Molecular Simulation Studies of Aqueous Interfacial Fluids

A THESIS
SUBMITTED TO THE FACULTY OF THE
UNIVERSITY OF MINNESOTA
BY

Jingyi Chen

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Avisor: J. Ilja Siepmann

February 2021

Acknowledgement

First of all, I would like to express my deepest gratitude to my advisor Prof. Ilja Siepmann for his guidance and support through my PhD studies. I have learned a lot from his research expertise, and more importantly, his attitudes towards science. It is truly a rewarding experience and a great honor to work under his guidance.

I would also like to thank the lovely members and alumni of the Siepmann group. Dr. Bai Xue offered me a lot of help when I just joined the group. Dr. Evgenii Fetisov kindly provided me with a lot of resources for learning first-principles Monte Carlo simulations. Yangzesheng Sun helped with code compiling and job submitting at the Argonne Leadership Computing Facility. There are many others with whom I have had inspiring conversations, and I feel very fortunate to be a member of such a friendly and lively group.

It is also a great pleasure to collaborate with Prof. Krishnan Mahesh, and his students Mrugank Bhatt, Filipe Brandao, and Karim Alame from the Aerospace Engineering and Mechanics Department. Interdisciplinary collaborations are surely a precious part of my PhD studies as they have broadened my knowledge and gave me a different perspective.

Finally, I want to acknowledge the financial support from the Abu Dhabi Petroleum Institute, the Office of Naval Research, and National Science Foundation. Part of the computer resources were provided by the Minnesota Supercomputing Institute and the Argonne Leadership Computing Facility.

Abstract

Aqueous heterogeneous fluids with interfaces are ubiquitous and play an important role in many natural and industrial processes. In this work, two types of these systems: water/oil liquid–liquid and bubbly water liquid–vapor systems are investigated using molecular simulations.

First, Monte Carlo simulations are performed to predict the interfacial tension (IFT) for water/n-dodecane, water/toluene, and water/(50 wt% n-dodecane + 50 wt% toluene) mixtures at elevated temperatures and pressures. Simulations for the binary mixtures are performed in the $N_1N_2p_{\mathrm{N}}AT$ ensemble and those for the ternary mixtures are performed in the $N_1\mu_2\mu_3p_{\mathrm{N}}AT$ ensemble. In order to control the bulk composition of the organic phase for ternary mixture simulations, separate reservoirs for toluene and n-dodecane molecules were utilized. Identity switch moves are applied to facilitate the sampling of the spatial distributions of the organic molecules. Calculations of the IFT for water/n-decane and water/benzene mixtures were used to benchmark the force fields and to develop mixing rules. With the modified Lennard-Jones cross-interaction parameters, the simulated IFTs agree well with the experimental data.

For the bubbly water systems, molecular dynamics simulations in the canonical ensemble were performed to probe a variety of thermophysical properties of both homogeneously stretched and bubbly water systems using both the coarse-grained single-site mW and atomistic TIP4P/2005 water model. Following two simulation protocols starting either from a homogeneous configuration or from a heterogeneous configuration with a single spherical cavity, spinodal cavitation and bubble collapse points were located separately. This behavior of a fluid in a box of fixed volume is analogous to the hysteresis observed for adsorption–desorption isotherms of a subcritical fluid in a mesoporous adsorbent. In terms of thermophysical properties, both water models give qualitatively consistent result, with a common step change observed as liquid ruptures and cavitation occurs. For the heterogeneous (bubbly) phase, the Young-Laplace relation is found to hold well.

In addition to equilibrium properties, molecular dynamic simulations for the collapse of a bubble in both neat water and water/nitrogen mixture are performed, with the largest system reaching half a micron in linear dimension using the mW water model. A significant system size dependence is observed. Supersonic wall speeds preceding the initial bubble collapse and extreme local heating at the center of the collapsing bubble are only observed for large bubbles. In contrast to previous equilibrium studies, two water models predict qualitatively different collapse behaviors, which could be explained with the help of Rayleigh-Plesset equation. The presence of nitrogen gas inside the bubble leads to less violent collapse and stronger oscillations in bubble size after initial collapse. The nitrogen molecules are found to be highly compressed to a liquid-like density near the initial collapse point due to the slow dissolution of nitrogen into surrounding liquid.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Aqueous heterogeneous fluids with interfaces are ubiquitous and play an important role in many natural and industrial processes. For example, in enhanced oil recovery, surfactant flooding is an important technique to reduce the interfacial tension (IFT) of water/oil systems and shift the wettability of the reservoir towards more water-wet so that residual oil is mobilized [1, 2]. However, a reservoir's ambient conditions depend on its depth and can be characterized by high temperatures and pressures [3], and experimental data of water/oil IFT, especially at elevated temperatures and pressures, are highly scattered. This can be illustrated by the large spread among experimental data obtained by different research groups for the IFT of the water/n-decane mixture (see Figure 1.1). The difference between two data points at essentially the same state point is found to be as large as 6 dyn/cm. The experimental data also show significant differences in the pressure dependence of the IFT. In this case, molecular simulations may serve as an alternative to experimental measurement. Thus the goal of the first part of this work is to predict the IFT of binary and ternary water/oil mixtures at elevated temperatures and pressures using Monte Carlo simulations with advanced sampling techniques. The interested reader is referred to Chapter 2 for a detailed description of this work.

In contrast to the previously mentioned water/oil liquid–liquid systems, bubbly water systems are composed of a liquid water phase and a vapor or gas phase. Here, the term "vapor" refers to a phase that condenses to a liquid phase upon the increase of pressure at constant temperature (i.e., just water vapor), whereas a gas phase contains noncondensable species at this temperature (e.g., nitrogen or oxygen). Consider a sur-

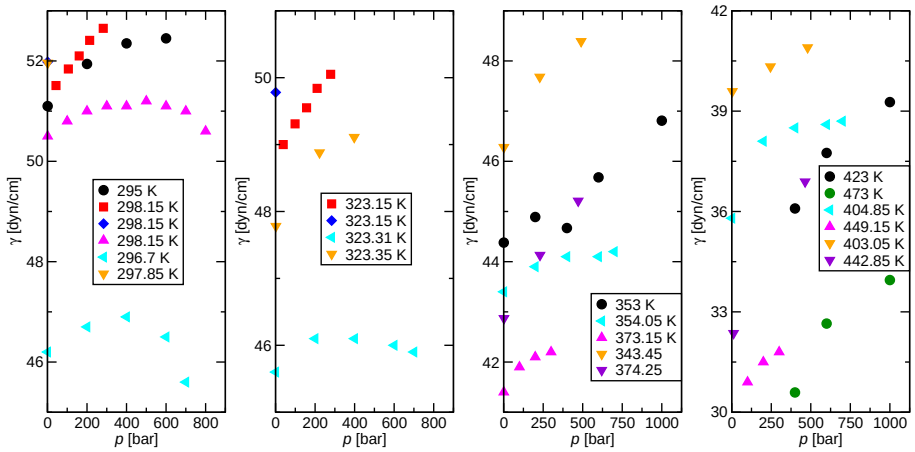


Figure 1.1: Experimental IFT data for the water/n-decane mixture grouped by temperature ranges. The black and green circles, red squares, blue diamonds, magenta up triangles, cyan left triangles, and orange and purple down triangles denote data reported by Wiegand and Franck [4], Cai *et al.* [5], Zeppieri *et al.* [6], Jennings [7], Michaels and Hauser [8], and Georgiadis *et al.* [9], respectively.

face ship or submerged vessel moving through a turbulent ocean environment, near the rotating propeller blades where the fluid is being stretched, cavitation occurs, and a large number of bubbles are formed (see Figure 1.2). Subsequent bubble collapse can potentially generate strong shock waves and high local temperatures that can affect propulsive efficiency, cause damage to the blades (see Figure 1.2), and influence wake signature [10, 11]. Despite the technological importance, understanding and predicting multi-phase cavitating flows remains a great challenge, as it involves various physical phenomena with a large span in both spatial and temporal scales. For example, molecular-scale (nanometer and picosecond) information is required to understand the collapse of a bubble, while periodic shedding of large clouds of bubbles from a hydrofoil can only be simulated at macroscopic scale (micron to millimeter) [12]. Due to this considerable complexity, this bubbly water project, with Professor Krishnan Mahesh from the Aerospace Engineering Department as the PI, involves collaborative efforts from multiple experimental and computational research groups. While all other computational groups are tackling macroscopic phenomena with computational fluid dynamics (CFD) simulations, molecular simulations are utilized in this work to study bubbly water systems.

First, molecular simulations are performed to calculate various thermophysical properties, such as pressure, energy, heat capacity, viscosity, and self-diffusion coefficient for bubbly water systems. To obtain a full picture of the thermodynamic behavior of such

Figure 1.2: Top: cavitating propeller model [13]; Bottom: cavitation damage evident on the propeller of a watercraft [13].

systems, homogeneously stretched systems are also included. Results obtained from molecular simulations can be reduced to provide necessary input parameters for CFD simulations (e.g. the correlation between system viscosity and bubble volume fraction), which are mostly empirical due to lack of experimental data at thermodynamically metastable or unstable conditions for bulk systems. The interested reader is referred to Chapter 3 for a detailed description of this work.

