins.[162] Additionally, integrating chargeNN water model with QM electron structures via a QM/MM interface will allow accurate descriptions of outer solvent shells for reactive centers. Moreover, incorporating empirical valence bond (EVB)[163] and reactive force fields (ReaxFF)[150,151] into ChargeNN will enable simulations of bond breaking and formation, facilitating efficient modeling of proton transfers through hydrogen-bond networks and chemical reactions.

Furthermore, accurate simulation of protein−ligand electrostatic interactions requires incorporating electronic polarizability, which induces screening effects that weaken electrostatics in the buried environment.[164] A recent study also shows that QM electronic polarization is essential for accurately producing spectral densities for proteins.[165] Therefore, we aim to extend the ChargeNN to other atoms and molecular systems, enabling generalizable predictions of atomic charges across diverse local environments and achieving QM-level charge assignments for complex macromolecules. Combining accurate charges and force fields will allow for the depiction of instantaneous polarization and transient charge transfer in molecular dynamics. Coupled with well-designed energy functionals, fast prediction of ab initio polarization from neural networks would significantly enhance the accuracy of atomistic simulations for large chemical systems other than water.

5. CONCLUSIONS

Modeling water has been pivotal task in theoretical chemistry, aiming at understanding its unique thermodynamic and electronic characteristics and propensities. Nevertheless, simulating polarization remains a challenge due to the volatile nature of electron distribution and its high sensitivity to local electric fields, which necessitates quantum mechanical treatments for accurate descriptions. Despite significant advancements in low-scaling QM electronic structures and advanced computing technology, QM calculations for large water systems remain prohibitively expensive. Machine learning offers promise for overcoming the cost obstacles and predicting atomic charges with QM accuracy. However, the existing ML charge models are limited to static calculations of specific systems, incapable of simulating polarization in both liquid and gas phase water during the structural and dynamical evolution.

In this work, we address the aforementioned problems by introducing a dynamic and polarizable water model that leverages machined-learned MP2-level partial charges. The charge model, termed ChargeNN, employs deep neural networks to map the interaction classified functions that accurately characterize the local environment onto the CM5 charges computed with MP2 electron density, demonstrating high accuracy and generalizability. Equipped with quantum atomic charges and charge transfer predicted by the neural networks, the ChargeNN water model successfully reproduces a variety of water properties across different temperatures and phases in excellent agreement with the experimental measurements, validating its capacity to generate electronic and thermodynamic quantities via fast and long molecular dynamics.

Additionally, we probed the origin of the strong local electric fields on the droplet surface by conducting molecular dynamics of both liquid water and a large droplet using the ChargeNN water model. The findings clearly indicate that the layer electric fields on the droplet surface are primarily induced by breakage of the hydrogen-bond network, which leads to distinct proportion of hydrogen atoms from the liquid phase. Furthermore, surface-to-interior charge transfer also considerably contributes to the layer electron density.

■ APPENDIX

■ MBE(3)-OSV-MP2 ELECTRONIC STRUCTURE

The atomic charges were calculated using the MBE(3)-OSV-MP2 method. The OSV approximation has demonstrated a

significant reduction in the computational effort required for correlated methods without considerable loss in accuracy.[129,166−171] The OSVs ($\mathbf{Q}_k$) can be obtained through a simple singular value decomposition of the diagonal pair MP2 amplitudes $\mathbf{T}_{kk}$:

$$\left[\mathbf{Q}_k^{\dagger}\mathbf{T}_{kk}\mathbf{Q}_k\right]_{\overline{\mu}_k\overline{\nu}_k} = \omega_{\overline{\mu}_k}\delta_{\overline{\mu\nu}} \tag{11}$$

The inherent sparsity allows the OSV space to be reduced by setting a cutoff for the singular values $\omega_{\overline{\mu}_k}$, which indicates the significance of OSVs. To further address the computational scaling challenge, a many-body expansion of OSV-MP2 amplitudes and density matrices with orbital-specific partitioning was introduced:

$$\mathbf{T}_{(ii,ii)} = \mathbf{T}_{(ii,ii)}^i + \sum_k \Delta\mathbf{T}_{(ii,ii)}^{i,k} + \sum_{k>l} \Delta\mathbf{T}_{(ii,ii)}^{i,k,l} \tag{12}$$

The population analysis can be performed using the total relaxed one particle density matrix, which can be computed as the sum of the density matrices of Hartree−Fock, relaxed MP2 and molecular orbital (MO) relaxation:

$$\mathbf{P} = \mathbf{P}^{HF} + \mathbf{P}^{MP2} + \mathbf{P}^{MO} \tag{13}$$

The relaxed OSV-MP2 density matrix in MO basis can be obtained with

$$\mathbf{P}^{MP2} = \sum_{ij} \mathbf{T}_{ii} < \mathbf{X}_{ij}^T + \mathbf{X}_{ji}^\perp> \tag{14}$$

where $\mathbf{T}_{ii}$ denotes the diagonal semicanonical amplitudes, while $\mathbf{X}_{ij}^T$ and $\mathbf{X}_{ji}^\perp$ represent the upper and lower diagonal blocks of the matrix containing OSV-MP2 responses.[169]

Additionally, we developed a large-scale parallelism strategy utilizing efficient memory management and data communication, which enables computations of MP2-level charges for large molecular systems.[129]

ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.4c01448.

