Signatures of a liquid–liquid transition in an ab initio deep neural network model for water

Thomas E. Gartner III[a], Linfeng Zhang[b], Pablo M. Piaggi[a], Roberto Car[a,b,c,d], Athanassios Z. Panagiotopoulos[d,e,1], and Pablo G. Debenedetti[e,1]

[a]Department of Chemistry, Princeton University, Princeton, NJ 08544; [b]Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ 08544; [c]Department of Physics, Princeton University, Princeton, NJ 08544; [d]Princeton Institute for the Science and Technology of Materials, Princeton University, Princeton, NJ 08544; and [e]Department of Chemical and Biological Engineering, Princeton University, Princeton, NJ 08544

Contributed by Pablo G. Debenedetti, August 27, 2020 (sent for review July 21, 2020; reviewed by Christoph Dellago and Carlos Vega)

The possible existence of a metastable liquid–liquid transition (LLT) and a corresponding liquid–liquid critical point (LLCP) in supercooled liquid water remains a topic of much debate. An LLT has been rigorously proved in three empirically parametrized molecular models of water, and evidence consistent with an LLT has been reported for several other such models. In contrast, experimental proof of this phenomenon has been elusive due to rapid ice nucleation under deeply supercooled conditions. In this work, we combined density functional theory (DFT), machine learning, and molecular simulations to shed additional light on the possible existence of an LLT in water. We trained a deep neural network (DNN) model to represent the ab initio potential energy surface of water from DFT calculations using the Strongly Constrained and Appropriately Normed (SCAN) functional. We then used advanced sampling simulations in the multithermal–multibaric ensemble to efficiently explore the thermophysical properties of the DNN model. The simulation results are consistent with the existence of an LLCP, although they do not constitute a rigorous proof thereof. We fit the simulation data to a two-state equation of state to provide an estimate of the LLCP's location. These combined results—obtained from a purely first-principles approach with no empirical parameters—are strongly suggestive of the existence of an LLT, bolstering the hypothesis that water can separate into two distinct liquid forms.

water | liquid–liquid transition | molecular simulations | machine learning

Water is ubiquitous in both the physical sciences and everyday life; as such, there is a very large body of literature characterizing its physical properties (1). This research has revealed that water exhibits a striking set of anomalies relative to most other liquids (2), including the fact that the liquid freezes into a less-dense solid phase at ambient and moderate pressures; liquid water's observed density maximum; and the sharp increases in compressibility, heat capacity, and the magnitude of the thermal expansion coefficient of liquid water upon cooling (2–4). Several competing theories exist that provide thermodynamically consistent descriptions of these anomalies (2, 4), including the liquid–liquid transition (LLT) hypothesis (5), the stability-limit conjecture (6), and the singularity-free scenario (7). Much of the behavior discussed in these theories occurs under supercooled conditions (i.e., below the melting temperature of ice); experimental data in this region are challenging to obtain due to rapid crystallization (4, 8), which renders the liquid state short lived (9). Because of these challenges, the available experimental work has been fruitfully augmented by a large number of computational and theoretical studies aimed at illuminating the key phenomena and weighing the relative merits of these scenarios (2, 4, 8, 10–12). For example, recent simulation work illustrated that the stability-limit conjecture likely is not compatible with the shape of the thermodynamic extrema exhibited by the popular TIP4P/2005 water model (13) but that the LLT and singularity-free scenarios could still be potential candidates (14).

According to the LLT hypothesis (5, 10), under strongly supercooled conditions and high pressures, water undergoes a first-order phase separation into high-density liquid (HDL) and low-density liquid (LDL) states, which are metastable with respect to crystalline ice. In the pressure–temperature plane, this line of liquid–liquid coexistence terminates in a metastable liquid–liquid critical point (LLCP). Emanating from the LLCP toward higher temperatures and lower pressures is the so-called Widom line (11, 15), which is a locus of maximum correlation length in the fluid; maxima in other properties such as isothermal compressibility and constant-pressure heat capacity converge with the Widom line near the LLCP. There is some recent experimental evidence for a maximum in the isothermal compressibility upon cooling liquid water at ambient pressure (16). This result has elicited debate (17, 18), and analogous experiments at higher pressure have not been done to date.

In simulations, small system sizes serve to suppress homogenous nucleation of ice, and the timescales for nucleation and structural relaxation tend to be well separated in molecular models of water. These factors enable computational study of liquid water in the strongly supercooled regime across the full range of relevant temperatures and pressures. Rigorous evidence for the LLT has been obtained for three classical molecular models for water with simple pairwise interaction potentials: ST2 (19–24), TIP4P/2005 (13), and TIP4P/Ice (25, 26). There are also numerous studies presenting evidence consistent with an LLT but without a rigorous

Significance

Water is central across much of the physical and biological sciences and exhibits physical properties that are qualitatively distinct from those of most other liquids. Understanding the microscopic basis of water's peculiar properties remains an active area of research. One intriguing hypothesis is that liquid water can separate into metastable high- and low-density liquid phases at low temperatures and high pressures, and the existence of this liquid–liquid transition could explain many of water's anomalous properties. We used state-of-the-art approaches in computational quantum chemistry, statistical mechanics, and machine learning and obtained evidence consistent with a liquid–liquid transition, supporting the argument for the existence of this phenomenon in real water.

Author contributions: T.E.G., A.Z.P., and P.G.D. designed research; T.E.G., L.Z., and P.M.P. performed research; T.E.G., L.Z., and P.M.P. analyzed data; and T.E.G., L.Z., P.M.P., R.C., A.Z.P., and P.G.D. wrote the paper.

Reviewers: C.D., University of Vienna; and C.V., Universidad Complutense de Madrid.

The authors declare no competing interest.

[1]To whom correspondence may be addressed. Email: azp@princeton.edu or pdebene@princeton.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2015440117/-/DCSupplemental.

First published October 2, 2020.

demonstration thereof (10, 27–29). However, it should be borne in mind that empirically parameterized models do not provide definitive proof of the presence or absence of an LLT in real water (10). In fact, other popular water models, such as the coarse-grained mW model (30, 31) and SPC/E (32–34), do not exhibit clear evidence of an LLT. There are also several more sophisticated models of water that explicitly represent many-body and/or polarizability effects and have been used to probe the supercooled liquid state, such as E3B3 (11, 27, 35), WAIL (28, 36), iAMOEBA (37, 38), RexPoN (39–41), MB-pol (42), and others (43). Some of these models are parameterized in order to reproduce experimental data (35), while others are fit to theoretical calculations (36, 41–43). Mixed approaches are also possible as is the case of the iAMOEBA model (38). These advanced models and methods have revealed significant insights about the structure and properties of liquid water (39, 40, 42) and in some cases, have even been suggestive of the existence of an LLT (10, 27, 28). However, in general these models are too computationally expensive to rigorously identify the LLT with current computational resources, especially in light of sluggish structural relaxation in the vicinity of the critical point.

