Ab-initio Structure and Thermodynamics of the RPBE-D3 Water/Vapor Interface by Neural-Network Molecular Dynamics

Oliver Wohlfahrt, Christoph Dellago, and Marcello Sega

1) University of Vienna, Faculty of Physics, Boltzmann-gasse 5, A-1090, Vienna, Austria
2) Forschungszentrum Jülich GmbH, Helmholtz Institute Erlangen-Nürnberg for Renewable Energy (IEK-11), Fürther Str. 248, D-90429, Nürnberg, Germany

(Dated: July 21, 2020)

Aided by a neural network representation of the density functional theory (DFT) potential energy landscape of water in the RPBE approximation corrected for dispersion, we calculate several structural and thermodynamic properties of its liquid/vapor interface. The neural network speed allows us to bridge the size and time scale gaps required to sample the properties of water along its liquid/vapor coexistence line with unprecedented precision.

I. INTRODUCTION

The strong, directed network of hydrogen bonds that confers water its rich phase diagram, and its numerous anomalous properties, is also responsible for the peculiar structure of its liquid/vapor interface. Thanks to its hydrogen bond network, water is one of the most cohesive liquids known, with exceptionally high surface tension.

The anomalous surface tension is, by far, not the only surprising property of water interfaces. Unlike simple liquids, water undergoes substantial structural changes in the interface’s first molecular layer. In its preferential configuration, a surface molecule at the liquid/vapor interface reorients its dipole vector along the liquid surface, with one OH bond and one lone pair both directed towards the vapor phase. The rearrangement of water molecules at the interface with other phases, with the corresponding entropic change, lies at the very heart of phenomena like the hydrophobic effect, one of the main actors of self-organization in living organisms, and of many physicochemical based applications, from detergents to novel materials based on microemulsions.

Understanding the relationship between the microscopic structure of the water/vapor interface and its thermodynamic properties is then key to obtaining better control over a vast array of processes. Due to its molecular-scale extension, interfacial water can be investigated experimentally by a small set of techniques, most importantly, X-ray reflectivity and sum-frequency generation spectroscopy. In this sense, computer simulations techniques, providing direct access to the molecular configurations, are a unique asset. However, the accuracy of the calculations and their computational cost are two significant challenges for the simulation of water in general and its interfaces in particular. Ideally, one would aim at an accurate parameter-free description. While much progress has been made to achieve better ab-initio simulations of liquid water, the computational resources required to simulate liquid interfaces using state-of-the-art methods are still daunting. Artificial neural networks have proven to be a valuable tool to reproduce the features of the potential energy landscapes of water. Using an artificial neural network to reproduce the forces acting between nuclei as computed with an ab-initio method of choice can decrease the computational cost of the problem by several orders of magnitude.

Here, we extend previously parameterized neural networks by including explicitly interfacial configurations and use the optimized potentials to study the phase diagram of water along its liquid/vapor coexistence line. The need for relatively large systems is particularly pressing for interfacial systems as they suffer more than bulk systems from finite-size effect. The computational advantage warranted by the neural network allowed us to simulate for many nanoseconds simulation boxes containing 1024 water molecules, based on the RPBE generalized gradient approximation for the exchange-correlation functional, supplemented by Grimme’s D3 dispersion corrections.
II. EXTENDING THE NEURAL NETWORK TRAINING SET

As a first step, we simulated a water/vapor interface using the neural network for the RPBE exchange-correlation functional with Grimme’s D3 corrections presented in Ref.12. This network was trained using configurations taken from the liquid, solid, and liquid/solid coexistence phases.19 As expected, the network could not recognize a relatively high fraction of configurations at the liquid/vapor interface, resulting in unphysically high vapor densities and large interfacial widths. A preliminary test with the SPC/F3 empirical model showed that using a training dataset of about 400 frames generated from equilibrium trajectories of a liquid/vapor interface (216 molecules) yields excellent convergence and physically sound results. We used these 400 frames to augment the dataset presented in Ref.12 taking care of perturbing the nuclear coordinates to enhance the sampling of the neighborhood of the free energy minima. The DFT energies and forces at the RPBE level were calculated for the whole training dataset using FHI-AIMS19 and D3 corrections were added. We then used this initial extension to train the neural network and generated a trajectory with a temperature ramp from 300 to 600 K over 1 ns of integration. Next, we took about 500 (unperturbed) frames from this trajectory to extend the training set further, performing in this way a self-consistent refinement step. We report additional methodological details in Sec.IV.

III. RESULTS

We ran molecular dynamics simulations in the canonical ensemble using the neural network potential for systems of 1024 water molecules in slab configurations within simulation boxes of size 3 × 3 × 10 nm, at 15 different temperature values ranging from 300 to 620 K. Each simulation started from a pre-equilibrated configuration of the SPC/F model at the corresponding temperature, using a timestep of 0.5 fs. After 100 ps, we observed no drift in the the potential energy of each trajectory, and we deemed them to have reached equilibrium. We saved configurations to disk during the subsequent 1 ns of trajectory at intervals of 1 ps for further analysis. Values of energy and pressure were dumped every 10 fs.