> Results of the convergence of atomic charges and charge transfer with respect to atomic distances, testing MAEs relative to the training set size, parity graphs of molecular charges and charge transfer, optimized parameters for the ChargeNN water model, computational details for benchmarking ChargeNN, and some supplementary benchmark results (PDF)

AUTHOR INFORMATION

Corresponding Authors

Qiujiang Liang − *Department of Chemistry, The University of Hong Kong, Hong Kong 999077, P.R. China; Hong Kong Quantum AI Lab Limited, Hong Kong 999077, P.R. China;* ⓞ orcid.org/0000-0001-9416-1240; Email: qliang@connect.hku.hk

Jun Yang − *Department of Chemistry, The University of Hong Kong, Hong Kong 999077, P.R. China; Hong Kong Quantum AI Lab Limited, Hong Kong 999077, P.R. China;* ⓞ orcid.org/0000-0001-8701-9297; Email: juny@hku.hk

Complete contact information is available at:

https://pubs.acs.org/10.1021/acs.jctc.4c01448

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

REFERENCES

(1) Ball, P. Water as an active constituent in cell biology. *Chem. Rev.* **2008**, *108*, 74−108.

(2) Levy, Y.; Onuchic, J. N. Water mediation in protein folding and molecular recognition. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 389−415.

(3) Alexiadis, A.; Kassinos, S. Molecular simulation of water in carbon nanotubes. *Chem. Rev.* **2008**, *108*, 5014−5034.

(4) Dang, L. X. Importance of polarization effects in modeling the hydrogen bond in water using classical molecular dynamics techniques. *J. Phys. Chem. B* **1998**, *102*, 620−624.

(5) Piquemal, J.-P.; Chelli, R.; Procacci, P.; Gresh, N. Key role of the polarization anisotropy of water in modeling classical polarizable force fields. *J. Phys. Chem. A* **2007**, *111*, 8170−8176.

(6) Koocher, N. Z.; Saldana-Greco, D.; Wang, F.; Liu, S.; Rappe, A. M. Polarization dependence of water adsorption to CH3NH3PbI3 (001) surfaces. *J. Phys. Chem. Lett.* **2015**, *6*, 4371−4378.

(7) Schienbein, P.; Marx, D. Assessing the properties of supercritical water in terms of structural dynamics and electronic polarization effects. *Phys. Chem. Chem. Phys.* **2020**, *22*, 10462−10479.

(8) Grabowski, S. J. What is the covalency of hydrogen bonding? *Chem. Rev.* **2011**, *111*, 2597−2625.

(9) Mahadevi, A. S.; Sastry, G. N. Cooperativity in noncovalent interactions. *Chem. Rev.* **2016**, *116*, 2775−2825.

(10) Elgabarty, H.; Khaliullin, R. Z.; Kühne, T. D. Covalency of hydrogen bonds in liquid water can be probed by proton nuclear magnetic resonance experiments. *Nat. Commun.* **2015**, *6*, 8318.

(11) 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.

(12) Qiu, L.; Wei, Z.; Nie, H.; Cooks, R. G. Reaction acceleration promoted by partial solvation at the gas/solution interface. *ChemPlusChem.* **2021**, *86*, 1362−1365.

(13) Qiu, L.; Morato, N. M.; Huang, K.-H.; Cooks, R. G. Spontaneous water radical cation oxidation at double bonds in microdroplets. *Front. Chem.* **2022**, *10*, No. 903774.

(14) Ben-Amotz, D. Electric buzz in a glass of pure water. *Science* **2022**, *376*, 800−801.

(15) Liang, Q.; Zhu, C.; Yang, J. Water charge transfer accelerates Criegee intermediate reaction with H_2O−radical anion at the aqueous Interface. *J. Am. Chem. Soc.* **2023**, *145*, 10159−10166.

(16) Hedegård, E. D.; Bendix, J.; Sauer, S. P. Partial charges as reactivity descriptors for nitrido complexes. *J. Mol. Struct. THEOCHEM* **2009**, *913*, 1−7.

(17) Cheng, X.; Li, F.; Zhao, Y.; Cheng, X.; Nie, K.; Han, Y.; Yang, Y. Stability, atomic charges, bond-order analysis, and the directionality of lone-electron pairs on nitriles and isocyanides. *J. Phys. Org. Chem.* **2022**, *35*, No. e4420.

(18) De Proft, F.; Van Alsenoy, C.; Peeters, A.; Langenaeker, W.; Geerlings, P. Atomic charges, dipole moments, and Fukui functions using the Hirshfeld partitioning of the electron density. *J. Comput. Chem.* **2002**, *23*, 1198−1209.

(19) Nikolova, V.; Cheshmedzhieva, D.; Ilieva, S.; Galabov, B. Atomic charges in describing properties of aromatic molecules. *J. Org. Chem.* **2019**, *84*, 1908−1915.

(20) Kiryanov, I. I.; Mukminov, F. H.; Tulyabaev, A. R.; Khalilov, L. M. Prediction of 13C NMR chemical shifts by artificial neural network. I. Partial charge model as atomic descriptor. *Chemom. Intell. Lab. Syst.* **2016**, *152*, 62−68.

(21) Mulliken, R. S. Electronic population analysis on LCAO−MO molecular wave functions I. *J. Chem. Phys.* **1955**, *23*, 1833−1840.

(22) Löwdin, P.-O. On the non-orthogonality problem connected with the use of atomic wave functions in the theory of molecules and crystals. *J. Chem. Phys.* **1950**, *18*, 365−375.

(23) Foster, J. P.; Weinhold, F. Natural hybrid orbitals. *J. Am. Chem. Soc.* **1980**, *102*, 7211−7218.

(24) Bader, R. F.; Matta, C. F. Atomic charges are measurable quantum expectation values: a rebuttal of criticisms of QTAIM charges. *J. Phys. Chem. A* **2004**, *108*, 8385−8394.