Ab initio methods could also present an opportunity to evaluate many-body effects in supercooled liquid water. For instance, one could use ab initio molecular dynamics (AIMD), in which the forces that drive the simulation are obtained on the fly from electronic structure calculations (44). These types of simulations do not rely on empirically derived parameters or training data and thus, could provide truly predictive first-principles evidence of an LLT. Specifically in the case of water, the Strongly Constrained and Appropriately Normed (SCAN) (45) exchange correlation (XC) functional has been shown to be especially accurate among nonempirical XC functionals in predicting the properties of water (e.g., the increase in density upon melting ice 1-h) due to its accurate description of the balance of covalent, intermediate-range van der Waals and hydrogen bonding interactions (46). Thus, searching for evidence of an LLT in a SCAN-based description of water is appealing but hitherto impossible due to the timescales currently accessible with traditional AIMD (a few hundred picoseconds at most).

Augmenting traditional molecular simulations with machine learning provides a potential avenue by which these challenges in the computational study of supercooled liquid water might be addressed (47). In one class of machine learning-based simulation approaches, a neural network is trained to represent the potential energy surface (PES) of a system obtained from ab initio calculations (48–53). The trained neural network model translates the local environment of an atom to ab initio-level atomic energies and forces at a fraction of the computational cost, which can then be used to drive a molecular dynamics (MD) simulation. This scheme allows for the evolution of systems according to the underlying many-body ab initio PES but at simulation speeds orders of magnitude faster than (for example) density functional theory (DFT)-based AIMD (44). In the present context, such an approach now enables first-principles computational study of water commensurate with the system sizes and timescales necessary for studies of phase behavior in the deeply supercooled regime.

In this work, we used the Deep Potential Generator (DP-GEN) active learning procedure (54) to generate a deep neural network (DNN) model of water. We prepared the training data with DFT calculations using the SCAN functional (45), which has recently shown promise in accurately predicting the thermophysical properties of water (46). Via isothermal–isobaric Deep Potential Molecular Dynamics (DPMD) simulations (51, 55), we confirmed that the DNN model qualitatively reproduces key trends in the structural, bulk, and dynamic properties of supercooled liquid water. We then used DPMD simulations in the multithermal–multibaric (MTMB) ensemble (56–58) to efficiently obtain the thermophysical properties (e.g., density, internal energy, isothermal compressibility, heat capacity) of this model over a wide range of temperatures and pressures. The simulation data were consistent with the existence of an LLCP. We then used a phenomenological two-state equation of state (TSEOS) model (59–63), which represents water as an interconvertible mixture of high- and low-density structures, to describe the MTMB simulation data. The TSEOS closely matched the data and provided an estimate for the LLCP location in this DNN model. Overall, these results are consistent with the existence of an LLT in water, providing ab initio evidence for this phenomenon for a model free of empirical parameterizations.

Results and Discussion

DPMD Simulations with DNN Model. We generated our DNN model using the DP-GEN active learning approach; refs. 54 and 64 have details. We performed DFT calculations using the Vienna ab initio simulation package (65, 66) within the SCAN functional approximation (45) on a total of 19,006 structures identified by the DP-GEN procedure, which covers temperature and pressure ranges of 0 to 600 K and 0 to 50 GPa, respectively. We then used the DeePMD-kit package (67) to train our DNN model to reproduce the ab initio-level atomic energies and forces from this DFT dataset. *SI Appendix* contains all details of the model training and validation procedures (*SI Appendix*, Fig. S1), as well as a comparison of the structure, thermophysical properties, and dynamics of the trained DNN model with available experimental data for supercooled liquid water (*SI Appendix*, Fig. S2). We also computed the ambient pressure ice 1-h melting point with this DNN model, which was in the vicinity of $T = 312.5$ K (*SI Appendix*, Fig. S4). The qualitative temperature and pressure dependence of water's properties was well captured by the DNN model, including the density maximum and increase in isothermal compressibility upon cooling and the increase in diffusivity in supercooled liquid water at increased pressure. We note that our classical simulations with the DNN model predicted key properties of water (e.g., liquid water structure, ambient pressure density maximum, ice 1-h melting point) at temperatures ∼47-K higher than the equivalent temperature in experiments, which we attribute to a combination of nuclear quantum effects (68) and limitations in accuracy of the underlying SCAN functional (52, 69) (we discuss this in detail in *SI Appendix*). Acknowledging these factors, our DNN model nevertheless reproduced, at an ab initio level, qualitative trends in the available experimental data for structural, bulk, and dynamic properties of supercooled liquid water. Given this correspondence, we now turn to examine whether this DNN model demonstrates evidence for an LLCP.

Due to the DNN model's high computational expense relative to classical atomistic models (albeit still much faster than AIMD at the DFT level), comprehensively exploring the thermophysical properties of this model in the deeply supercooled regime is not feasible using standard (unconstrained) MD simulations. To resolve this issue, we used advanced sampling simulations in the MTMB ensemble (58), which enabled us to sample a wide range of temperatures and pressures with a single simulation. Briefly, MTMB simulations bias the potential (i.e., configurational) energy and volume of a simulation to achieve a desired probability distribution that spans the full range that would be sampled in multiple unbiased simulations at different temperatures and pressures of interest. Then, the results from the biased simulation are reweighted as described in ref. 58 to calculate desired observables in the unbiased probability distribution at a given (T, P). Herein, we used two MTMB simulations to sample the range 240 K ≤ T ≤ 330 K and 1 bar ≤ P ≤ 3,000 bar [*SI Appendix*, Fig. S5 has an illustration of how this range of (T, P) corresponds with the experimental water phase diagram]. We obtained the thermophysical properties of this DNN model at ∼4,700 state points [i.e., individual (T, P)] within this range from just two simulations, at a total computational cost comparable with only ∼100 individual

isothermal–isobaric simulations (*Materials and Methods* and *SI Appendix* have details). Additionally, we thus avoided having to run many separate slow to equilibrate simulations at the sluggish low-temperature state points of interest.