In addition to equilibrium properties, molecular dynamics simulations are performed to probe bubble collapse dynamics, especially the local extreme conditions, such as temperature and pressure, reached for the bubble interior and surface near the collapse point. This topic is of interest for two reasons. First, experimental measurements of temperature and pressure inside the bubble during collapse are mostly indirect due to technical difficulties, and even the most expensive high-speed cameras cannot capture dynamics at gigahertz and submicron resolution [10, 14]. Second, CFD simulations are not able to resolve the dynamics near the end of bubble collapse, as the underlying continuum assumptions break down when the bubble size gets sufficiently small. The interested reader is referred to Chapter 4 for a detailed description of this work.

Chapter 2

A Monte Carlo simulation study of the interfacial tension for water/oil mixtures at elevated temperatures and pressures[a]

2.1 Introduction

This work is carried out to address the 9th Industrial Fluid Property Simulation Challenge (IFPSC) [15]. The challenge for 2016 was to predict the interfacial tension (IFT) for water/*n*-dodecane, water/toluene, and water/(50 wt% *n*-dodecane + 50 wt% toluene) mixtures at a pressure of 1.83 MPa and four temperatures (383.15, 403.15, 423.15, and 443.15 K). Although water/oil IFTs are crucial for many industrial processes such as enhanced oil recovery [16] and control of detergency [17], experimental measurements of them, especially at elevated temperatures and pressures, remain a challenge [18]. This can be illustrated by the large spread among experimental data obtained by different research groups for the IFT of the water/*n*-decane mixture (see Figure 1.1).

[a]This chapter is adapted from the publication where I am the first author: Chen, J. L.; Xue, B.; Harwood, D. B.; Chen, Q. P.; Peters, C. J.; Siepmann, J. I. A Monte Carlo simulation study of the interfacial tension for water/oil mixtures at elevated temperatures and pressures: Water/*n*-dodecane, water/toluene, and water/(*n*-dodecane + toluene). *Fluid Ph. Equilibria* **2018**, 476, 16–24. My contribution includes implementation of the modified Janeček long-range correction, running the simulations for IFT calculations for binary mixtures, and writing up the manuscript. Dr. Bai Xue was responsible for running the simulations for k_{ij} parameterization and IFT calculations for ternary mixtures. Dr. Qile Paul Chen was responsible for designing the impurity molecules for the identity switch move.

The difference between two data points at essentially the same state point is found to be as large as 6 dyn/cm. The experimental data also show significant differences in the pressure dependence of the IFT.

Molecular simulations may offer a tool to quantitatively predict the IFT of water/oil mixtures at elevated temperatures and pressures. Furthermore, they can also provide molecular-level insight into interfacial enrichment and orientational order. Previous simulation studies for water/oil interfacial systems were mostly concerned with short-chain alkanes and were carried out either at low temperatures or at low pressures [19–23]. Therefore, the 9th IFPSC stimulates an effort to investigate water/oil interfacial systems at both elevated temperatures and pressures. In this work, the accuracy of the combination of the TraPPE force fields for aliphatic and aromatic hydrocarbons with the TIP4P/2005 water model is assessed based on available experimental data for the IFT of water/*n*-decane and water/benzene mixtures. Then modifications of the energetic Lennard-Jones (LJ) cross-interaction parameters are introduced and the modified force fields are applied to predict the IFT for water/*n*-dodecane, water/toluene, and water/(50 wt% *n*-dodecane + 50 wt% toluene) mixtures at elevated temperatures and pressures.

2.2 Simulation methods

2.2.1 Force fields

The LJ 12–6 and Coulomb potentials are used to describe the intermolecular interactions of water, *n*-alkanes, benzene, and toluene, and also the nonbonded intramolecular interactions for the flexible alkane chains:

$$u_{ij}(r_{ij}) = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \tag{2.1}$$

where r_{ij}, ϵ_{ij}, σ_{ij}, q_i, q_j, and ϵ_0 are the distance between interaction sites i and j, the LJ well depth, the LJ diameter, the partial charges on interaction sites i and j, and the permittivity of vacuum, respectively. The modified Lorentz–Berthelot (LB) combining rules [24] are used to determine the LJ parameters for unlike interactions:

$$\varepsilon_{ij} = k_{ij}\sqrt{\varepsilon_{ii}\varepsilon_{jj}} \qquad \sigma_{ij} = \frac{\sigma_{ii} + \sigma_{jj}}{2} \tag{2.2}$$

where k_{ij} is a parameter used to scale the unlike well depth to mimic departures from the Berthelot rule; $k_{ij} = 1.0$ corresponds to the standard Berthelot combining rule. In all simulations, water is represented by the rigid four-site TIP4P/2005 model [25] because, in addition to overall excellent performance for neat water, this model also yields very accurate predictions for the liquid–vapor surface tension of neat water [26]. n-Decane and n-dodecane are represented by ten- and twelve-site semi-flexible united-atom models with the interaction parameters taken from the TraPPE–UA force field [27] that is also known to yield accurate predictions for the surface tension [28]. Benzene and toluene are represented by rigid twelve-site models with the interaction parameters for carbon and explicit hydrogen atoms being part of the aromatic ring taken from the TraPPE–EH force field [29] and parameters for the united-atom methyl group of toluene taken from the TraPPE–UA force field [27]. The partial charges of toluene are chosen following the convention used for other substituted aromatic compounds [30]. The surface tensions of toluene obtained from simulations using this force field are in a fairly good agreement with experimental data (Table SA1 and Figure SA1). The Supporting Information (Tables SA2 and SA3) provides a detailed description of all force field parameters and the charge parameterization procedure.

2.2.2 Calculation method used for the interfacial tension

The IFT, γ_{KB}, is calculated based on the Kirkwood and Buff definition [31, 32]. The simulations are performed in an elongated (along the z-direction) box with periodic boundary conditions in all three directions that contains two liquid phases and two planar interfaces perpendicular to the z-direction. In this case, γ_{KB} can be determined as follows:

$$\gamma_{\mathrm{KB}} = \frac{1}{2} \langle P_{\mathrm{N}} - P_{\mathrm{T}} \rangle L_z \tag{2.3}$$

where L_z is the length of the periodic simulation box in the z-direction, P_{N} and P_{T} are the normal and tangential components, respectively, of the pressure tensor, and the factor of $1/2$ accounts for the presence of two interfaces in the periodic box.

2.2.3 Long-range corrections for Lennard-Jones interactions

The Janeček long-range correction [33] is utilized for the LJ part of the energy and pressure calculations, whereas the Ewald summation is used for the Coulomb part (see below). The simulation box is divided into a fixed number, N_s, of thin slabs along the z-direction. In the current work, N_s is set to 400 yielding a slab width of $\Delta z = L_z/400$ that fluctuates in proportion to L_z in the $N_1N_2p_\mathrm{N}AT$ and $N_1\mu_2\mu_3p_\mathrm{N}AT$ ensembles. The total long-range correction to the potential energy is calculated by summing the contributions from all individual slabs:

$$U_\mathrm{LRC} = \frac{1}{2}\sum_{n=1}^{N_\mathrm{s}} u_\mathrm{LRC}(z_n) \tag{2.4}$$

where z_n denotes the position of the center of each slab. The contribution from each slab is calculated from

$$u_\mathrm{LRC}(z_n) = \sum_{i=1}^{t_\mathrm{LJ}}\sum_{j=1}^{t_\mathrm{LJ}} \rho_i(z_n)V_\mathrm{s}\sum_{m=1}^{N_\mathrm{s}} \rho_j(z_m)\; w_{ij}(|z_m - z_n|)\Delta z \tag{2.5}$$

where t_LJ is the number of types of LJ interaction sites, $\rho_i(z_n)$ and $\rho_j(z_n)$ are the number densities for interaction sites of a specific type in slab n located at z_n, and V_s is the volume of each slab. In the present work, the density profiles are evaluated on the fly during each Monte Carlo trial move (i.e., accounting for minor differences in $\rho(z_n)$ between trial and old configurations).

The function $w_{ij}(|z_m - z_n|)$ is computed via the following equation by assuming a uniform distribution of interaction sites in the slab beyond the spherical potential truncation at R_c:

$$w_{ij}(\zeta) = \begin{cases} 4\pi\epsilon_{ij}(\sigma_{ij})^2\left[\frac{1}{5}\left(\sigma_{ij}/R_\mathrm{c}\right)^{10} - \frac{1}{2}\left(\sigma_{ij}/R_\mathrm{c}\right)^4\right] & \zeta \le R_\mathrm{c} \\ 4\pi\epsilon_{ij}(\sigma_{ij})^2\left[\frac{1}{5}\left(\sigma_{ij}/\zeta\right)^{10} - \frac{1}{2}\left(\sigma_{ij}/\zeta\right)^4\right] & \zeta > R_\mathrm{c} \end{cases} \tag{2.6}$$

The long-range correction for the IFT is obtained here through calculating the long-range correction for different components of the pressure tensor in an analogous manner. Details can be found in Ref. [33].