(25) Hirshfeld, F. L. Bonded-atom fragments for describing molecular charge densities. *Theor. Chim. Acta* **1977**, *44*, 129−138.

(26) Marenich, A. V.; Jerome, S. V.; Cramer, C. J.; Truhlar, D. G. Charge model 5: An extension of Hirshfeld population analysis for the accurate description of molecular interactions in gaseous and condensed phases. *J. Chem. Theory Comput.* **2012**, *8*, 527−541.

(27) Wiberg, K. B.; Rablen, P. R. Atomic charges. *J. Org. Chem.* **2018**, *83*, 15463−15469.

(28) Houlding, S.; Liem, S.; Popelier, P. A polarizable high-rank quantum topological electrostatic potential developed using neural networks: Molecular dynamics simulations on the hydrogen fluoride dimer. *Int. J. Quantum Chem.* **2007**, *107*, 2817−2827.

(29) Handley, C. M.; Hawe, G. I.; Kell, D. B.; Popelier, P. L. Optimal construction of a fast and accurate polarisable water potential based on multipole moments trained by machine learning. *Phys. Chem. Chem. Phys.* **2009**, *11*, 6365−6376.

(30) Bleiziffer, P.; Schaller, K.; Riniker, S. Machine learning of partial charges derived from high-quality quantum-mechanical calculations. *J. Chem. Inf. Model.* **2018**, *58*, 579−590.

(31) Martin, R.; Heider, D. ContraDRG: automatic partial charge prediction by machine learning. *Front. Genet.* **2019**, *10*, 990.

(32) Wang, J.; Cao, D.; Tang, C.; Chen, X.; Sun, H.; Hou, T. Fast and accurate prediction of partial charges using atom-path-descriptor-based machine learning. *Bioinformatics* **2020**, *36*, 4721−4728.

(33) Wang, X.; Gao, J. Atomic partial charge predictions for furanoses by random forest regression with atom type symmetry function. *RSC Adv.* **2020**, *10*, 666−673.

(34) Kancharlapalli, S.; Gopalan, A.; Haranczyk, M.; Snurr, R. Q. Fast and accurate machine learning strategy for calculating partial atomic charges in metal−organic frameworks. *J. Chem. Theory Comput.* **2021**, *17*, 3052−3064.

(35) Han, B.; Isborn, C. M.; Shi, L. Incorporating polarization and charge transfer into a point-charge model for water using machine learning. *J. Phys. Chem. Lett.* **2023**, *14*, 3869−3877.

(36) Korolev, V. V.; Mitrofanov, A.; Marchenko, E. I.; Eremin, N. N.; Tkachenko, V.; Kalmykov, S. N. Transferable and extensible machine learning-derived atomic charges for modeling hybrid nanoporous materials. *Chem. Mater.* **2020**, *32*, 7822−7831.

(37) Nebgen, B.; Lubbers, N.; Smith, J. S.; Sifain, A. E.; Lokhov, A.; Isayev, O.; Roitberg, A. E.; Barros, K.; Tretiak, S. Transferable dynamic molecular charge assignment using deep neural networks. *J. Chem. Theory Comput.* **2018**, *14*, 4687−4698.

(38) Sifain, A. E.; Lubbers, N.; Nebgen, B. T.; Smith, J. S.; Lokhov, A. Y.; Isayev, O.; Roitberg, A. E.; Barros, K.; Tretiak, S. Discovering a transferable charge assignment model using machine learning. *J. Phys. Chem. Lett.* **2018**, *9*, 4495−4501.

(39) Unke, O. T.; Meuwly, M. PhysNet: A neural network for predicting energies, forces, dipole moments, and partial charges. *J. Chem. Theory Comput.* **2019**, *15*, 3678−3693.

(40) Kato, K.; Masuda, T.; Watanabe, C.; Miyagawa, N.; Mizouchi, H.; Nagase, S.; Kamisaka, K.; Oshima, K.; Ono, S.; Ueda, H.; et al. High-precision atomic charge prediction for protein systems using fragment molecular orbital calculation and machine learning. *J. Chem. Inf. Model.* **2020**, *60*, 3361−3368.

(41) Gallegos, M.; Guevara-Vela, J. M.; Pendás, Á. M. NNAIMQ: A neural network model for predicting QTAIM charges. *J. Chem. Phys.* **2022**, *156*, No. 014112.

(42) Gallegos, M.; Martín Pendás, Á. Developing a User-Friendly Code for the Fast Estimation of Well-Behaved Real-Space Partial Charges. *J. Chem. Inf. Model.* **2023**, *63*, 4100−4114.

(43) Gallegos, M.; Vassilev-Galindo, V.; Poltavsky, I.; Martín Pendás, Á.; Tkatchenko, A. Explainable chemical artificial intelligence from accurate machine learning of real-space chemical descriptors. *Nat. Commun.* **2024**, *15*, 4345.

(44) Raza, A.; Sturluson, A.; Simon, C. M.; Fern, X. Message passing neural networks for partial charge assignment to metal−organic frameworks. *J. Phys. Chem. C* **2020**, *124*, 19070−19082.

(45) Wang, J.; Sun, H.; Chen, J.; Jiang, D.; Wang, Z.; Wu, Z.; Chen, X.; Cao, D.; Hou, T. DeepChargePredictor: a web server for predicting QM-based atomic charges via state-of-the-art machine-learning algorithms. *Bioinformatics* **2021**, *37*, 4255−4257.