In Fig. 1*A*, we plot bulk mass density (ρ) along isobars obtained from MTMB simulations (squares and solid lines) and standard isothermal–isobaric (NPT) simulations (circles). The MTMB and NPT simulations agreed closely over the full range of T and P studied, and the uncertainty from the MTMB simulations (95% CIs obtained from a bootstrap analysis) (*Materials and Methods*) was significantly smaller than from NPT simulations. The isobars showed a strongly positive slope at low temperatures and intermediate pressures, reminiscent of the (ρ, T, P) behavior of other water models definitively shown to exhibit an LLCP [e.g., ST2 (59), TIP4P/2005 (60)]. Fig. 1*B* shows the isothermal compressibility (κ_T), calculated via numerical evaluation of $\kappa_T = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T$ from the MTMB results. The κ_T along isobars exhibited a maximum upon cooling, and the height of this maximum increased sharply with increasing pressure while shifting toward progressively lower temperatures, strongly suggestive of the existence of a Widom line terminating in an LLCP. In order to estimate the location of the LLCP predicted via this DNN model (if it exists), we fit the simulation data to a TSEOS.

Estimated LLCP Location and Thermodynamic Anomalies. In order to explore the compatibility of the DNN simulation results with the existence of an LLCP, we fit the simulation data with a phenomenological TSEOS model (60). The TSEOS represents water as a binary mixture of interconvertible (high-entropy, high-density and low-entropy, low-density) local configurations and constructs the Gibbs free energy of the system as a function of (T, P) and the equilibrium low-density fraction (x_e) at any given state point. Then, relevant thermodynamic quantities, such as density, internal energy, heat capacity, isothermal compressibility, etc., can be calculated via the appropriate partial derivatives of the free energy. The TSEOS approach has been successfully shown to reproduce the thermodynamics of many popular models of water [including mW (70), ST2 (59), TIP4P/2005 (60, 61)] and also successfully describes a large set of available experimental data for liquid water (62, 63). Other versions of the two-state approach also successfully capture the anomalies of water and other tetrahedral liquids (71–73). Fig. 2 shows the predictions of our TSEOS model for the density and internal energy (configurational + kinetic energy), compared with the MTMB simulations. The correspondence between the simulation results and TSEOS predictions was excellent over the majority of the state points, with appreciable deviations only at the limits of the (T, P) range of interest. The liquid–liquid critical parameters obtained from the TSEOS were T_c = 224 ± 3 K, P_c = 2,687 ± 68 bar, and ρ_c = 1,072 ± 16 kg/m^3. Uncertainties in the critical parameters were obtained by independently fitting the TSEOS to 200 block-bootstrap samples generated from the MTMB data, as described in *Materials and Methods*. Interestingly, the TSEOS fit to this DNN model predicts a relatively mild difference in the densities at coexistence between the LDL and HDL phases. For example, in this work ρ_{HDL} = 1,102 kg/m^3 and ρ_{LDL} = 1,047 kg/m^3 at a reduced temperature T/T_c = 0.965, whereas for the classical ST2 model, ρ_{HDL} = 1,150 kg/m^3 and ρ_{LDL} = 900 kg/m^3 at the same reduced temperature (19). This relatively subtle difference between the densities of the LDL and HDL states, combined with the small free energy barrier between them (19), may contribute to the difficulty in rigorously establishing the LLT in both simulations and experiments.

Recalling *SI Appendix*, Fig. S2, in which we applied a −47 K temperature shift to the simulation data to achieve correspondence between the DNN and experimental results, the T_c near 224 K in this DNN model might correspond to an experimental temperature near 177 K, which is in the range of T_c values recently obtained for the popular and realistic classical TIP4P/2005 and TIP4P/Ice models (25). However, as seen in some other studies with advanced water models, the less sharp increase in compressibility at 1 bar compared with the experimental results (*SI Appendix*, Fig. S2*C*) suggests that the P_c predicted by this DNN model is somewhat higher than might be expected in experiments (11). We also note recent work that suggests an LLCP in AIMD simulations of water with the SCAN functional at small negative pressures (74); that work reported a van der Waals loop in a series of isothermal-isochoric simulations at only six state points and for a system much smaller than those considered here. Furthermore, their simulation timescales are two to four orders of magnitude smaller than those used in this work, which is likely not enough to reliably equilibrate the system in the deeply supercooled regime. Therefore, we consider our estimate for the location of the LLCP in this model to be within the range of estimates obtained from other models and approaches (10).

Fig. 3 shows water's thermodynamic anomalies in the pressure–temperature plane from the MTMB simulations and the TSEOS. The TSEOS provides a good representation of the loci of maximum density, isothermal compressibility, and constant-pressure

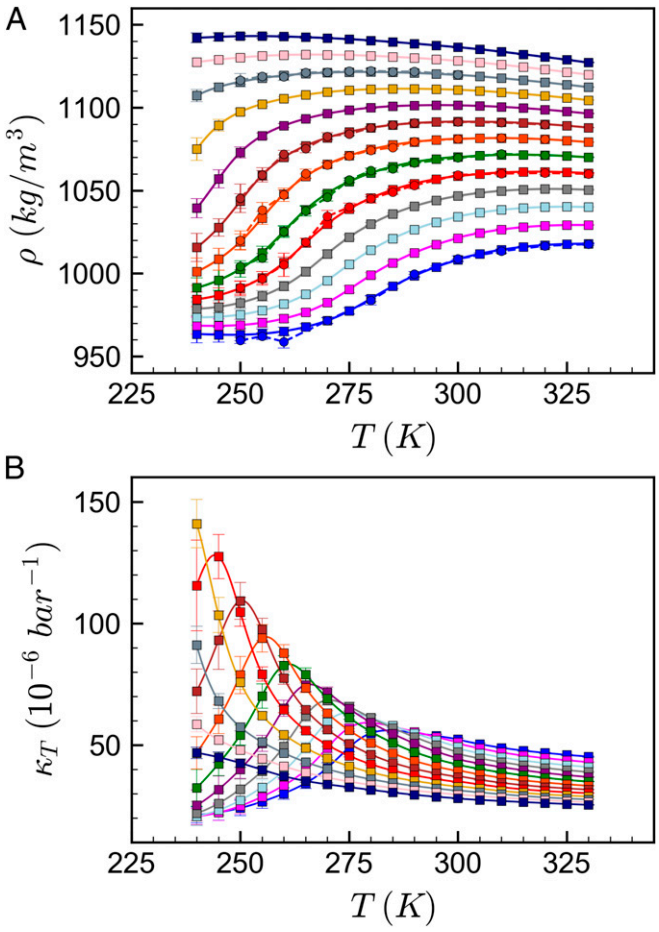

Fig. 1. MTMB and NPT simulation results along isobars. (*A*) Mass density, ρ, and (*B*) isothermal compressibility, κ_T, from NPT (circles and dashed lines) and MTMB (squares and solid lines) simulations using the DNN model. In both plots, the blue isobar is P = 1 bar, and the magenta to dark blue isobars range from P = 250 bar to P = 3,000 bar in steps of 250 bar. Note that the solid lines are MTMB results reweighted across the entire temperature range; squares are only included every 5 K for visual clarity.