2.2.4 Simulation details

For all simulations in this work, a spherical potential truncation at $R_c = 16$ Å is utilized for the intermolecular LJ interactions. The Ewald summation method [34] is utilized to calculate the Coulomb interactions with a screening parameter of $\kappa = 3.2/R_c$ and an upper bound of the reciprocal space summation at $K_{max,i} = \text{int}(\kappa L_i) + 1$, where i can be x, y, or z and L_i is the length of the corresponding box dimension. An additional hard-sphere potential with $\sigma_{HS} = 1.2$ Å is used to avoid unphysical configurations of two water molecules where a hydrogen site with partial positive charge belonging to one molecule would nearly coincide with the negatively charged M site of another molecule.

All simulations are carried out using the Monte Carlo for Complex Chemical Systems–Minnesota (MCCCS–MN, version 16.1) code developed in house. In order to sample the classical phase space, the following types of moves are performed: center-of-mass translations [35] for all types of molecules, rotations around the center of mass [36] for all types of molecules, conformational changes via the dual cut-off, coupled-decoupled configurational-bias Monte Carlo (DCCD-CBMC) algorithm [37–39] only for alkane molecules, and volume changes with an external reservoir [40]. For simulations of the ternary mixture, particle transfers are performed using DCCD-CBMC with multiple trial sites for the insertion of the first bead [38–41]. During the equilibration period, the maximum displacements are adjusted to yield an acceptance ratio of about 40% for translational, rotational and volume moves. The probabilities of each type of move are distributed to approximately reflect the number of degrees of freedom of each type to be sampled. Details about the distribution of move probabilities are provided in Tables SA3 and SA4.

Simulations for the parametrization of k_{ij}. Water/n-decane and water/benzene IFT simulations with the normal and modified LB combining rules are run to assess the accuracy of the force fields. The simulations are carried out in the isobaric-isothermal ensemble with fixed interfacial area ($N_W N_O p_N A T$, where the subscripts W and O denote water and organic compounds). These simulations are initialized using the following procedure. First, simulations in the canonical (NVT) ensemble are performed to obtain stable liquid configurations for unary systems consisting of 1500 water, 150 n-decane, or 290 benzene molecules using the same box size of $32 \times 32 \times 50$ Å^3 for all three systems. Then the water box is appended to the n-decane (or benzene) box along the z-direction.

After the water and n-decane (or benzene) boxes are combined, the system is briefly equilibrated in the canonical ($N_{\mathrm{W}}N_{\mathrm{O}}VT$) ensemble to remove energetically unfavorable configurations. Thereafter, the simulations are switched to the $N_{\mathrm{W}}N_{\mathrm{O}}p_{\mathrm{N}}AT$ ensemble and only the z-dimension of the simulation box is allowed to fluctuate. The binary systems are further equilibrated in the $N_{\mathrm{W}}N_{\mathrm{O}}p_{\mathrm{N}}AT$ ensemble for at least 50,000 Monte Carlo cycles (MCCs), where one MCC consisting of N randomly chosen MC moves, where $N = N_{\mathrm{W}} + N_{\mathrm{O}}$ is the total number of molecules in the system. For the k_{ij} parameters used for the challenge entry, statistics are gathered from 8 independent simulations, each consisting of at least 100,000 MCCs for production. After submission of the challenge entry, the lengths of the production periods are extended to ensure that the uncertainty of each data point is reduced to be less than 2.1 dyn/cm, and additional low-temperature, low-pressure state points were investigated to find improved k_{ij} parameters.

Simulations for binary mixtures. Simulations for water/n-dodecane and water/toluene mixtures are performed in the $N_{\mathrm{W}}N_{\mathrm{O}}p_{\mathrm{N}}AT$ ensemble at $T = 383.15$, 403.15, 423.15, 443.15 K and $p_{\mathrm{N}} = 1.83$ MPa. These systems contain 1500 water molecules and either 120 n-dodecane or 260 toluene molecules. The cross-sectional area of the simulation box is 1024 Å^2 with $L_x = L_y = 32$ Å. The simulations are set-up using the procedure outlined above for the water/n-decane and water/benzene mixtures. Statistics for the IFT prediction are gathered from 16 independent simulations, each consisting of at least 140,000 MCCs for the production periods. The predictions using this simulation protocol are denoted in the following sections as O1. After the submission of the IFPSC entry, simulations with the same k_{ij} parameters are all elongated to 200,000 MCCs for the production period. The predictions using this simulation protocol are denoted in the following sections as O2. Furthermore, 16 independent simulations with the improved k_{ij} parameters are also carried out using 200,000 MCCs for the production periods. The predictions using this simulation protocol are denoted in the following as I2.

Simulations for ternary mixtures. Simulations for the water/(50 wt% n-dodecane + 50 wt% toluene) mixture are carried out in the $N_{\mathrm{W}}\mu_{\mathrm{D}}\mu_{\mathrm{T}}p_{\mathrm{N}}AT$ ensemble, where the subscripts D and T denote n-dodecane and toluene, respectively. Figure 2.1 illustrates the simulation set-up. In order to control the organic-phase composition, the simulations with the MCCCS–MN software utilize a 3-box $N_{\mathrm{W}}N_{\mathrm{D}}N_{\mathrm{T}}(p_{\mathrm{N}}A)p_{\mathrm{D}}p_{\mathrm{T}}T$ osmotic

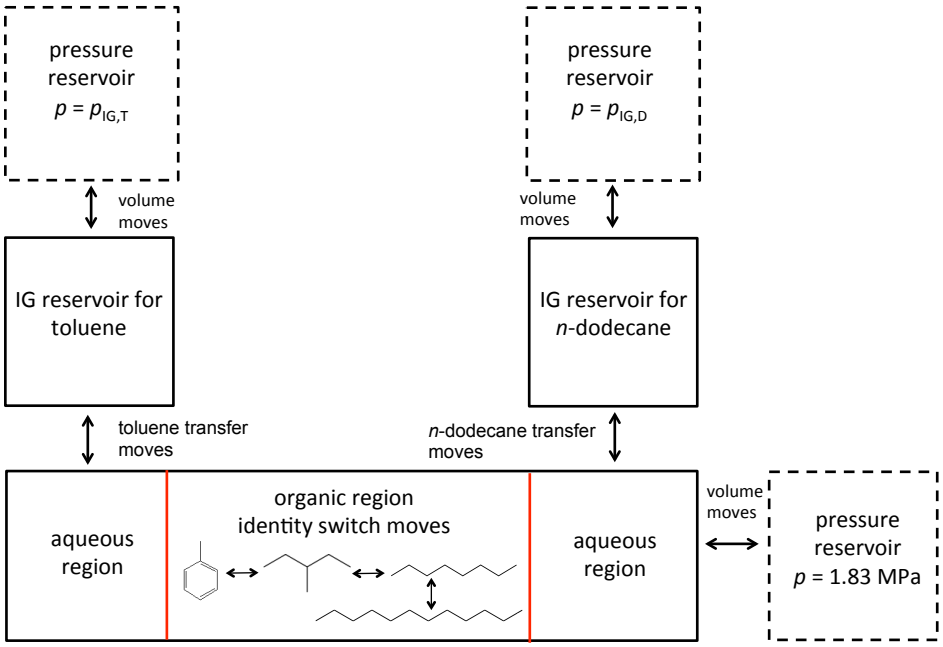


Figure 2.1: System set-up for simulations of the ternary systems

Gibbs ensemble set-up [42, 43]. The elongated liquid box that contains the two interfaces is initialized following the same procedure as for the binary simulations, but replacing the neat organic phase with a 50 wt% n-dodecane + 50 wt% toluene mixture. Only the z-dimension of the liquid box is allowed to fluctuate with p_N = 1.83 MPa. In order to facilitate sampling of the spatial distribution of the n-dodecane and toluene molecules in the organic-phase region, identity switch moves with the DCCD-CBMC algorithm [44–47] are performed utilizing two types of intermediate impurities: 3-methyl-pentane and n-octane; only one molecule of each impurity type is introduced into the liquid-phase box and these molecules are not allowed to transfer into another box.