From the mass density profiles reported in Fig.2 one can appreciate the progressive broadening of the slab width, accompanied by the density increase in the vapor phase and corresponding decrease in the liquid one. In Fig.3 we report the density values of the 1 nm-wide regions located in the middle of the liquid and vapor slab. To extrapolate the critical point location, we performed the best fit of the simulation results to the expression proposed by Wilding22

\[ \rho(T) = \rho_c + a |T - T_c| \pm b(T - T_c)^β, \]

where \( \rho_c \) and \( T_c \) are the critical density and temperature, and the sign plus and minus applies to the liquid and vapor branch, respectively. We used the scaling exponent \( \beta \) as a fitting parameter, in addition to \( a \) and \( b \), obtaining as best fit estimates \( \rho_c = 310 \pm 3 \text{ kg/m}^3 \), \( T_c = 632 \pm 2 \text{ K} \), \( \beta = 0.27 \pm 0.01 \), \( a = 0.45 \pm 0.01 \) and \( b = 94 \pm 5 \).

In Fig.3 we report also the IAPWS95 curve23 which matches the experimental data, and the points from the only other estimate of the coexistence line of RPBE-D3 water we are aware of, taken from the work of Schienbein and Marx.20 Notice the steeper trend of Schienbein and Marx’s data, obtained via ab-initio Gibbs ensemble Monte Carlo of 128 water molecules, which is arguably a finite-size effect yielding an effective critical temperature \( T_c(L) \) that shifts toward higher values with
decreasing system (linear) size $L$ as:

$$T_c(L) - T_c \simeq \frac{1}{L^{1/\nu}}, \quad (2)$$

where $\nu$ is the critical exponent governing the scaling of the correlation length.

From the average values of the pressure tensor elements, $p_{ij}$, it is straightforward to compute the surface tension $\gamma$ using the mechanical route, as

$$\gamma = L_z \frac{p_{zz} - (p_{xx} + p_{yy})}{2}, \quad (3)$$

where $z$ identifies the direction normal to the macroscopic interface. In Fig. 4 we report the surface tension as a function of temperature, the result of the best fit to Eq. (1) from Ref. 27 and we compare them with the interpolated experimental values. Even though the present results do not match the experimental curve as well as the best empirical potential models like TIP4P-2005, the agreement is still very good, and superior to several other mainstream empirical models.

Density and surface tension values along the coexistence line are all reported in Tab. I, together with the estimated critical point.

With plenty of configurations at hand, it is now possible to investigate, in a statistically meaningful way, how water’s structural features are dependent on whether the molecules are located right at the interface, or below it. Molecular dynamics simulations with empirical potentials show that, at the liquid/vapor interface, water exhibits a much weaker order than van-der-Waals liquids, and many structural and dynamical properties reach their bulk value roughly after the second interfacial layer. In our investigation, we witnessed the same behavior, and here we only concentrate on the properties of the first molecular layer, detected on a per-frame basis using the Pytim analysis package as described in Sec. IV. For brevity, we will refer to “bulk properties” extracted from the trajectories presented in this work, as those properties computed from the molecules beyond the third surface layer, where the observables already converged to position-independent values.

We turn first to the distribution of bond lengths and angles, which are reported, for the 300 K system, in Fig. 5.
The bond length and angle distributions in the first layer are very similar to their bulk counterparts. However, one can notice a shift to lower values in both first-layer distributions, indicating that the molecules in the first layer are less stretched, with a difference of 0.3 pm and 0.3 deg for the bond length and angle, respectively. These differences become less pronounced when the temperature is raised.

Next, we characterize the orientation of the water molecules at the surface. To this purpose, we employ two angles. The first one, \( \theta \), is the angle that the molecular axis vector (pointing from the oxygen atom to the midpoint between the hydrogen ones) is making with the macroscopic surface normal (pointing from the liquid to the vapor phase). The second one, \( \phi \), is the angle that the molecule would need to rotate along its molecular axis to have the \( \text{H-H} \) vector aligned parallel to the surface, above 550K all orientations are accessible within an energy of 0.2 \( k_B T \) (see Supplementary Material).

The free energy plot shows a prevalence of molecular orientations (the minima) when the dipole moment is aligned parallel to the surface plane, or pointing slightly below it (\( \cos \theta \approx -0.25 \)) and when the \( \text{OH} \) bonds either laying in the surface plane (\( \phi \approx 0 \)) or pointing out of it (\( \phi \approx \pi/2 \)). This result agrees qualitatively with previous experimental findings and simulations with empirical potentials. The orientations with the dipole moment pointing toward the vapor (\( \cos \theta = 1 \)) or the liquid (\( \cos \theta = -1 \)) are less likely to be found than the parallel orientation, although all orientations are accessible within an energy of \( k_B T \), where \( k_B \) is the Boltzmann constant. When the temperature increases the free energy map becomes gradually flatter and although the preferential orientation of the dipole vector is always parallel to the surface, above 550K all orientations are accessible within an energy of 0.2 \( k_B T \) (see Supplementary Material).