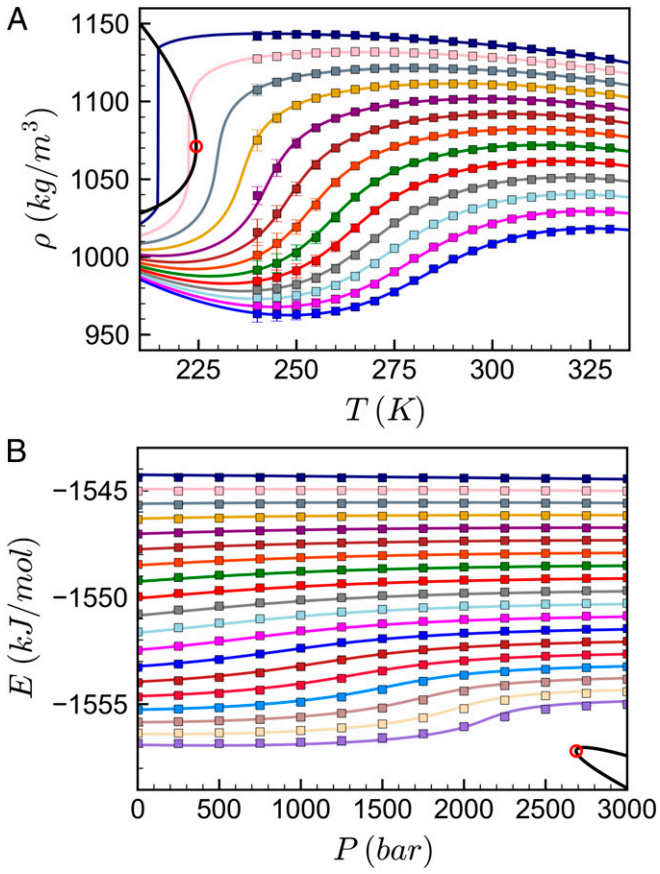

Fig. 2. Fits of the TSEOS to the MTMB simulation data. (*A*) Mass density, ρ, and (*B*) internal energy, *E*, from MTMB simulations (squares) and TSEOS predictions (thick solid lines). In *A*, colors denote different isobars, with blue representing $P = 1$ bar and the magenta to dark blue isobars ranging from $P = 250$ bar to $P = 3{,}000$ bar in steps of 250 bar. In *B*, colors denote different isotherms, ranging from $T = 240$ K (lavender) to $T = 330$ K (dark blue) in steps of 5 K. The liquid–liquid binodal and LLCP location from the TSEOS are denoted by the thick black line and red circle, respectively.

heat capacity calculated from our simulations with the DNN model. These data exhibit the expected trends, with the loci of maximum heat capacity and compressibility converging along the Widom line near the LLCP, as observed in other classical models for water (60, 61). The LLT transition line extends from the critical point to lower temperatures and higher pressures with slope and curvature given by the equality of Gibbs energy between the high- and low-density states (*Materials and Methods*). Overall, these results suggest that the thermophysical properties of this SCAN-based DNN model for water can be well described by the two-state picture. While as discussed in detail in previous TSEOS work (60), this analysis does not constitute definitive proof of an LLT or LLCP in the SCAN PES, the close agreement between simulation and TSEOS results and the qualitative shape of the density and compressibility isobars together provide strong evidence consistent with the existence of an LLCP. Furthermore, recent simulation and experimental scattering data provide evidence in support of the two-structure view of liquid water (75).

Conclusions

We have combined state-of-the-art approaches in first-principles quantum chemistry, molecular simulation, statistical mechanics, and machine learning to explore the possibility of an LLT in the SCAN ab initio PES of water. We developed a DNN potential free of empirical parameters that captures the key features of experimental supercooled liquid water, at a computational cost orders of magnitude smaller than other comparable advanced models. We then connected this DNN model with MTMB sampling techniques to further extend the regimes of timescales and length scales accessible. Using this platform, we probed the properties of the SCAN description of water over a wide swath of the supercooled liquid water phase diagram. Our simulation results for density and compressibility along isobars were strongly suggestive of an LLCP, and we fit the simulation data to a TSEOS model. The trends in density and internal energy were compatible with the two-state picture, and the estimates for the critical parameters were within the ranges of other recent computational studies. Overall, the results reported herein strengthen the interpretation of supercooled liquid water as possessing two metastable liquid forms, providing first-principles evidence for this phenomenon free from empirical extrapolation.

In the future, we envision exploring in more detail the low-temperature phase behavior of this DNN model, including extending the lines of maximum density and ice 1-h melting point to negative pressures, as well as probing the effects of the choice of XC functional and the treatment of quantum fluctuations on water's LLT. We note recent work that compared classical and path-integral simulations with a similar SCAN-based neural network model of water and found that the limitations of the SCAN XC functional may be the key contributor to discrepancies between simulation and experiment (69). It would be instructive to perform similar tests on the effect of nuclear quantum effects on the signatures of the LLT shown herein, perhaps by taking inspiration from prior work in which rigid models were used to estimate the effect of nuclear quantum effects with a more modest computational expense (76). One could also consider generating even more accurate DNN models from accurate but expensive quantum chemistry methods. Lastly, we envision using free energy techniques at state points predicted by the TSEOS in order to definitely establish liquid–liquid coexistence from first principles, further pushing the boundaries of problems accessible through this approach.