The other two boxes are separate "ideal-gas" reservoirs for n-dodecane and toluene molecules, respectively, and these molecules are allowed to transfer between the liquid-phase box and their corresponding reservoirs to equilibriate their chemical potentials. The reasons for using reservoirs with explicit molecules but all intermolecular interactions switched-off are three-fold. First, the dual cut-off CBMC formalism reduces computational demands for particle transfer and conformational moves, but does not allow for the calculation of chemical potentials via the Rosenbluth weights [38]. Second, the calculation of Gibbs free energies of transfer is more precise and less prone to finite-

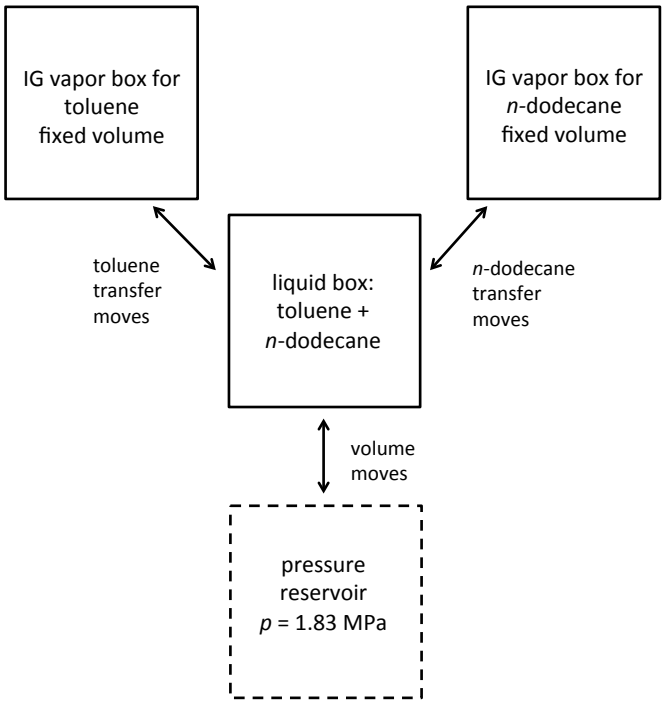


Figure 2.2: System set-up for simulations to determine the ideal gas pressure used in the ternary simulations

size effects than the calculation of excess chemical potentials [48]. Third, the volume of the liquid-phase box is allowed to fluctuate in response to the normal pressure, whereas the standard grand canonical ensemble requires a fixed volume.

The external pressures, p_{D} and p_{T}, of the separate n-dodecane and toluene ideal-gas reservoirs are set to be equal to those corresponding to the number densities of ideal-gas reservoirs that are in contact with a 50/50 wt% toluene and n-dodecane bulk (i.e., no interfaces) liquid mixture at $p = 1.83$ MPa and a specific temperature. Isotropic volume moves are used to maintain the pressure of these ideal-gas reservoirs.

The pressures p_{D} and p_{T} are determined by the following procedure. First, several sets of 3-box $N_{\mathrm{D}}N_{\mathrm{T}}p_{\mathrm{liq}}V_{\mathrm{D}}V_{\mathrm{T}}T$ Gibbs ensemble simulations are carried out using a cubic liquid-phase box and two separate ideal gas reservoirs both with fixed volume for n-dodecane and toluene molecules, respectively. Simulations are performed at four temperatures using multiple different n-dodecane/toluene overall numbers of molecules in the system: 159/241, 150/250, 146/254, and 140/260 (correspodning to weight ratios 54.9/45.1, 52.6/47.4, 51.5/48.5, and 49.9/50.1 wt%, respectively). The simulation set-up is illustrated in Figure 2.2. All molecules are initially placed in the liquid-phase box, and isotropic volume moves are performed on this box in contact with an external

pressure reservoir at $p = 1.83$ MPa. n-Dodecane and toluene molecules are allowed to transfer between the liquid-phase box and their corresponding ideal-gas reservoirs (see Figure SA3 for simulation set-up). In order to improve the statistics, the sizes of the vapor boxes are adjusted throughout the equilibration period, so that they contain on average approximately 20 molecules. Statistics are gathered from 8 independent simulations, each consisting of 100,000 MCCs for the production period. Fig. 2.3 (numerical data are provided in Table SA6) shows the the calculated average number densities of toluene and n-dodecane molecules in their corresponding ideal-gas boxes as function of the average mole fraction in the liquid phase at all four pressure-temperature conditions. The ideal-gas number densities of toluene and n-dodecane molecules that correspond to a 50/50 wt% ($x_{\text{toluene}} = 0.649$) in the liquid-phase box are determined through linear interpolation, and then converted into pressures using the ideal gas law (numerical data can be found in Table SA7).

As for the binary mixtures, predictions for the ternary mixture are based on 16 independent simulations, and these are performed using the protocols O1, O2, and I2.

With the set-up used here for the ternary mixture, one concern is that the global free energy minimum would be a state with most of the organic molecules in their ideal-gas reservoirs and a single aqueous phase in the elongated box (i.e., the free energy cost for the two interfaces would be removed). To this end, we monitor the number of n-dodecane and toluene molecules in the liquid-phase box. Of course, the fluctuations in the numbers of the two types of organic molecules are coupled through the volume of the organic region (plus an offset due to interfacial enrichment). For the current simulations the number of n-dodecane molecules is found to range from 31 to 77 and those for toluene from 91 to 179 with the extrema observed at the highest temperature. Thus, the current trajectories would not have been affected by hard lower boundaries of 30 n-dodecane and 90 toluene molecules. In addition, we performed another set of simulations at $T = 423.15$ K with the improved set of k_{ij} parameters during which volumes moves are not used for the liquid box, i,e., an $N_{\text{W}}N_{\text{D}}N_{\text{T}}V_{\text{liq}}p_{\text{D}}p_{\text{T}}T$ Gibbs ensemble. In this case, the external pressure is maintained through the transfers of n-dodecane and toluene molecules with chemical potentials that combine to a bulk liquid-phase pressure of 1.83 MPa. Since the free energy cost for the water/vapor surface is larger than that for the water/organic interface and the volume of the aqueous region is constrained by the

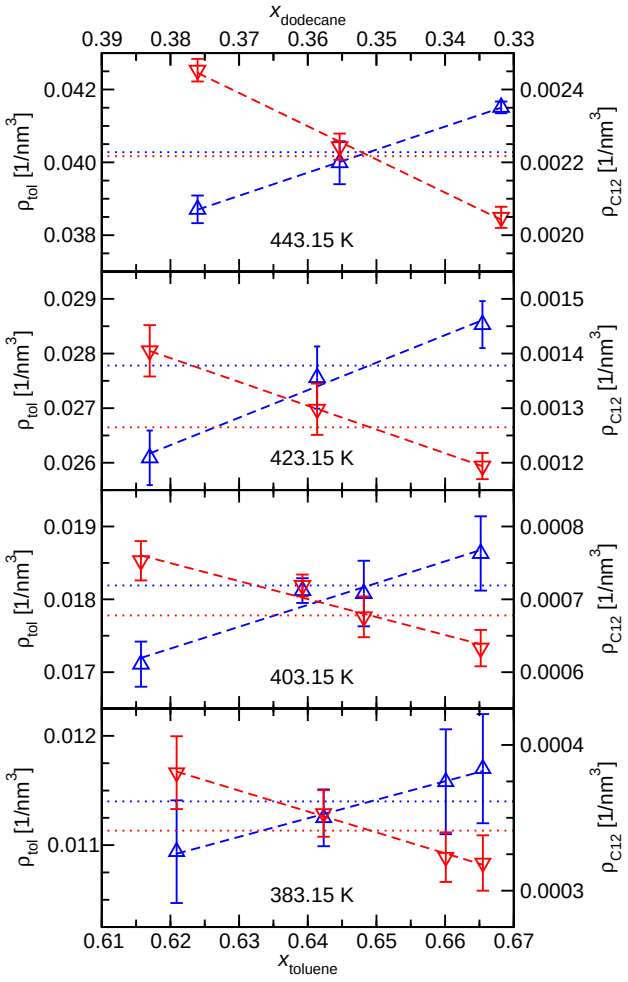


Figure 2.3: Ideal-gas number densities of toluene (blue up triangles) and n-dodecane (red down triangles) reflecting the corresponding chemical potentials as function of toluene mole fraction in the liquid phase at $T = 383.15$, 403.15, 423.15 and 443.15 K, and $p = 1.83$ MPa. Dashed lines show linear fits of the data and the dotted-lines indicate the ideal gas number densities for a 50/50 wt% ($x_{\text{toluene}} = 0.649$) mixture.

number of water molecules, the remainder of the elongated box will always be occupied by a liquid region of the organic components. The IFT values obtained with fluctuating and fixed volume for the liquid-phase box are found to be statistically indistinguishable, thereby confirming that the approach using a fluctuating volume is valid as long as the fluctuations are constrained either through a hard boundary or through a sufficient nucleation free energy barrier.

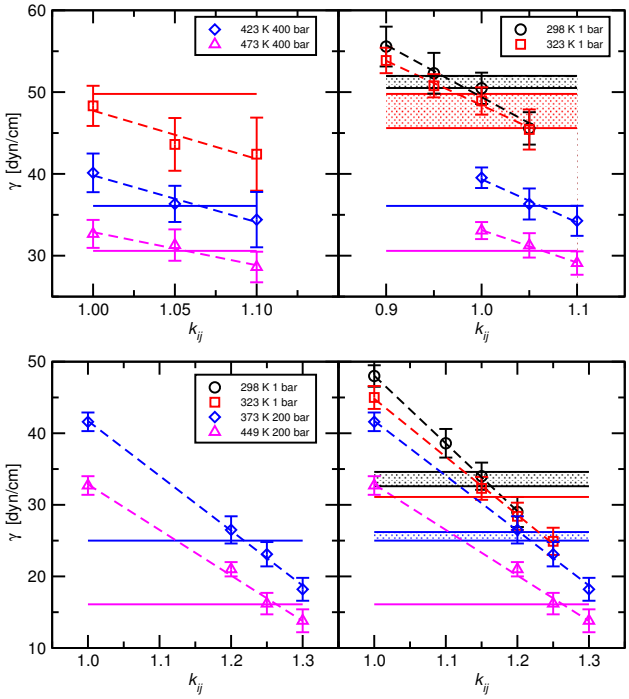


Figure 2.4: Determination of k_{ij} parameters. The left column shows the IFT data for the water/n-decane (top) and the water/benzene (bottom) mixtures versus k_{ij} at different state points that are used for the k_{ij} parameterization of the IFPSC entry. The right column shows IFT data obtained from longer simulations and also at additional state points that are used for the subsequent improved k_{ij} parameterization. Horizontal lines and shaded areas indicate specific values and ranges of experimental data [4–9].