(46) Wang, J.; Cao, D.; Tang, C.; Xu, L.; He, Q.; Yang, B.; Chen, X.; Sun, H.; Hou, T. DeepAtomicCharge: a new graph convolutional network-based architecture for accurate prediction of atomic charges. *Briefings Bioinf.* **2021**, *22*, No. bbaa183.

(47) Metcalf, D. P.; Jiang, A.; Spronk, S. A.; Cheney, D. L.; Sherrill, C. D. Electron-passing neural networks for atomic charge prediction in systems with arbitrary molecular charge. *J. Chem. Inf. Model.* **2021**, *61*, 115−122.

(48) Jiang, D.; Sun, H.; Wang, J.; Hsieh, C. Y.; Li, Y.; Wu, Z.; Cao, D.; Wu, J.; Hou, T. Out-of-the-box deep learning prediction of quantum-mechanical partial charges by graph representation and transfer learning. *Briefings Bioinf.* **2022**, *23*, No. bbab597.

(49) Lehner, M. T.; Katzberger, P.; Maeder, N.; Schiebroek, C. C.; Teetz, J.; Landrum, G. A.; Riniker, S. DASH: Dynamic Attention-Based Substructure Hierarchy for Partial Charge Assignment. *J. Chem. Inf. Model.* **2023**, *63*, 6014−6028.

(50) Wang, Y.; Pulido, I.; Takaba, K.; Kaminow, B.; Scheen, J.; Wang, L.; Chodera, J. D. EspalomaCharge: Machine Learning-Enabled Ultrafast Partial Charge Assignment. *J. Phys. Chem. A* **2024**, *128*, 4160−4167.

(51) Bultinck, P.; Van Alsenoy, C.; Ayers, P. W.; Carbó-Dorca, R. Critical analysis and extension of the Hirshfeld atoms in molecules. *J. Chem. Phys.* **2007**, *126*, 144111.

(52) Demerdash, O.; Wang, L.; Head-Gordon, T. Advanced models for water simulations. *WIREs Mol. Sci.* **2018**, *8*, No. e1355.

(53) Berendsen, H. J.; Postma, J. P.; van Gunsteren, W. F.; Hermans, J. Interaction models for water in relation to protein hydration. *Intermolecular forces: proceedings of the fourteenth Jerusalem symposium on quantum chemistry and biochemistry held in jerusalem* **1981**, *14*, 331−342.

(54) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.

(55) Mahoney, M. W.; Jorgensen, W. L. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J. Chem. Phys.* **2000**, *112*, 8910−8922.

(56) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J. Chem. Phys.* **2004**, *120*, 9665−9678.

(57) Abascal, J. L. F.; Sanz, E.; García Fernández, R.; Vega, C. A potential model for the study of ices and amorphous water: TIP4P/Ice. *J. Chem. Phys.* **2005**, *122*, 234511.

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

(59) Izadi, S.; Anandakrishnan, R.; Onufriev, A. V. Building water models: a different approach. *J. Phys. Chem. Lett.* **2014**, *5*, 3863−3871.

(60) Izadi, S.; Onufriev, A. V. Accuracy limit of rigid 3-point water models. *J. Chem. Phys.* **2016**, *145*, No. 074501.

(61) Wu, Y.; Tepper, H. L.; Voth, G. A. Flexible simple point-charge water model with improved liquid-state properties. *J. Chem. Phys.* **2006**, *124*, No. 024503.

(62) 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.

(63) Habershon, S.; Markland, T. E.; Manolopoulos, D. E. Competing quantum effects in the dynamics of a flexible water model. *J. Chem. Phys.* **2009**, *131*, No. 024501.

(64) Berendsen, H. J.; Grigera, J.-R.; Straatsma, T. P. The missing term in effective pair potentials. *J. Phys. Chem.* **1987**, *91*, 6269−6271.

(65) Caldwell, J. W.; Kollman, P. A. Structure and properties of neat liquids using nonadditive molecular dynamics: water, methanol, and N-methylacetamide. *J. Phys. Chem.* **1995**, *99*, 6208−6219.

(66) Shi, Y.; Xia, Z.; Zhang, J.; Best, R.; Wu, C.; Ponder, J. W.; Ren, P. Polarizable atomic multipole-based AMOEBA force field for proteins. *J. Chem. Theory Comput.* **2013**, *9*, 4046−4063.

(67) Laury, M. L.; Wang, L.-P.; Pande, V. S.; Head-Gordon, T.; Ponder, J. W. Revised parameters for the AMOEBA polarizable atomic multipole water model. *J. Phys. Chem. B* **2015**, *119*, 9423−9437.

(68) Liu, C.; Piquemal, J.-P.; Ren, P. AMOEBA+ classical potential for modeling molecular interactions. *J. Chem. Theory Comput.* **2019**, *15*, 4122−4139.

(69) Liu, C.; Piquemal, J.-P.; Ren, P. Implementation of geometry-dependent charge flux into the polarizable AMOEBA+ potential. *J. Phys. Chem. Lett.* **2020**, *11*, 419−426.

(70) Burnham, C. J.; Li, J.; Xantheas, S. S.; Leslie, M. The parametrization of a Thole-type all-atom polarizable water model from first principles and its application to the study of water clusters (n= 2−21) and the phonon spectrum of ice Ih. *J. Chem. Phys.* **1999**, *110*, 4566−4581.

(71) Burnham, C. J.; Xantheas, S. S. Development of transferable interaction models for water. IV. A flexible, all-atom polarizable potential (TTM2-F) based on geometry dependent charges derived from an ab initio monomer dipole moment surface. *J. Chem. Phys.* **2002**, *116*, 5115−5124.