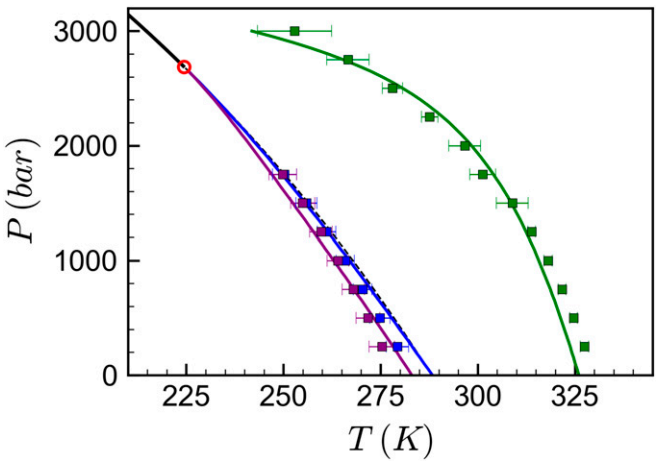

Fig. 3. Thermodynamic anomalies from MTMB simulations and the TSEOS. The liquid–liquid coexistence line, the LLCP, and the Widom line predicted by the TSEOS (*Materials and Methods*) are plotted with the thick black line, red circle, and thin dashed black line, respectively. Green, blue, and purple symbols and lines denote the loci of maximum density, isothermal compressibility, and constant-pressure heat capacity, respectively, with squares representing MTMB simulation results and lines representing TSEOS predictions. Note that the ambient pressure ice 1-h melting temperature for this model is near $T = 312.5$ K (*SI Appendix*, Fig. S4).

Table 1. Optimized TSEOS parameters

Parameter	Value
T_c, K	224
P_c, bar	2,687
λ	1.660
a	0.3603
b	−0.2694
w_0	0.1268
c_{00}	−831.7
c_{01}	2.091
c_{02}	−0.01794
c_{11}	0.2733
c_{20}	−6.066
c_{12}	−0.1736
c_{21}	−0.1816
c_{30}	−0.8471
c_{22}	0.4378
c_{31}	0.3749
c_{40}	1.968
c_{23}	0.1490
c_{32}	−0.4847
c_{33}	−0.1868

Materials and Methods

Simulation Methods. We performed MD simulations with our SCAN-based DNN water model using the DPMD-augmented LAMMPS (77) software package [DeepMD-kit (64, 67) version 1.0 and LAMMPS version 9 Aug 2019]. We performed NPT simulations with 512 water molecules over a temperature range 250 K $\leq T \leq$ 330 K and a pressure range 1 bar $\leq P \leq$ 2,500 bar. We maintained temperature and pressure using the LAMMPS Nosé–Hoover-type thermostat and barostat, with relaxation times 50 and 500 fs, respectively. We equilibrated all NPT simulations at least two times longer than the structural relaxation time determined by the decay of the peak value of the self-intermediate scattering function $F(k^*,t)$ to a value of 1/e; most simulations were equilibrated 10 to 500 times longer than that. We computed thermodynamic averages over a sampling period of 4 to 32 ns depending on the state point. Because configurational properties obtained from classical simulations do not depend on the mass of the atoms, we used a hydrogen mass of 2 atomic mass units in the time integration to enable a time step size of 0.5 fs but used water's average molar mass of 18.015 g/mol for analysis. In the NPT simulations, we calculated the isothermal compressibility via $\kappa_T = \left(\langle V^2\rangle - \langle V\rangle^2\right)/k_B T\langle V\rangle$. Error bars in all quantities reported from the unconstrained simulations represent 95% CIs obtained from 10-block averages.

We performed the MTMB simulations (58) by patching the DPMD-LAMMPS implementation with the PLUMED software (78) using the Variationally Enhanced Sampling (VES) module (79). The MTMB approach is built upon the VES framework (80), in which a bias potential is added to the PES to achieve a desired marginal probability distribution of a chosen set of collective variables. In MTMB simulations, the potential energy (i.e., configurational energy) and volume are biased in order to sample a range of temperatures and pressures with a single simulation. In an initial stage, one first optimizes the functional form of the bias potential to achieve the MTMB probability distribution in potential energy and volume, and then in a sampling stage, one samples the MTMB distribution with a stationary (or nearly stationary) bias potential. Then, one can reweight the biased results to obtain a given observable $O(\boldsymbol{R}, V)$ that depends on the atomic coordinates $\boldsymbol{R}$ and the system volume V at a desired temperature (T') and pressure (P') via the following formula:

$$O(\boldsymbol{R}, V)_{T',P'} = \frac{\langle O(\boldsymbol{R}, V) w(\boldsymbol{R}, V)\rangle_{T,P,V}}{\langle w(\boldsymbol{R}, V)\rangle_{T,P,V}}, \quad [1]$$

in which T and P are the nominal temperature and pressure of the MTMB simulation, respectively, and $w(\boldsymbol{R}, V)$ is a properly chosen weight for that configuration (*SI Appendix*). We refer the reader to ref. 58 for a full discussion of the MTMB method.

We performed two independent MTMB simulations with 192 water molecules to calculate the properties of the DNN model. We selected this smaller system size due to the large computational expense of the DNN simulations. In both MTMB simulations, we targeted the PLUMED TD_MULTITHERMAL_MULTIBARIC probability distribution in both energy and volume to a free energy threshold of 1 k_BT with a 10th-order Legendre polynomial bias potential in each dimension. We ran the first simulation at a nominal temperature and pressure of T = 270 K and P = 1,250 bar, respectively, and a second at T = 290 K and P = 1,250 bar, respectively. In the first simulation, the polynomials were defined in the potential energy range −1,567.71 kJ/mol $\leq U \leq$ −1,557.29 kJ/mol and volume range 4.875 nm^3 $\leq V \leq$ 6.375 nm^3, and in the second simulation, the ranges were −1,567.71 kJ/mol $\leq U \leq$ −1,552.08 kJ/mol and 4.875 nm^3 $\leq V \leq$ 6.375 nm^3 to cover the desired set of temperatures and pressures. We used the reweighted results of the first simulation to calculate all properties $T \leq$ 300 K and the second simulation to calculate properties in the temperature range 300 K $< T \leq$ 330 K, and both simulations spanned the pressure range 1 bar $\leq P \leq$ 3,000 bar. For each MTMB simulation, we optimized the bias potential coefficients for ∼40 ns, followed by 150 ns of sampling with frozen coefficients for simulation 1 and 100 ns of sampling for simulation 2, with data saved every 0.25 ps. We used four independent walkers for each simulation to improve the efficiency of the sampling and the convergence of the bias potential. *SI Appendix*, Figs. S6 and S7 show the evolution of the MTMB bias potential coefficients, the form of the bias potential and statistical weights $w(\boldsymbol{R}, V)$ used in this work, and the potential energy-volume values visited by the MTMB and NPT simulations. Error bars for plots related to the MTMB simulations are 95% CIs obtained from sampling 200 block-bootstrap samples resampled from 40 blocks generated from the MTMB results.