2.3 Results and discussion

2.3.1 Parametrization of k_{ij}

IFT data assessing the accuracy of the IFT predictions with the TraPPE–UA/EH//TIP4P/2005 combination of force fields are shown in Fig. 2.4 (numerical data are provided in Tables SA7 to SA10). Using the standard LB combining rule ($k_{ij} = 1.0$), it is found that the IFT for the water/n-decane mixture is slightly underestimated at $T = 323$ K and $p = 1$ bar compared to the experimental data reported by Zeppieri *et al.* [6], but overestimated by about 10% at elevated temperatures and pressures compared to the experimental data reported by Wiegand and Franck [4]. In contrast, the IFT for the water/benzene mixture is overestimated by factors of 1.6 and 2.1 at $p = 20.0$ MPa and $T = 373$ and 449 K, respectively, compared to the experimental data reported by Jennings [7]. Considering that the molecular models used here yield fairly accurate liquid–vapor surface tensions, the large overestimation of the water/benzene IFT can be attributed to the use of non-polarizable force fields that apparently do not capture

Table 2.1: k_{ij} values used for the IFT predictions

	O/CH_x		O/(C or H)
	alkane	toluene	aromatic ring
IFPSC entry	1.03	1.03	1.25
Improved values	1.01	1.10	1.19

the significant polarization of the π-electrons (and weak charge transfer due to hydrogen bond formation) in the highly polarizable aromatic ring. In contrast, the alkane is less polarizable and also does not act as an electron donor for the formation of hydrogen bonds.

One approach to implicitly account for induced polarization effects is to modify the Berthelot combining rule for unlike interactions, and a few studies have already tried to increase the scaling parameter k_{ij} to improve water/alkane mutual solubilities [49–53]. Therefore, additional simulations for the water/n-decane mixture are run exploring $k_{ij} =$ 1.05 and 1.10 for the cross interaction between the O atom in water (only the oxygen site carries a LJ interaction in the TIP4P/2005 water model) and the $-CH_3$ and $-CH_2-$ groups in n-decane. Similarly, simulations for the water/benzene mixture are run using $k_{ij} = 1.20$, 1.25, and 1.30 for the interactions between the oxygen site in water and the aromatic interaction sites at carbon and hydrogen nuclei in the TraPPE–EH benzene model. The IFT data for both water/n-decane and water/benzene mixtures indicate a weak temperature dependence with larger deviations from the standard Berthelot rule at higher temperatures (and pressures). Nevertheless, the state-point independent k_{ij} values for our IFPSC entry are determined by averaging the values corresponding to the intersections of the selected experimental data (three and two state points for water/n-decane and water/benzene, respectively) and linear fits to the k_{ij}-dependent calculated IFT data. Furthermore, the same k_{ij} value is used for the $-CH_3$ group of toluene as for the n-alkanes. The selected k_{ij} values are summarized in Table 2.1.

After the submission of the IFPSC entry, some changes are made to the determination of k_{ij} parameters to improve the accuracy of the predictions. First, to reflect the presumably higher accuracy of experimental data at near-ambient conditions, additional state points are incorporated for both water/n-decane and water/benzene mixtures, so that two low-temperature/low-pressure state points and two high-temperature/high-pressure state points are used for each mixture. Second, longer production periods

and additional k_{ij} values were used to improve the precision of the k_{ij} parameterization. Third, multiple experimental values at the selected state points are used for the determination of the intersection points. In order to account for the differences in the number of experimental data sources among the chosen state points, the improved k_{ij} values are obtained as follows: An unweighted average of the intersection points for multiple experimental values at a given state point is calculated. Then, a weighted average is calculated from these state-point specific k_{ij} values with the weights proportional to the square root of the number of experimental data available at that state point. Fourth, since the $-CH_3$ group in toluene replaces a hydrogen site in benzene, the corresponding k_{ij} is taken to be the average between O–CH_x (aliphatic) and O–C (aromatic). Another choice would have been to fit different k_{ij} values for the carbon and hydrogen sites of benzene with, maybe, the latter using the same k_{ij} value as for the alkanes. When fitting the k_{ij} value for O–CH_x (aliphatic), one may only use the experimental data from Georgiadis *et al* [9] as it is considered to be more accurate. However, this will yield a k_{ij} value of 1.00. Considering the current uncertainties in IFT values, this 0.01 difference in k_{ij} value would have a negligible effect on the results.

2.3.2 Prediction of IFT for binary and ternary mixtures

The IFT predictions from the current simulation study are presented in Fig. 2.5 and numerical data are provided in Table 2.2. It should be emphasized first that the statistical uncertainties in the calculated IFT values are quite large; the 95% confidence intervals range from 1.1 to 2.3 dyn/cm and correspond to relative uncertainties in the range from 3 to 7%. Since the IFT values decrease on average only by 10% as the temperature is increased by 20 K, there is considerable scatter in the data over this narrow temperature interval. On the other hand, it should be noted that the spread of experimental data between independent measurements from different groups also indicates combined uncertainties of about 2 dyn/cm (see Figure 1.1).

To smooth out the noise in the predicted IFT data (e.g., the values at $T = 383.15$ and 403.15 K for the water/toluene mixture show an increase with increasing temperature that is unphysical for a regular miscibility gap ending in an upper critical solution point), our IFPSC entry is based on a linear fit to the data at the four temperatures. The IFT values of our IFPSC entry underpredict the experimental benchmark data by

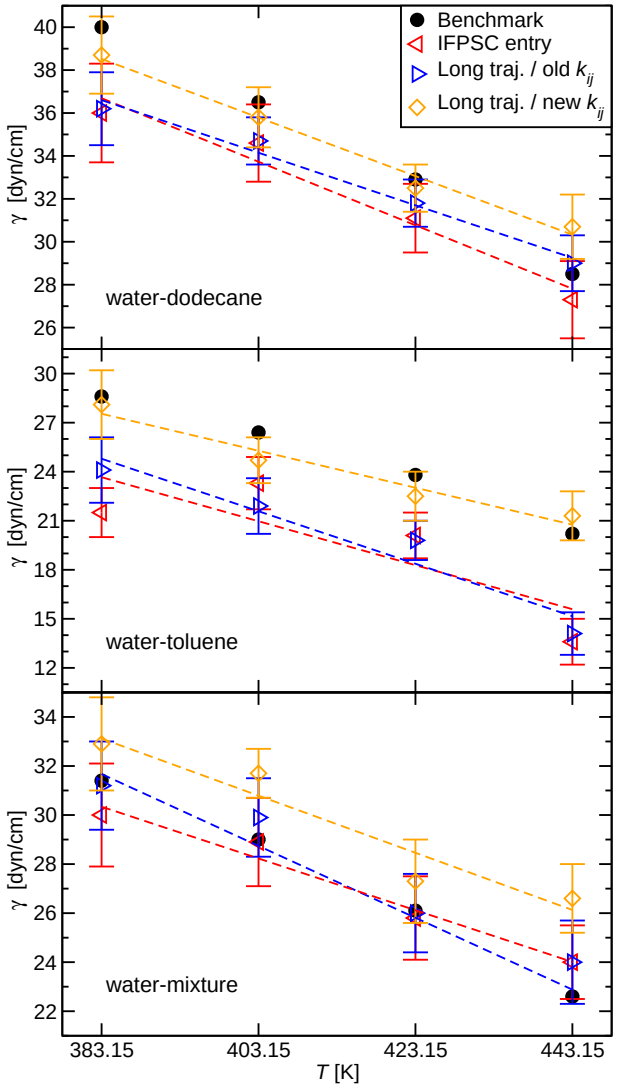


Figure 2.5: Predicted IFT data for water/n-dodecane, water/toluene, and water/(50 wt% toluene + 50 wt% n-dodecane) as function of temperature at $p = 1.83$ MPa. Open symbols denote averages obtained from Monte Carlo simulations at a specific state point, dashed lines represent the linear fits using all four temperatures, and filled black circles denote the benchmark data. The red left triangles, blue right triangles, and orange diamonds correspond to data obtained with the O1, O2, and I2 simulation protocols.

a mean signed error (MSE) of −2.2 dyn/cm (mean unsigned percentage error (MUPE) of 6%) for the water/n-dodecane mixture and with MSE of −5.1 dyn/cm (MUPE of 21%) for the water/toluene mixture, whereas the predictions for the ternary mixture are in fortuitously excellent agreement with an MSE of only −0.1 dyn/cm (MUPE of 3%). MSE and MUPE are defined in the following equations,

$$\mathrm{MSE} = \frac{1}{4}\sum_{n}(\gamma_{\mathrm{sim},n} - \gamma_{\mathrm{ben},n}) \tag{2.7}$$

Table 2.2: Benchmark and simulation data of the IFT (in units of dyn/cm) for the water/n-dodecane (W/D), water/toluene (W/T), and water/(50 wt% toluene + 50 wt% n-dodecane) (W/DT) mixtures at four temperatures and $p = 1.83$ MPa. The uncertainties denote the 95% confidence intervals. The data in bold font are our IFPSC entry. In the Protocol column, the letters A and L stand for the average obtained from 16 independent simulations at a specific state point or the value determined from a linear fit to the averages at all state points.