Having access to the set of surface molecules, it is possible to build a local reference frame \((x, y, \zeta(x, y))\) on the corrugated interface and use it to compute the intrinsic density profile, as reported in Fig. 7. The usual density profile expresses the correlation between molecular positions in the system and the location of its center of mass. In contrast, the intrinsic density profile, \( \rho_I(z') \), represents the correlation between molecular positions and the local position of the interface,

\[
\rho_I(z') = \frac{m}{A} \left\langle \sum_i \delta(z' - z_i + \zeta(x_i, y_i)) \right\rangle, \quad (4)
\]
where $A$ is the simulation box cross-sectional area, $N$ is the number of molecules, $m$ their mass, and angular brackets stand for the canonical average. The molecules in the surface layer are located at $z' = 0$, giving rise to a delta peak, which is not shown in Fig. 7. By convention, positive values of $z'$ are located in the vapor phase. At relatively low temperature, where the vapor density is small, and the liquid’s cohesive strength is the highest, the structure of the local packing emerges with a clear peak and a small shoulder, similar to the results from simulations with empirical potentials. With increasing temperature, the local structure at the interface becomes less pronounced and almost disappears at 620 K. On the vapor side, one observes a similar behavior, with a peak of the vapor density next to the interface that vanishes at high temperature. One can expect condensing vapor at the interface because the molecules in the vapor phase feel both the attractive dispersion forces and residual dipolar interactions. Upon a closer look at the vapor phase region in a semi-logarithmic scale (Fig. 7, lower panel), we notice that the density, beyond the local maximum, decays exponentially toward the bulk vapor density like $\exp(-z/\xi)$, as it is expected for a dilute vapor next to a weakly attractive surface (the liquid phase) in the mean field approximation. The closer to the critical temperature, the more the profile tends toward homogeneity (with vapor density approaching the liquid one) except for an excluded volume region right at the interface that resembles the square-well one we would expect from a hard-sphere.

IV. METHODS

The neural network potential has been set up using the approach by Behler and Parrinello,\textsuperscript{13} using the implementation n2p2 of Singraber and coworkers.\textsuperscript{34} We used the same selection of symmetry functions and cutoff values reported in Ref. \textsuperscript{12}. The network was able to reproduce the forces acting on atoms within 1.5 meV/Å (1 standard deviation, $\approx 68\%$), 2.2 meV/Å (2 standard deviations, $\approx 95\%$) and 4 meV/Å (3 standard deviations, $\approx 99.7\%$), with a typical rms error of 3\% (see Supplementary Material). DFT calculations were performed using the FHI-AIMS package\textsuperscript{19} with second tier level of basis functions for oxygen and hydrogen atoms. Neural network molecular dynamics simulations were performed using LAMMPS (http://lammps.sandia.gov)\textsuperscript{35} linked to the neural network library of Singraber and coworkers (https://github.com/CompPhysVienna/n2p2).\textsuperscript{34} Each system consisted of 1024 water molecules simulated in the canonical ensemble using the Nosé–Hoover thermostat (damping constant 0.5 ps) and an integration timestep of 0.5 fs. The surface layer was determined using the MDAnalysis-based pymt package (https://github.com/Marcello-Sega/pymt)\textsuperscript{20} via a combination of the ITIM\textsuperscript{37} method (probe sphere radius 2 Å) and DBSCAN\textsuperscript{38} based prefiltering of the vapor phase\textsuperscript{29}, using the automatic determination of the density threshold for the clustering procedure.

V. CONCLUSIONS

We have performed what is arguably the most accurate calculation, to-date, of the liquid/vapor coexistence of water described by the RPBE exchange-correlation functional, supplemented by dispersion corrections. This result was possible thanks to the use of a neural network-based fit of the DFT potential energy surface. We reported the coexistence curve of the system, estimated the critical temperature of the model ($T_c = 632 \pm 2$ K), the surface tension curve as a function of temperature, and two order parameters, namely, the density profile and the orientation of water molecules in the surface layer. While the most refined empirical potential models are still superior in describing some aspects of the thermodynamics of water at interfaces, ab-initio calculations are becoming increasingly more accurate, albeit still very expensive computationally. Neural network-based approaches like the present one, alone or by exploiting promotion to the DFT level of choice,\textsuperscript{40} open up the possibility to explore with superior statistical accuracy systems that were, until now, almost exclusively in the realm of empirical potential-based simulations.

SUPPLEMENTARY MATERIAL

See supplementary material for the neural network training forces histogram and the orientation free energy maps.

ACKNOWLEDGMENTS

The computational results presented have been achieved using the Vienna Scientific Cluster (VSC).

We thank Andreas Singraber for providing help with the n2p2 package.

DATA AVAILABILITY

The datasets of the neural network training configurations with forces, the optimal weights, as well as the input files are openly available in Zenodo at http://doi.org/10.5281/zenodo.3944892, reference number 3944892.

REFERENCES

\textsuperscript{1}R. Kumar, J. R. Schmidt, and J. L. Skinner, J. Chem. Phys. \textbf{126}, 204107 (2007)