In the lower-temperature simulation (T = 270 K), we observed that the system periodically entered a "sluggish" or "glassy" state, in which the molecules did not translate and the energy and volume of the system did not traverse the desired range (*SI Appendix*, Figs. S8–S10). Given that these configurations were highly correlated due to the lack of molecular motion, we excluded them from our MTMB reweighting. *SI Appendix*, Fig. S8 shows a quantification of the average molecular motion throughout the simulations, and *SI Appendix*, Figs. S9 and S10 show an illustration of how the sluggish state affected the energy and volume results. However, whether we included or excluded these points in the MTMB reweighting did not significantly affect our obtained results; including the sluggish state resulted in critical parameters less than 5% different from those reported herein (i.e., with the sluggish points excluded). We note that this sluggish state was not observed in the NPT simulations, so we conjecture it is a consequence of the large temperature range sampled via the MTMB algorithm.

Two-State Equation of State. Following recent work, which used a TSEOS to describe the properties of the TIP4P/2005 (60, 61) and ST2 (59) models of water, we represent the molar Gibbs free energy of the system at a given state point by assuming that water is a mixture of two interconvertible states—high-density form A and low-density form B—with the fraction of molecules assuming local state B given by x:

$$G(T,P) = G^A + xG^{BA} + G^{mix}. \quad [2]$$

In Eq. **2**, G^A is the Gibbs energy of pure structure A at a given (T, P), represented as a polynomial expansion of temperature and pressure:

$$G^A(T,P) = RT_c \sum_{m,n} c_{mn} \Delta\hat{T}^m \Delta\hat{P}^n. \quad [3]$$

In Eq. **3** and the following expressions, the temperature and pressure are nondimensionalized relative to the critical temperature T_c and critical pressure P_c by $\Delta\hat{T} = \frac{T-T_c}{T_c}$ and $\Delta\hat{P} = \frac{P-P_c}{P_c}$. G^{BA} is the difference in free energy between structure B and structure A, and we represent it with a simple nonlinear form:

$$G^{BA}(T,P) = RT\lambda\left(\Delta\hat{T} + a\Delta\hat{P} + b\Delta\hat{T}\Delta\hat{P}\right). \quad [4]$$

In the expression for G^{BA}, λ represents the entropy difference between states A and B. The set of temperatures and pressures that yield $G^{BA} = 0$ defines the LLT (for $T < T_c$) and Widom (for $T > T_c$) lines, in which a sets the slope of the LLT/Widom lines at the LLCP and b determines their curvature. The free energy of mixing between state A and state B is represented via a sum of ideal and nonideal terms:

$$G^{mix}(T,P) = RT\left[x\ln x + (1-x)\ln(1-x) + \frac{2 + w_0\Delta\hat{P}}{T/T_c}x(1-x)\right]. \quad [5]$$

The temperature and pressure dependence of the term $\frac{2+w_0\Delta\hat{P}}{T/T_c}$ is chosen to

reproduce a "regular solution"-type nonideal enthalpy of mixing, which was previously shown to successfully capture the two-state thermodynamics of the ST2 and TIP4P/2005 models (59, 60).

In the above expressions, T_c, P_c, and the coefficients $c_{m,n}$, λ, a, b, and w_0 are fitting parameters that we optimized to match the predictions of the TSEOS to the MTMB simulation data. Our fitting procedure was as follows: first, we obtained the equilibrium low-density fraction (x_e) at a given (T, P) by solving $\left(\frac{\partial G}{\partial x}\right)_{T,P} = 0$ numerically, followed by the evaluation of the molar density and internal energy as $\rho_m = \left[\left(\frac{\partial G}{\partial P}\right)_T\right]^{-1}$ and $E = G + TS - PV = G - T\left(\frac{\partial G}{\partial T}\right)_P - P\left(\frac{\partial G}{\partial P}\right)_T$ via numerical differentiation of Eq. **2**. We repeated this process for the desired range of (T, P) and then optimized the parameters of the TSEOS to minimize the total sum of squared error between the TSEOS predictions and MTMB simulation results. The optimized parameters are listed in Table 1. We obtained the critical density ρ_c by finding the temperature maximum of the binodal line defined by $G^{BA} = 0$ for $T \leq T_c$. In order to provide an error estimate for the critical parameters, we fit an independent set of parameters to each of the 200 block-bootstrap samples of the MTMB results to obtain 95% CIs.

Data Availability. All data related to this work, including software installation instructions, DP-GEN, LAMMPS, and PLUMED input files, trained DNN model, raw simulation trajectory data, and analysis scripts are publicly available for download at https://doi.org/10.34770/45m3-am91 (81).

ACKNOWLEDGMENTS. We acknowledge H. Wang for assistance with DNN model development. T.E.G., L.Z., R.C., and A.Z.P. were supported by the "Chemistry in Solution and at Interfaces" Center funded by US Department of Energy Award DE-SC001934. P.M.P. was supported by an Early Postdoc.Mobility fellowship from the Swiss National Science Foundation. Calculations were performed on computational resources managed and supported by Princeton Research Computing, a consortium of groups including the Princeton Institute for Computational Science and Engineering (PICSciE) and the Office of Information Technology's High Performance Computing Center and Visualization Laboratory at Princeton University.