(72) Fanourgakis, G. S.; Xantheas, S. S. Development of transferable interaction potentials for water. V. Extension of the flexible, polarizable, Thole-type model potential (TTM3-F, v. 3.0) to describe the vibrational spectra of water clusters and liquid water. *J. Chem. Phys.* **2008**, *128*, No. 074506.

(73) Heindel, J. P.; Sami, S.; Head-Gordon, T. Completely Multipolar Model as a General Framework for Many-Body Interactions as Illustrated for Water. *J. Chem. Theory Comput.* **2024**, *20*, 8594−8608.

(74) Symons, B. C.; Bane, M. K.; Popelier, P. L. DL_FFLUX: a parallel, quantum chemical topology force field. *J. Chem. Theory Comput.* **2021**, *17*, 7043−7055.

(75) Hughes, Z. E.; Ren, E.; Thacker, J. C.; Symons, B. C.; Silva, A. F.; Popelier, P. L. A FFLUX water model: Flexible, polarizable and with a multipolar description of electrostatics. *J. Comput. Chem.* **2020**, *41*, 619−628.

(76) Symons, B. C.; Popelier, P. L. Application of quantum chemical topology force field FFLUX to condensed matter simulations: Liquid water. *J. Chem. Theory Comput.* **2022**, *18*, 5577−5588.

(77) Brown, M. L.; Skelton, J. M.; Popelier, P. L. Application of the FFLUX Force Field to Molecular Crystals: A Study of Formamide. *J. Chem. Theory Comput.* **2023**, *19*, 7946−7959.

(78) Lamoureux, G.; MacKerell, A. D.; Roux, B. A simple polarizable model of water based on classical Drude oscillators. *J. Chem. Phys.* **2003**, *119*, 5185−5197.

(79) Lamoureux, G.; Harder, E.; Vorobyov, I. V.; Roux, B.; MacKerell, A. D., Jr A polarizable model of water for molecular dynamics simulations of biomolecules. *Chem. Phys. Lett.* **2006**, *418*, 245−249.

(80) Wang, X.; Tse, Y.-L. S. Flexible Polarizable Water Model Parameterized via Gaussian Process Regression. *J. Chem. Theory Comput.* **2022**, *18*, 7155−7165.

(81) Yu, W.; Lopes, P. E. M.; Roux, B.; MacKerell, A. D. Six-site polarizable model of water based on the classical Drude oscillator. *J. Chem. Phys.* **2013**, *138*, No. 034508.

(82) Xiong, Y.; Izadi, S.; Onufriev, A. V. Fast polarizable water model for atomistic simulations. *J. Chem. Theory Comput.* **2022**, *18*, 6324−6333.

(83) Rappe, A. K.; Goddard, W. A., III Charge equilibration for molecular dynamics simulations. *J. Phys. Chem.* **1991**, *95*, 3358−3363.

(84) Rick, S. W.; Stuart, S. J.; Berne, B. J. Dynamical fluctuating charge force fields: Application to liquid water. *J. Chem. Phys.* **1994**, *101*, 6141−6156.

(85) Oliferenko, A. A.; Pisarev, S. A.; Palyulin, V. A.; Zefirov, N. S. Atomic charges via electronegativity equalization: generalizations and perspectives. *Adv. Quantum Chem.* **2006**, *51*, 139−156.

(86) Patel, S.; Brooks, C. L., III CHARMM fluctuating charge force field for proteins: I parameterization and application to bulk organic liquid simulations. *J. Comput. Chem.* **2004**, *25*, 1−16.

(87) Leven, I.; Hao, H.; Tan, S.; Guan, X.; Penrod, K. A.; Akbarian, D.; Evangelisti, B.; Hossain, M. J.; Islam, M. M.; Koski, J. P.; et al. Recent advances for improving the accuracy, transferability, and efficiency of reactive force fields. *J. Chem. Theory Comput.* **2021**, *17*, 3237−3251.

(88) Chen, C.; Deng, Z.; Tran, R.; Tang, H.; Chu, I.-H.; Ong, S. P. Accurate force field for molybdenum by machine learning large materials data. *Phys. Rev. Mater.* **2017**, *1*, No. 043603.

(89) Banks, J. L.; Kaminski, G. A.; Zhou, R.; Mainz, D. T.; Berne, B.; Friesner, R. A. Parametrizing a polarizable force field from ab initio data. I. The fluctuating point charge model. *J. Chem. Phys.* **1999**, *110*, 741−754.

(90) Nistor, R. A.; Polihronov, J. G.; Müser, M. H.; Mosey, N. J. A generalization of the charge equilibration method for nonmetallic materials. *J. Chem. Phys.* **2006**, *125*, No. 094108.

(91) Müser, M. H. The chemical hardness of molecules and the band gap of solids within charge equilibration formalisms: Toward force field-based simulations of redox reactions. *Eur. Phys. J. B* **2012**, *85*, 135.

(92) Dapp, W. B.; Müser, M. H. Towards time-dependent, non-equilibrium charge-transfer force fields: Contact electrification and history-dependent dissociation limits. *Eur. Phys. J. B* **2013**, *86*, 337.

(93) Morita, A.; Kato, S. Molecular dynamics simulation with the charge response kernel: Diffusion dynamics of pyrazine and pyrazinyl radical in methanol. *J. Chem. Phys.* **1998**, *108*, 6809−6818.

(94) Nakano, H.; Yamamoto, T.; Kato, S. A wave-function based approach for polarizable charge model: Systematic comparison of polarization effects on protic, aprotic, and ionic liquids. *J. Chem. Phys.* **2010**, *132*, No. 044106.