Mixture	Protocol	T [K]			
		383.15	403.15	423.15	443.15
W/D	Benchmark	40.0	36.5	32.9	28.5
	O1A	36.0 ± 2.3	34.6 ± 1.8	31.1 ± 1.6	27.3 ± 1.8
	O1L	**36.7**±1.6	**33.7**±1.1	**30.8**±1.2	**27.8**±1.7
	O2A	36.2 ± 1.7	34.7 ± 1.1	31.8 ± 1.1	29.0 ± 1.3
	I2A	38.7 ± 1.8	35.8 ± 1.4	32.5 ± 1.1	30.7 ± 1.5
W/T	Benchmark	28.6	26.4	23.8	20.2
	O1A	21.5 ± 1.5	23.3 ± 1.6	20.1 ± 1.4	13.6 ± 1.4
	O1L	**23.7**±1.9	**21.0**±1.1	**18.3**±1.0	**15.6**±1.6
	O2A	24.1 ± 2.0	21.9 ± 1.7	19.8 ± 1.2	14.1 ± 1.3
	I2A	28.1 ± 2.1	24.7 ± 1.4	22.5 ± 1.5	21.3 ± 1.5
W/DT	Benchmark	31.4	29.0	26.1	22.6
	O1A	30.0 ± 2.1	28.9 ± 1.8	25.8 ± 1.7	24.0 ± 1.5
	O1L	**30.3**±1.3	**28.2**±1.0	**26.1**±1.1	**24.0**±1.6
	O2A	31.2 ± 1.8	29.9 ± 1.6	26.0 ± 1.6	24.0 ± 1.7
	I2A	32.9 ± 1.9	31.7 ± 1.0	27.3 ± 1.7	26.6 ± 1.4

$$\text{MUPE} = \frac{1}{4} \sum_{n} \frac{|\gamma_{\text{sim},n} - \gamma_{\text{ben},n}|}{\gamma_{\text{ben},n}} \times 100\% \tag{2.8}$$

where n labels different state points (4 in total) of a system. $\gamma_{\text{sim},n}$ and $\gamma_{\text{ben},n}$ are simulated and benchmark IFT, respectively. The longer simulations with the same k_{ij} parameters reduce the MUPEs only very slightly to 5, 20, and 3%, respectively, thus any failures are due to the model and not due to insufficient statistics. The underprediction for both binary mixtures indicates that the k_{ij} parameters used for our IFPSC entry are too large and over-correct the IFT. Our IFPSC predictions yield values for the ternary mixture that exceed the average of the two binary mixtures by a factor of 1.05. In contrast, for the benchmark data, the IFT of the ternary mixture falls below the average of the binary mixtures by a factor of 1.09. The positive excess IFT for the ternary mixture found for our IFPSC entry indicates that adjusting k_{ij} parameters for the LJ interactions leads to a global effect (explicitly up to the cut-off and implicitly through the tail corrections), whereas a more local change may be needed to yield a

negative excess IFT. An attempt to construct a localized k_{ij} parameter was discussed in the SI. Utilizing this localized k_{ij} parameter, a negative excess IFT was able to be reproduced only at the lowest temperature, while failed at the other three temperatures. This indicates that it remains a challenge to develop a good non-polarizable force field to mimic very localized induced polarization and charge transfer effect.

As is evident from Fig. 2.5, the improved k_{ij} parameters based on additional state points and longer simulations yield vastly better IFT predictions for the two binary mixtures with MSEs of only -0.1 and -0.6 dyn/cm (MUPEs of 4 and 5%) for the water/n-dodecane and water/toluene mixtures, respectively. However, the MSE and MUPE for the ternary mixture are now +2.4 dyn/cm and 9%, respectively. With the improved k_{ij} parameters, the predictions for the ternary mixture exceed the average of the two binary mixtures only by a factor of 1.01, but still a positive excess IFT.

For the predictions with the improved k_{ij} parameters, the IFT values at $T = 383.15$ K exceed those at 443.15 K by factors of 1.26, 1.32, and 1.24 (note that the last digit is not statistically significant) for water/n-dodecane, water/toluene, and the ternary mixture, respectively. These factors are somewhat smaller than those for the benchmark data (1.40, 1.42, and 1.39, respectively). However, both experimental data and predictions agree that there is no significant difference in the relative decrease of the IFT for the three mixtures over the temperature range studied.

2.3.3 Interfacial enrichment for the ternary mixture

Since the IFT of the binary water/n-dodecane mixture significantly exceeds that of the water/toluene mixture, one may expect an enrichment of toluene molecules at the water/organic interface for the ternary mixture. To provide molecular-level insight, the toluene weight fraction is determined as function of relative z-position in the elongated liquid-phase box; here the center of mass of the water region is calculated for each individual configuration and used to align the profile. As illustrated by the composition profiles shown in Fig. 2.6, the simulations indeed yield a significant toluene enrichment at the water/organic interface. This enrichment decreases with increasing temperature, e.g., the maximum values of the local toluene weight fraction are 0.77 and 0.70 at $T = 383.15$ and 443.15 K, respectively. The enrichment only extends over about 10 Å before the bulk composition of the organic phase imposed by the reservoirs is reached.

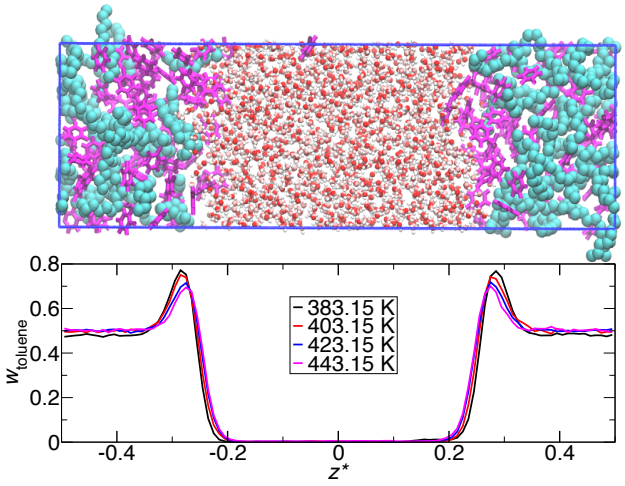


Figure 2.6: Interfacial structure for the ternary mixture using the improved k_{ij} parameters. (Top) Snapshot of a representative configuration at $T = 443.15$ K, where the oxygen and hydrogen atoms of water molecules are shown as red and white spheres, respectively, the toluene and n-dodecane molecules are represented as magenta wireframes and cyan spheres for the united atoms, respectively. Bottom: Toluene weight fraction profile along the z-dimension. The center of mass of the water region is shifted to 0, and z^* is the scaled z-coordinate with $z^* = z/L_z$.

Numerical values for the composition of the entire organic region, w_{T}, and for the bulk-like organic region, $w_{\mathrm{T}}^{\mathrm{b}}$, in the liquid-phase box are provided in Table 2.3. At the three higher temperatures, $w_{\mathrm{T}}^{\mathrm{b}}$ is close to 50 wt%, the desired composition for the ternary mixture. At $T = 383.15$ K, however, $w_{\mathrm{T}}^{\mathrm{b}} \approx 48$ wt% for both sets of k_{ij} parameters. This points to a problem with the specific reservoir pressures used at this state point, but it should be noted that the relative uncertainties in the ideal-gas number densities are significantly larger at $T = 383.15$ K (see Figure 2.3) and 50 wt% likely falls within the combined uncertainty.

More importantly, the fact that $w_{\mathrm{T}} \geq 57$ wt% for all four state points and both sets of k_{ij} parameters demonstrates the need to carry out simulations for the ternary mixture either in an open ensemble to control the composition or using an extremely large system size so that any interfacial enrichment does not lead to a significant depletion of the interfacially more active component in the bulk-like organic region. The average enrichment factors ($w_{\mathrm{T}}/w_{\mathrm{T}}^{\mathrm{c}}$) for the current system size are found to be essentially the same for the k_{ij} parameters of our IFPSC entry and the improved set (1.158 and 1.155, respectively).