1. D. S. Eisenberg, W. Kauzmann, *The Structure and Properties of Water*, (Clarendon Press, 2006).
2. P. G. Debenedetti, Supercooled and glassy water. *J. Phys. Condens. Matter* **15**, R1669–R1726 (2003).
3. R. J. Speedy, C. A. Angell, Isothermal compressibility of supercooled water and evidence for a thermodynamic singularity at −45 degrees C. *J. Chem. Phys.* **65**, 851–858 (1976).
4. O. Mishima, H. E. Stanley, The relationship between liquid, supercooled and glassy water. *Nature* **396**, 329–335 (1998).
5. P. H. Poole, F. Sciortino, U. Essmann, H. E. Stanley, Phase-behavior of metastable water. *Nature* **360**, 324–328 (1992).
6. R. J. Speedy, Stability-limit conjecture—an interpretation of the properties of water. *J. Phys. Chem.* **86**, 982–991 (1982).
7. S. Sastry, P. G. Debenedetti, F. Sciortino, H. E. Stanley, Singularity-free interpretation of the thermodynamics of supercooled water. *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics* **53**, 6144–6154 (1996).
8. P. H. Handle, T. Loerting, F. Sciortino, Supercooled and glassy water: Metastable liquid(s), amorphous solid(s), and a no-man's land. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 13336–13344 (2017).
9. J. A. Sellberg *et al.*, Ultrafast X-ray probing of water structure below the homogeneous ice nucleation temperature. *Nature* **510**, 381–384 (2014).
10. J. C. Palmer, P. H. Poole, F. Sciortino, P. G. Debenedetti, Advances in computational studies of the liquid-liquid transition in water and water-like models. *Chem. Rev.* **118**, 9129–9151 (2018).
11. N. J. Hestand, J. L. Skinner, Perspective: Crossing the Widom line in no man's land: Experiments, simulations, and the location of the liquid-liquid critical point in supercooled water. *J. Chem. Phys.* **149**, 140901 (2018).
12. P. Gallo *et al.*, Water: A tale of two liquids. *Chem. Rev.* **116**, 7463–7500 (2016).
13. J. L. F. Abascal, C. Vega, A general purpose model for the condensed phases of water: TIP4P/2005. *J. Chem. Phys.* **123**, 234505 (2005).
14. M. A. González, C. Valeriani, F. Caupin, J. L. F. Abascal, A comprehensive scenario of the thermodynamic anomalies of water using the TIP4P/2005 model. *J. Chem. Phys.* **145**, 054505 (2016).
15. L. Xu *et al.*, Relation between the Widom line and the dynamic crossover in systems with a liquid-liquid phase transition. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 16558–16562 (2005).
16. K. H. Kim *et al.*, Maxima in the thermodynamic response and correlation functions of deeply supercooled water. *Science* **358**, 1589–1593 (2017).
17. F. Caupin *et al.*, Comment on "Maxima in the thermodynamic response and correlation functions of deeply supercooled water.". *Science* **360**, eaat1634 (2018).
18. K. H. Kim *et al.*, Response to comment on "Maxima in the thermodynamic response and correlation functions of deeply supercooled water.". *Science* **360**, eaat1729 (2018).
19. J. C. Palmer *et al.*, Metastable liquid-liquid transition in a molecular model of water. *Nature* **510**, 385–388 (2014).
20. F. Sciortino, I. Saika-Voivod, P. H. Poole, Study of the ST2 model of water close to the liquid-liquid critical point. *Phys. Chem. Chem. Phys.* **13**, 19759–19764 (2011).
21. Y. Liu, J. C. Palmer, A. Z. Panagiotopoulos, P. G. Debenedetti, Liquid-liquid transition in ST2 water. *J. Chem. Phys.* **137**, 214505 (2012).
22. J. C. Palmer, R. Car, P. G. Debenedetti, The liquid-liquid transition in supercooled ST2 water: A comparison between umbrella sampling and well-tempered metadynamics. *Faraday Discuss.* **167**, 77–94 (2013).
23. T. A. Kesselring, G. Franzese, S. V. Buldyrev, H. J. Herrmann, H. E. Stanley, Nanoscale dynamics of phase flipping in water near its hypothesized liquid-liquid critical point. *Sci. Rep.* **2**, 474 (2012).
24. F. H. Stillinger, A. Rahman, Improved simulation of liquid water by molecular-dynamics. *J. Chem. Phys.* **60**, 1545–1557 (1974).
25. P. G. Debenedetti, F. Sciortino, G. H. Zerze, Second critical point in two realistic models of water. *Science* **369**, 289–292 (2020).
26. J. L. F. Abascal, E. Sanz, R. García Fernández, C. Vega, A potential model for the study of ices and amorphous water: TIP4P/Ice. *J. Chem. Phys.* **122**, 234511 (2005).
27. Y. Ni, J. L. Skinner, Evidence for a liquid-liquid critical point in supercooled water within the E3B3 model and a possible interpretation of the kink in the homogeneous nucleation line. *J. Chem. Phys.* **144**, 214501 (2016).
28. Y. Li, J. Li, F. Wang, Liquid-liquid transition in supercooled water suggested by microsecond simulations. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 12209–12212 (2013).
29. J. L. F. Abascal, C. Vega, Widom line and the liquid-liquid critical point for the TIP4P/2005 water model. *J. Chem. Phys.* **133**, 234502 (2010).
30. E. B. Moore, V. Molinero, Structural transformation in supercooled water controls the crystallization rate of ice. *Nature* **479**, 506–508 (2011).
31. V. Molinero, E. B. Moore, Water modeled as an intermediate element between carbon and silicon. *J. Phys. Chem. B* **113**, 4008–4016 (2009).
32. S. Harrington, P. H. Poole, F. Sciortino, H. E. Stanley, Equation of state of supercooled water simulated using the extended simple point charge intermolecular potential. *J. Chem. Phys.* **107**, 7443–7450 (1997).
33. F. Sciortino, E. La Nave, P. Tartaglia, Physics of the liquid-liquid critical point. *Phys. Rev. Lett.* **91**, 155701 (2003).
34. H. J. C. Berendsen, J. R. Grigera, T. P. Straatsma, The missing term in effective pair potentials. *J. Phys. Chem.* **91**, 6269–6271 (1987).
35. C. J. Tainter, L. Shi, J. L. Skinner, Reparametrized E3B (explicit three-body) water model using the TIP4P/2005 model as a reference. *J. Chem. Theory Comput.* **11**, 2268–2277 (2015).
36. O. Akin-Ojo, Y. Song, F. Wang, Developing ab initio quality force fields from condensed phase quantum-mechanics/molecular-mechanics calculations through the adaptive force matching method. *J. Chem. Phys.* **129**, 064108 (2008).
37. H. Pathak *et al.*, The structural validity of various thermodynamical models of supercooled water. *J. Chem. Phys.* **145**, 134507 (2016).
38. L. P. Wang *et al.*, Systematic improvement of a classical molecular model of water. *J. Phys. Chem. B* **117**, 9956–9972 (2013).
39. S. Naserifar, W. A. Goddard III, Liquid water is a dynamic polydisperse branched polymer. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 1998–2003 (2019).
40. S. Naserifar, W. A. Goddard III, Anomalies in supercooled water at ∼230 K arise from a 1D polymer to 2D network topological transformation. *J. Phys. Chem. Lett.* **10**, 6267–6273 (2019).
41. S. Naserifar, W. A. Goddard III, The quantum mechanics-based polarizable force field for water simulations. *J. Chem. Phys.* **149**, 174502 (2018).
42. S. K. Reddy *et al.*, On the accuracy of the MB-pol many-body potential for water: Interaction energies, vibrational frequencies, and classical thermodynamic and dynamical properties from clusters to liquid water and ice. *J. Chem. Phys.* **145**, 194504 (2016).
43. G. A. Cisneros *et al.*, Modeling molecular interactions in water: From pairwise to many-body potential energy functions. *Chem. Rev.* **116**, 7501–7528 (2016).
44. R. Car, M. Parrinello, Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.* **55**, 2471–2474 (1985).
45. J. Sun, A. Ruzsinszky, J. P. Perdew, Strongly constrained and appropriately normed semilocal density functional. *Phys. Rev. Lett.* **115**, 036402 (2015).
46. M. Chen *et al.*, Ab initio theory and modeling of water. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 10846–10851 (2017).
47. F. Noé, A. Tkatchenko, K. R. Müller, C. Clementi, Machine learning for molecular simulation. *Annu. Rev. Phys. Chem.* **71**, 361–390 (2020).
48. J. Behler, M. Parrinello, Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.* **98**, 146401 (2007).
49. T. Morawietz, A. Singraber, C. Dellago, J. Behler, How van der Waals interactions determine the unique properties of water. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 8368–8373 (2016).
50. J. Han, L. Zhang, R. Car, W. E, Deep potential: A general representation of a many-body potential energy surface. *Commun. Comput. Phys.* **23**, 629–639 (2018).
51. L. Zhang, J. Han, H. Wang, R. Car, W. E, Deep potential molecular dynamics: A scalable model with the accuracy of quantum mechanics. *Phys. Rev. Lett.* **120**, 143001 (2018).
52. B. Cheng, E. A. Engel, J. Behler, C. Dellago, M. Ceriotti, Ab initio thermodynamics of liquid and solid water. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 1110–1115 (2019).
53. A. Singraber, T. Morawietz, J. Behler, C. Dellago, Density anomaly of water at negative pressures from first principles. *J. Phys. Condens. Matter* **30**, 254005 (2018).