(95) Verstraelen, T.; Ayers, P. W.; Van Speybroeck, V.; Waroquier, M. ACKS2: Atom-condensed Kohn-Sham DFT approximated to second order. *J. Chem. Phys.* **2013**, *138*, No. 074108.

(96) Verstraelen, T.; Vandenbrande, S.; Ayers, P. W. Direct computation of parameters for accurate polarizable force fields. *J. Chem. Phys.* **2014**, *141*, 194114.

(97) Holt, A.; Karlström, G.; Lindh, R. The charge capacity of the chemical bond. *Chem. Phys. Lett.* **2007**, *436*, 297−301.

(98) Holt, A.; Karlström, G.; Roos, B. O. The charge capacitance of the chemical bond: Application to bonds containing metals. *Int. J. Quantum Chem.* **2009**, *109*, 618−628.

(99) Jensen, F. Using atomic charges to model molecular polarization. *Phys. Chem. Chem. Phys.* **2022**, *24*, 1926−1943.

(100) Jensen, F. Unifying Charge-Flow Polarization Models. *J. Chem. Theory Comput.* **2023**, *19*, 4047−4073.

(101) 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.

(102) Bukowski, R.; Szalewicz, K.; Groenenboom, G. C.; van der Avoird, A. Polarizable interaction potential for water from coupled cluster calculations. II. Applications to dimer spectra, virial coefficients, and simulations of liquid water. *J. Chem. Phys.* **2008**, *128*, No. 094314.

(103) Huang, X.; Braams, B. J.; Bowman, J. M. Ab initio potential energy and dipole moment surfaces of $(H_2O)_2$. *J. Phys. Chem. A* **2006**, *110*, 445−451.

(104) Wang, Y.; Huang, X.; Shepler, B. C.; Braams, B. J.; Bowman, J. M. Flexible, ab initio potential, and dipole moment surfaces for water. I. Tests and applications for clusters up to the 22-mer. *J. Chem. Phys.* **2011**, *134*, No. 094509.

(105) Wang, Y.; Bowman, J. M. Ab initio potential and dipole moment surfaces for water. II. Local-monomer calculations of the infrared spectra of water clusters. *J. Chem. Phys.* **2011**, *134*, 154510.

(106) Wang, Y.; Shepler, B. C.; Braams, B. J.; Bowman, J. M. Full-dimensional, ab initio potential energy and dipole moment surfaces for water. *J. Chem. Phys.* **2009**, *131*, No. 054511.

(107) 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*, 1103−1114.

(108) Babin, V.; Medders, G. R.; Paesani, F. Toward a universal water model: First principles simulations from the dimer to the liquid phase. *J. Phys. Chem. Lett.* **2012**, *3*, 3765−3769.

(109) 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.

(110) Babin, V.; Medders, G. R.; Paesani, F. Development of a first principles water potential with flexible monomers. II: Trimer potential energy surface, third virial coefficient, and small clusters. *J. Chem. Theory Comput.* **2014**, *10*, 1599−1607.

(111) Medders, G. R.; Babin, V.; Paesani, F. Development of a first-principles water potential with flexible monomers. III. Liquid phase properties. *J. Chem. Theory Comput.* **2014**, *10*, 2906−2910.

(112) Behler, J.; Parrinello, M. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.* **2007**, *98*, No. 146401.

(113) 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*, No. 136403.

(114) Bartók, A. P.; Kondor, R.; Csányi, G. On representing chemical environments. *Physical Review B—Condensed Matter and Materials. Physics* **2013**, *87*, No. 184115.

(115) Rupp, M.; Tkatchenko, A.; Müller, K.-R.; Von Lilienfeld, O. A. Fast and accurate modeling of molecular atomization energies with machine learning. *Phys. Rev. Lett.* **2012**, *108*, No. 058301.

(116) Unke, O. T.; Chmiela, S.; Sauceda, H. E.; Gastegger, M.; Poltavsky, I.; Schütt, K. T.; Tkatchenko, A.; Müller, K.-R. Machine learning force fields. *Chem. Rev.* **2021**, *121*, 10142−10186.

(117) Yu, Q.; Qu, C.; Houston, P. L.; Nandi, A.; Pandey, P.; Conte, R.; Bowman, J. M. A status report on Gold Standard machine-learned potentials for water. *J. Phys. Chem. Lett.* **2023**, *14*, 8077−8087.

(118) Thacker, J. C.; Wilson, A. L.; Hughes, Z. E.; Burn, M. J.; Maxwell, P. I.; Popelier, P. L. Towards the simulation of biomolecules: optimization of peptide-capped glycine using FFLUX. *Mol. Simul.* **2018**, *44*, 881−890.

(119) Bader, R. F.; Molecules, A. I. *A quantum theory*; Clarendon: Oxford, UK, 1990,.

(120) Bader, R. F. A quantum theory of molecular structure and its applications. *Chem. Rev.* **1991**, *91*, 893−928.

(121) Reed, A. E.; Weinstock, R. B.; Weinhold, F. Natural population analysis. *J. Chem. Phys.* **1985**, *83*, 735−746.

(122) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma,

K.; Farkas, O.; Foresman, J. B.; Fox, D. J *Gaussian 1∼6 Revision C.01*; Gaussian Inc: Wallingford CT, 2016.

(123) Lu, T.; Chen, F. Multiwfn: A multifunctional wavefunction analyzer. *J. Comput. Chem.* **2012**, *33*, 580−592.