Table 2.3: Toluene weight fraction in the entire organic region, w_{T}, and only within the bulk-like part, $|z^*| \geq 0.4$, of the organic region, $w_{\mathrm{T}}^{\mathrm{b}}$, of the elongated liquid-phase box. The uncertainties denote the 95% confidence intervals.

Region	Protocol	T [K]			
		383.15	403.15	423.15	443.15
w_{T} [%]	O2	56.5 ± 0.3	58.3 ± 0.4	58.6 ± 0.2	57.4 ± 0.3
w_{T} [%]	I2	57.0 ± 0.3	57.8 ± 0.2	57.4 ± 0.2	57.0 ± 0.2
$w_{\mathrm{T}}^{\mathrm{b}}$ [%]	O2	48.2 ± 0.6	50.3 ± 0.4	50.7 ± 0.3	50.1 ± 0.3
$w_{\mathrm{T}}^{\mathrm{b}}$ [%]	I2	48.1 ± 0.4	49.9 ± 0.3	50.3 ± 0.3	50.2 ± 0.2

2.4 Conclusion

This work demonstrates a simulation protocol to predict the interfacial tension of water/oil mixtures that can account for the interfacial enrichment of some components of the oil mixture. The TIP4P/2005∥TraPPE–UA combination of force fields can yield fairly accurate results for the water/alkane IFT, whereas the TIP4P/2005∥TraPPE–EH combination significantly overestimates the IFT for water/arene mixtures. Although modifications of the energetic Lennard-Jones cross-interaction parameters can lead to a more accurate prediction of the IFTs for binary mixtures, care should be taken when fitting k_{ij} parameters to account for uncertainties in the experimental data. Furthermore, the current study indicates a weak state-point dependence of the binary k_{ij} parameters. Nevertheless, when the state points for the k_{ij} parameterization are chosen appropriately and long simulation trajectories are used, then mean signed errors for the IFT of the water/n-dodecane and water/toluene mixtures are found to be less than 1 dyn/cm.

With regards to the prediction of the IFT for the water/(50 wt% n-dodecane + 50 wt% toluene) mixture, it needs to be emphasized that the simulations with both sets of k_{ij} parameters yield a value that slightly exceeds the average of the two binary mixtures. That is, we observe a positive excess IFT for the ternary mixture despite the significant enrichment of toluene at the interface, whereas the experimental data demonstrate a negative excess IFT. This observation points to a failure of using non-local k_{ij} parameters to compensate for very local induced polarization and charge transfer effects.

Chapter 3

Molecular simulations probing the thermophysical properties of homogeneously stretched and bubbly water systems[a]

3.1 Introduction

Cavitation refers to the formation of vapor cavities inside a liquid when the pressure drops below the saturated vapor pressure. It is widely observed and is often undesirable. For example, around the blades of a working propeller, bubbles are formed due to localized low-pressure regions. The subsequent collapse of bubbles can generate strong shock waves or high local temperatures that can cause damage to the blades [54]. An image of sheet cavitation on a hydrofoil is shown in Figure 3.1. Understanding and predicting cavitating flows is inherently challenging due to the variety of phenomena involved that can span large temporal and spatial scales. Currently many of these phenomena are modeled empirically or are not included in the field of CFD, and much

[a]This chapter is adapted from the publication where I am the first author: Chen, J. L.; Xue, B.; Mahesh, K.; Siepmann, J. I. Molecular simulations probing the thermophysical properties of homogeneously stretched and bubbly water systems. *J. Chem. Eng. Data* **2019**, 64, 3755—3771. My contribution includes implementation of different simulation protocols, developing an efficient MPI-parallelized program for analyzing properties of bubble, running majority of the simulations, analyzing majority of the thermophysical properties, and writing up the manuscript. Dr. Bai Xue helped with preparing the initial version of the input files for LAMMPS and generating preliminary data.

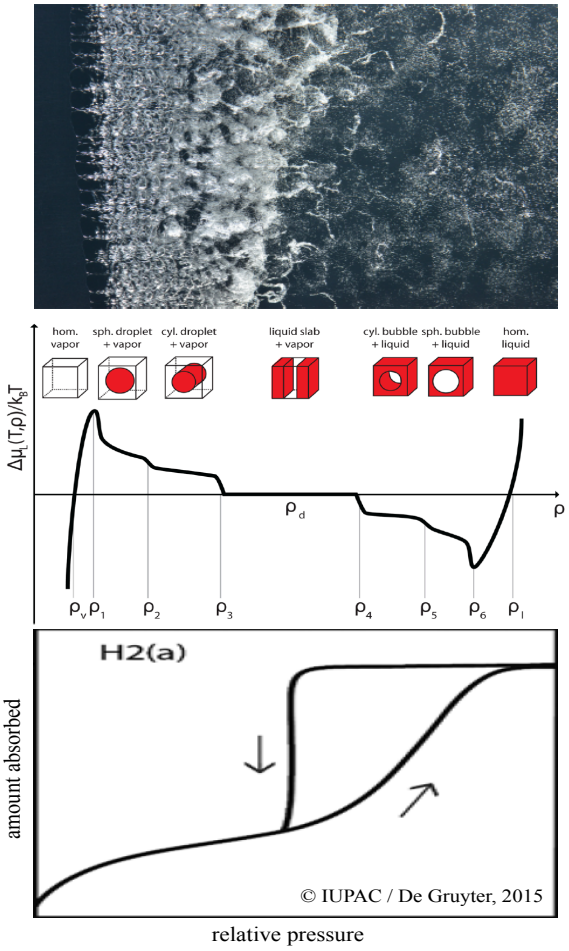


Figure 3.1: (Top) High-resolution photograph of sheet cavitation on a hydrofoil [57]. (Middle) Chemical potential and morphology as a function of density for the LJ fluid; reprinted from Binder, Block, Virnau, and Tröster, *Am. J. Phys.* **2012**, *80*, 1099–1109, with the permission of the American Association of Physics Teachers [58]. (Bottom) Type H2(a) hysteresis loop observed for an adsorption-desorption isotherm in a mesoporous adsorbent with delayed, but steep desorption due to caviatition-induced evaporation [59].

ongoing research is devoted to improving the underlying physical models and eliminating empiricism [55, 56].

Although molecular dynamics simulations can only access much smaller temporal and spatial scales compared to CFD, they can potentially be very useful because of the explicit consideration of the molecular nature of fluids and the ability to probe finite-size model system under conditions that are not stable for bulk systems. In addition, the scales of some physical phenomena involved in the cavitating flow, such as bubble nucleation and growth, are inherently compatible with molecular simulations. In fact, homogeneous bubble nucleation in liquid water has been explored using molecular simulations [60–62].

Molecular simulations can also be utilized to assess some of the underlying equations applied in CFD for predicting cavitating multi-phase flow. For example, in the commonly used homogeneous mixture approach, liquid and vapor regions are assumed to have equal velocity and behave like a single compressible continuum in a given CFD cell (i.e., the vapor and liquid regions are assumed to form a homogenous mixture within this cell) [63]. In this approach, the governing equations describing the conservation laws are written for the homogeneous mixture and an analytic or tabulated equation of state (EOS) for this mixture is needed for system closure [64, 65]. In addition, a set of equations describing the properties of the mixture, such as viscosity, are often needed as well [65, 66]. However, the mixture EOS and property equations utilized in practice are rarely validated through experiments for multi-phase systems [65]. For example, a mixture EOS is usually chosen based on its capability of reproducing the experimentally measured sound speed, yet available experimental data, though already sparse, are mostly concerned with liquid/gas mixtures, and reliable data for liquid/vapor mixtures probably do not exist [67, 68]. Another example is the variety of equations that relate liquid/vapor mixture viscosity with bulk-phase viscosity and vapor-volume fraction [69,70]. These equations are mostly not derived directly from experimental viscosity data but estimated using experimental data for pressure drops of two-phase flow. In these cases, molecular simulation offers an alternative tool for assessing these equations and can potentially lead to better options.