54. L. Zhang, D. Y. Lin, H. Wang, R. Car, W. E, Active learning of uniformly accurate interatomic potentials for materials simulation. *Phys. Rev. Mater.* **3**, 023804 (2019).
55. L. Zhang *et al*., "End-to-end Symmetry Preserving Inter-atomic Potential Energy Model for Finite and Extended Systems" in *Advances in Neural Information Processing Systems 31*, S. Bengio, Ed. *et al*. (Curran Associates, Inc., Red Hook, NY, 2018), pp. 4436–4446.
56. H. Okumura, Y. Okamoto, Molecular dynamics simulations in the multibaric-multithermal ensemble. *Chem. Phys. Lett.* **391**, 248–253 (2004).
57. M. S. Shell, P. G. Debenedetti, A. Z. Panagiotopoulos, Generalization of the Wang-Landau method for off-lattice simulations. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **66**, 056703 (2002).
58. P. M. Piaggi, M. Parrinello, Multithermal-multibaric molecular simulations from a variational principle. *Phys. Rev. Lett.* **122**, 050601 (2019).
59. V. Holten, J. C. Palmer, P. H. Poole, P. G. Debenedetti, M. A. Anisimov, Two-state thermodynamics of the ST2 model for supercooled water. *J. Chem. Phys.* **140**, 104502 (2014).
60. R. S. Singh, J. W. Biddle, P. G. Debenedetti, M. A. Anisimov, Two-state thermodynamics and the possibility of a liquid-liquid phase transition in supercooled TIP4P/2005 water. *J. Chem. Phys.* **144**, 144504 (2016).
61. J. W. Biddle *et al*., Two-structure thermodynamics for the TIP4P/2005 model of water covering supercooled and deeply stretched regions. *J. Chem. Phys.* **146**, 034502 (2017).
62. V. Holten, M. A. Anisimov, Entropy-driven liquid-liquid separation in supercooled water. *Sci. Rep.* **2**, 713 (2012).
63. V. Holten, J. V. Sengers, M. A. Anisimov, Equation of state for supercooled water at pressures up to 400 MPa. *J. Phys. Chem. Ref. Data* **43**, 043101 (2014).
64. Y. Zhang *et al*., DP-GEN: A concurrent learning platform for the generation of reliable deep learning based potential energy models. *Comput. Phys. Commun.* **253**, 107206 (2020).
65. G. Kresse, J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B Condens. Matter* **54**, 11169–11186 (1996).
66. G. Kresse, J. Furthmuller, Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15–50 (1996).
67. H. Wang, L. Zhang, J. Han, W. E, DeePMD-kit: A deep learning package for many-body potential energy representation and molecular dynamics. *Comput. Phys. Commun.* **228**, 178–184 (2018).
68. M. Ceriotti *et al*., Nuclear quantum effects in water and aqueous systems: Experiment, theory, and current challenges. *Chem. Rev.* **116**, 7529–7550 (2016).
69. Y. Yao, Y. Kanai, Temperature dependence of nuclear quantum effects on liquid water via artificial neural network model based on SCAN meta-GGA functional. *J. Chem. Phys.* **153**, 044114 (2020).
70. V. Holten, D. T. Limmer, V. Molinero, M. A. Anisimov, Nature of the anomalies in the supercooled liquid state of the mW model of water. *J. Chem. Phys.* **138**, 174501 (2013).
71. H. Tanaka, Simple physical model of liquid water. *J. Chem. Phys.* **112**, 799–809 (2000).
72. J. Russo, H. Tanaka, Understanding water's anomalies with locally favoured structures. *Nat. Commun.* **5**, 3556 (2014).
73. J. Russo, K. Akahane, H. Tanaka, Water-like anomalies as a function of tetrahedrality. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E3333–E3341 (2018).
74. G. Zhao *et al*., Equation of state of water based on the SCAN meta-GGA density functional. *Phys. Chem. Chem. Phys.* **22**, 4626–4631 (2020).
75. R. Shi, H. Tanaka, Direct evidence in the scattering function for the coexistence of two types of local structures in liquid water. *J. Am. Chem. Soc.* **142**, 2868–2875 (2020).
76. C. McBride, E. G. Noya, J. L. Aragones, M. M. Conde, C. Vega, The phase diagram of water from quantum simulations. *Phys. Chem. Chem. Phys.* **14**, 10140–10146 (2012).
77. S. Plimpton, Fast parallel algorithms for short-range molecular-dynamics. *J. Comput. Phys.* **117**, 1–19 (1995).
78. G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, G. Bussi, PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* **185**, 604–613 (2014).
79. O. Valsson, VES Code, a library that implements enhanced sampling methods based on Variationally Enhanced Sampling written by O. Valsson. https://www2.mpip-mainz.mpg.de/~valsson/ves-code/. Accessed 25 October 2019.
80. O. Valsson, M. Parrinello, Variational approach to enhanced sampling and free energy calculations. *Phys. Rev. Lett.* **113**, 090601 (2014).
81. T. E. Gartner III *et al*, Data from "Signatures of a liquid-liquid transition in an ab-initio deep neural network model for water." Princeton DataSpace. https://doi.org/10.34770/45m3-am91. Deposited 2 July 2020.