(124) Lu, T. A comprehensive electron wavefunction analysis toolbox for chemists, Multiwfn. *J. Chem. Phys.* **2024**, *161*, No. 082503.

(125) Maheshwary, S.; Patel, N.; Sathyamurthy, N.; Kulkarni, A. D.; Gadre, S. R. Structure and stability of water clusters $(H_2O)_n$, n = 8−20: an ab initio investigation. *J. Phys. Chem. A* **2001**, *105*, 10525−10537.

(126) Clough, S. A.; Beers, Y.; Klein, G. P.; Rothman, L. S. Dipole moment of water from Stark measurements of H2O, HDO, and D2O. *J. Chem. Phys.* **1973**, *59*, 2254−2259.

(127) Behler, J. Atom-centered symmetry functions for constructing high-dimensional neural network potentials. *J. Chem. Phys.* **2011**, *134*, No. 074106.

(128) Behler, J. Perspective: Machine learning potentials for atomistic simulations. *J. Chem. Phys.* **2016**, *145*, 170901.

(129) Liang, Q.; Yang, J. Third-order many-body expansion of OSV-MP2 wave function for low-order scaling analytical gradient computation. *J. Chem. Theory Comput.* **2021**, *17*, 6841−6860.

(130) Miliordos, E.; Aprà, E.; Xantheas, S. S. Optimal geometries and harmonic vibrational frequencies of the global minima of water clusters $(H_2O)_n$, n= 2−6, and several hexamer local minima at the CCSD (T) level of theory. *J. Chem. Phys.* **2013**, *139*, 114302.

(131) Miliordos, E.; Xantheas, S. S. An accurate and efficient computational protocol for obtaining the complete basis set limits of the binding energies of water clusters at the MP2 and CCSD (T) levels of theory: Application to $(H_2O)_m$, m= 2−6, 8, 11, 16, and 17. *J. Chem. Phys.* **2015**, *142*, 234303.

(132) Han, B.; Isborn, C. M.; Shi, L. Determining partial atomic charges for liquid water: assessing electronic structure and charge models. *J. Chem. Theory Comput.* **2021**, *17*, 889−901.

(133) Eastman, P.; Friedrichs, M. S.; Chodera, J. D.; Radmer, R. J.; Bruns, C. M.; Ku, J. P.; Beauchamp, K. A.; Lane, T. J.; Wang, L.-P.; Shukla, D.; et al. OpenMM 4: a reusable, extensible, hardware independent library for high performance molecular simulation. *J. Chem. Theory Comput.* **2013**, *9*, 461−469.

(134) Huang, J.; Lemkul, J. A.; Eastman, P. K.; MacKerell, A. D., Jr. Molecular dynamics simulations using the drude polarizable force field on GPUs with OpenMM: Implementation, validation, and benchmarks. *J. Comput. Chem.* **2018**, *39*, 1682−1689.

(135) Ramachandran, P.; Zoph, B.; Le, Q. V. Searching for activation functions. *arXiv preprint arXiv:1710.05941* 2017

(136) Liu, J.; Miller, W. H.; Fanourgakis, G. S.; Xantheas, S. S.; Imoto, S.; Saito, S. Insights in quantum dynamical effects in the infrared spectroscopy of liquid water from a semiclassical study with an ab initio-based flexible and polarizable force field. *J. Chem. Phys.* **2011**, *135*, 244503.

(137) Soniat, M.; Rick, S. W. The effects of charge transfer on the aqueous solvation of ions. *J. Chem. Phys.* **2012**, *137*, No. 044511.

(138) Rick, S. W.; Thompson, W. H. Effects of polarizability and charge transfer on water dynamics and the underlying activation energies. *J. Chem. Phys.* **2023**, *158*, 194504.

(139) Kumar Sinha, S.; Mehta, M.; Patel, S. A charge equilibration formalism for treating charge transfer effects in MD simulations: Application to water clusters. *J. Comput. Chem.* **2017**, *38*, 1389−1409.

(140) Dyke, T. R.; Muenter, J. Electric dipole moments of low J states of H_2O and D_2O. *J. Chem. Phys.* **1973**, *59*, 3125−3127.

(141) Dyke, T. R.; Mack, K. M.; Muenter, J. The structure of water dimer from molecular beam electric resonance spectroscopy. *J. Chem. Phys.* **1977**, *66*, 498−510.

(142) Jancso, G.; Van Hook, W. A. Condensed phase isotope effects. *Chem. Rev.* **1974**, *74*, 689−750.

(143) Krynicki, K.; Green, C. D.; Sawyer, D. W. Pressure and temperature dependence of self-diffusion in water. *Faraday Discuss. Chem. Soc.* **1978**, *66*, 199−208.

(144) Fernandez, D. P.; Mulev, Y.; Goodwin, A.; Sengers, J. L. A database for the static dielectric constant of water and steam. *J. Phys. Chem. Ref. Data* **1995**, *24*, 33−70.

(145) Badyal, Y.; Saboungi, M.-L.; Price, D.; Shastri, S.; Haeffner, D.; Soper, A. Electron distribution in water. *J. Chem. Phys.* **2000**, *112*, 9206−9208.

(146) Cardamone, S.; Hughes, T. J.; Popelier, P. L. Multipolar electrostatics. *Phys. Chem. Chem. Phys.* **2014**, *16*, 10367−10387.

(147) Del Ben, M.; Schönherr, M.; Hutter, J.; VandeVondele, J. Bulk liquid water at ambient temperature and pressure from MP2 theory. *J. Phys. Chem. Lett.* **2013**, *4*, 3753−3759.

(148) Sorenson, J. M.; Hura, G.; Glaeser, R. M.; Head-Gordon, T. What can x-ray scattering tell us about the radial distribution functions of water? *J. Chem. Phys.* **2000**, *113*, 9149−9161.

(149) Skinner, L. B.; Huang, C.; Schlesinger, D.; Pettersson, L. G. M.; Nilsson, A.; Benmore, C. J. Benchmark oxygen-oxygen pair-distribution function of ambient water from x-ray diffraction measurements with a wide Q-range. *J. Chem. Phys.* **2013**, *138*, No. 074506.

(150) Van Duin, A. C.; Dasgupta, S.; Lorant, F.; Goddard, W. A. ReaxFF: a reactive force field for hydrocarbons. *J. Phys. Chem. A* **2001**, *105*, 9396−9409.

(151) Leven, I.; Hao, H.; Das, A. K.; Head-Gordon, T. A reactive force field with coarse-grained electrons for liquid water. *J. Phys. Chem. Lett.* **2020**, *11*, 9240−9247.

(152) Max, J.-J.; Chapados, C. Isotope effects in liquid water by infrared spectroscopy. III. H_2O and D_2O spectra from 6000 to cm^{-1}. *J. Chem. Phys.* **2009**, *131*, 184505.

(153) Valle, J. V.; Mendonça, B. H.; Barbosa, M. C.; Chacham, H.; de Moraes, E. E. Accuracy of TIP4P/2005 and SPC/Fw Water Models. *J. Phys. Chem. B* **2024**, *128*, 1091−1097.

(154) Kell, G. S. Density, thermal expansivity, and compressibility of liquid water from 0.deg. to 150.deg. Correlations and tables for atmospheric pressure and saturation reviewed and expressed on 1968 temperature scale. *J. Chem. Eng. Data* **1975**, *20*, 97−105.

(155) Resta, R.; Vanderbilt, D. *Physics of ferroelectrics: a modern perspective*; Springer, 2007; 31−68.

(156) Inoue, K.-I; Ahmed, M.; Nihonyanagi, S.; Tahara, T. Reorientation-induced relaxation of free OH at the air/water interface revealed by ultrafast heterodyne-detected nonlinear spectroscopy. *Nat. Commun.* **2020**, *11*, 5344.

(157) Pezzotti, S.; Galimberti, D. R.; Gaigeot, M.-P. 2D H-bond network as the topmost skin to the air−water interface. *J. Phys. Chem. Lett.* **2017**, *8*, 3133−3141.

(158) Hao, H.; Leven, I.; Head-Gordon, T. Can electric fields drive chemistry for an aqueous microdroplet? *Nat. Commun.* **2022**, *13*, 280.

(159) Heindel, J. P.; Hao, H.; LaCour, R. A.; Head-Gordon, T. Spontaneous formation of hydrogen peroxide in water microdroplets. *J. Phys. Chem. Lett.* **2022**, *13*, 10035−10041.

(160) Heindel, J. P.; LaCour, R. A.; Head-Gordon, T. The role of charge in microdroplet redox chemistry. *Nat. Commun.* **2024**, *15*, 3670.

(161) Poli, E.; Jong, K. H.; Hassanali, A. Charge transfer as a ubiquitous mechanism in determining the negative charge at hydrophobic interfaces. *Nat. Commun.* **2020**, *11*, 901.

(162) Prabhu, N.; Sharp, K. Protein- solvent interactions. *Chem. Rev.* **2006**, *106*, 1616−1623.

(163) Warshel, A.; Weiss, R. M. An empirical valence bond approach for comparing reactions in solutions and in enzymes. *J. Am. Chem. Soc.* **1980**, *102*, 6218−6226.

(164) Rocklin, G. J.; Boyce, S. E.; Fischer, M.; Fish, I.; Mobley, D. L.; Shoichet, B. K.; Dill, K. A. Blind prediction of charged ligand binding affinities in a model binding site. *J. Mol. Biol.* **2013**, *425*, 4569−4583.

(165) Zuehlsdorff, T. J.; Hong, H.; Shi, L.; Isborn, C. M. Influence of electronic polarization on the spectral density. *J. Phys. Chem. B* **2020**, *124*, 531−543.

(166) Yang, J.; Kurashige, Y.; Manby, F. R.; Chan, G. K. L. Tensor factorizations of local second-order Møller−Plesset theory. *J. Chem. Phys.* **2011**, *134*, No. 044123.

(167) Kurashige, Y.; Yang, J.; Chan, G. K. L.; Manby, F. R. Optimization of orbital-specific virtuals in local Møller-Plesset perturbation theory. *J. Chem. Phys.* **2012**, *136*, 124106.

(168) Yang, J.; Chan, G. K.-L.; Manby, F. R.; Schütz, M.; Werner, H.-J. The orbital-specific-virtual local coupled cluster singles and doubles method. *J. Chem. Phys.* **2012**, *136*, 144105.

(169) Zhou, R.; Liang, Q.; Yang, J. Complete OSV-MP2 Analytical Gradient Theory for Molecular Structure and Dynamics Simulations. *J. Chem. Theory Comput.* **2020**, *16*, 196−210.

(170) Ng, W.-P.; Liang, Q.; Yang, J. Low-data deep quantum chemical learning for accurate MP2 and coupled-cluster correlations. *J. Chem. Theory Comput.* **2023**, *19*, 5439−5449.

(171) Yang, J. Making quantum chemistry compressive and expressive: Toward practical ab-initio simulation. *WIREs Mol. Sci.* **2024**, *14*, No. e1706.