In this work, molecular simulations in the canonical ensemble at $T = 298$ K were performed to probe a variety of thermophysical properties of neat water for densities ranging from the saturated liquid density (ρ_{sat}) to 800 kg/m^3 for periodic simulation boxes ranging from 4 to 32 nm. The lower bound at 800 kg/m^3 was chosen such that all inhomogeneous phases observed in the simulations have a single spherical cavity (denoted as bubbly water phase), rather than a cylindrical cavity or even a slab structure (see Figure 3.1). Binder *et al.* [58] have previously studied the phase behavior of the LJ fluid for densities ranging from below the saturated vapor density to above the satured liquid density for a periodic system with fixed linear dimension. As illustrated in Figure 3.1, the LJ fluid undergoes a transition from the homogeneous liquid phase to the bubbly phase as the density is reduced from ρ_{l} to below ρ_6. In our work, a similar phase behavior was observed for water as well. By following two different simulation protocols,

where one starts with a homogeneously stretched configuration (protocol 1), while the other starts with a spherical cavity (protocol 2), we were able to locate the spinodal cavitation (SC) and bubble collapse (BC) points separately. At these points, the free energy barrier for transition to a more stable phase becomes sufficiently small so that this phase transition occurs rapidly in the early part of the simulation. The differences observed between simulations utilizing protocols 1 (desorption) and 2 (adsorption) are analogous to the hysteresis loops observed for adsorption-desorption isotherms in mesoporous adsorbents [59]. In particular, type H2(a) hysteresis can be attributed to a system with wetting pore walls and delayed desorption due to cavitation-induced evaporation [59] (see Figure 3.1). Similarly, the formation of a bubble can be delayed for simulations in the canonical ensemble due to the free energy barrier for bubble nucleation; i.e., as for the desorption branch, the metastable stretched liquid phase can persist. On the other hand, the vapor-liquid interface of the bubble initially present using protocol 2 acts like a wetting wall, the bubble shrinks as the overall density is increased and then disappears when the homogeneously stretched liquid state is lower in free energy than the bubbly state. It should be noted that the hysteresis loops observed in adsorption-desorption isotherms are experimentally reproducible and "permanent", i.e., increasing the timescale for the sorption measurements from hours to months leads only to negligible changes in the hysteresis loop. Considering the timescale of flow experiments and CFD simulations, it is likely that, in the absence of seeds or surfaces, bubbles form through spinodal decomposition instead of homogeneous nucleation. In this case, the metastable homogeneous states observed with protocol 1 provide important information for the modeling of multi-phase flow.

The thermophysical properties investigated in this work include pressure (P), potential energy (U), residual isochoric heat capacity ($C_{V,\mathrm{res}}$), viscosity (η), and diffusion coefficient (D_{self}). Previous molecular simulation studies have investigated the thermophysical properties of stretched homogeneous water at similar conditions, [71–74] and comparisons are made here to these earlier studies. To our knowledge, the thermophysical properties of two-phase bubbly water have not been investigated previously. Since the spatial scale of the cells involved in CFD is usually microns or larger and, hence, orders of magnitude larger than those amenable to molecular simulation, attention needs to be given to finite-size effects. For simulations in the canonical ensemble, such as the

current work, system-size effects refers to changes of observables with system size (increasing linear dimension of the simulation box) at fixed system density. However, for inhomogeneous bubbly water systems, an increase in system size at fixed density necessarily also requires a corresponding change in bubble radius and volume, whereas the bubble volume fraction is less affected. The results obtained from molecular simulations are compared with some of the available equations utilized in CFD. Analytic equations fitted to simulation data are proposed for describing the properties of the bubbly water state if possible.

3.2 Simulation methods

3.2.1 Force fields

Two types of water models were used in this work. The first one is the rigid TIP4P/2005 water model [75], which consists of a single LJ site on the position of the oxygen nucleus, two equal partial charges on the position of the hydrogen nuclei, and a negative partial charge located at an off-site location along the H-O-H bisector to keep the molecule neutral. The intermolecular interactions are described by pairwise additive LJ 12–6 and Coulomb potentials:

$$u_{ij}(r_{ij}) = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \tag{3.1}$$

where r_{ij}, ϵ_{ij}, σ_{ij}, q_i, q_j, and ϵ_0 are the distance between interaction sites i and j, the LJ well depth, the LJ diameter, the partial charges on interaction sites i and j, and the permittivity of vacuum, respectively. In our simulations, periodic boundary conditions were applied to cubic simulation boxes in all three dimensions. LJ interactions were truncated at a distance $r_{\text{cut}} = 1.40$ nm. Since the bubbly water systems are inhomogeneous, use of analytical tail corrections for LJ energy and pressure that assume a uniform pair distribution beyond the truncation distance, would be inappropriate. Therefore, for consistency, no LJ tail corrections were applied throughout simulations and the same r_{cut} value is used for all system sizes. In the development of the TIP4P/2005 water model [75], the LJ potential was truncated at 0.85 nm and tail corrections were employed. Thus, the larger r_{cut} at 1.40 nm used here mitigates the

truncation error. To compute charge–charge interactions under periodic boundary conditions, the Ewald summation method [76, 77] with a screening parameter $\kappa = \pi/r_{\text{cut}}$ and an upper bound of the reciprocal space summation at $K_{\max} = \text{int}(\kappa L) + 1$ was utilized for Monte Carlo simulations, whereas the particle-particle particle-mesh Ewald summation technique [78, 79] with a relative force error of 10^{-5} was utilized for molecular dynamics simulations. [78, 80] The real-space site–site electrostatic interactions were truncated at r_{cut} as well.

The other water model utilized in this work is the coarse-grained mW water model [81] where each water molecule is represented by a single interaction site. In addition to a two-body interaction term, this water model also introduces a three-body interaction term favoring the tetrahedral packing of water molecules in the condensed phase. The potential takes the form

$$U = \sum_i \sum_{j>i} \phi_2(r_{ij}) + \sum_i \sum_{j\neq i} \sum_{k>j} \phi_3(r_{ij}, r_{ik}, \theta_{ijk}) \tag{3.2}$$

$$\phi_2(r_{ij}) = A\varepsilon \left[B \left(\frac{\sigma}{r_{ij}} \right)^m - \left(\frac{\sigma}{r_{ij}} \right)^n \right] \exp\left(\frac{\sigma}{r_{ij} - a\sigma} \right) \tag{3.3}$$

$$\phi_3(r_{ij}, r_{ik}, \theta) = \lambda\varepsilon(\cos\theta - \cos\theta_0)^2 \exp\left(\frac{\gamma\sigma}{r_{ij} - a\sigma} \right) \exp\left(\frac{\gamma\sigma}{r_{ik} - a\sigma} \right) \tag{3.4}$$

where $A = 7.049556277$, $B = 0.6022245584$, $m = 4$, $n = 0$, $\gamma = 1.2$, $a = 1.8$, $\theta_0 = 109.47°$, $\lambda = 23.15$, $\epsilon = 6.189$ kcal/mol, and $\sigma = 0.23925$ nm. In this model, both two-body and three-body interaction terms become negligible as pair distance approaches $a\sigma = 0.43065$ nm. Due to this short-ranged nature and fewer interaction sites per molecule, simulations typically gain at least 2 orders of magnitude in efficiency compared with the TIP4P/2005 water model, thus allowing for much larger system sizes.

3.2.2 Monte Carlo simulation details

Isochoric–isothermal (NVT) Gibbs ensemble Monte Carlo (GEMC) simulations [82, 83] were performed using MCCCS-MN software [84] developed in house to determine the vapor–liquid equilibrium (VLE) properties (densities of both phases and vapor pressure) at $T = 298$ K. These simulations used two boxes to represent liquid and vapor phases. The total number of molecules N was set to 2000. To ensure a sufficient number of

molecules in the vapor box, the initial box lengths for the liquid and vapor boxes were set to 4 nm and 100 nm, respectively. Due to the large vapor box utilized, r_{cut} was set to 40 nm instead of 1.4 nm for the vapor box to reduce computational cost for the reciprocal-space part of the Ewald sum. For each simulation, the equilibration period consisted of at least 50,000 Monte Carlo cycles (MCCs; one MCC consists of N randomly selected moves), and the production period consisted of 100,000 MCCs. Statistical uncertainties were estimated from the standard error of the mean at 95% confidence interval obtained from production periods of eight independent simulations.

For the TIP4P/2005 water model, the phase space was sampled via translational moves [85], rotational moves [86], volume exchange moves [83], and particle transfer moves between phases using the dual cut-off configurational-bias Monte Carlo algorithm [45, 87] with optimized control parameters and growth protocol. The maximum displacements for translational, rotational, and volume moves were adjusted during equilibration to yield acceptance ratios of approximately 40%. To increase sampling efficiency and also to avoid unphysical "fused" configurations where a positively charged interaction site of one water molecule is nearly co-located with the negatively charged site of another water molecule, a hard-sphere potential with $r_{\text{HS}} = 0.12$ nm was applied for all site-site interactions. For the mW water model, phase space was sampled via translational, volume exchange, and particle transfer moves. A hard-sphere potential with $r_{\text{HS}} = 0.05$ nm was utilized as well to improve sampling efficiency and to avoid underflow errors in the calculation of Boltzmann weights.

3.2.3 Molecular dynamics simulation details

Molecular dynamics simulations were carried out using the LAMMPS software package. [88] All simulations were performed in the NVT ensemble. The time step was set to 1 fs or 5 fs for the TIP4P/2005 or mW water models, respectively. The temperature was maintained at 298 K using the Nose–Hoover thermostat [89], and its relaxation time was set to 100 fs. For the rigid TIP4P/2005 water model, the O–H bonds and H-O-H angle were constrained using the SHAKE algorithm [90].

Simulations were first performed to calculate the surface tension for the TIP4P/2005 and mW water models. Following the Kirkwood and Buff definition [91], the surface tension was calculated from the ensemble averages of the normal and tangential com-

ponents of the pressure tensor. The simulation procedure consisted of two steps. First, a stable liquid configuration was obtained by performing simulations in the NVT ensemble with $N = 1024$ and a cubic box of length $L = 3$ nm. Then the simulation box was elongated in the z dimension to $L_z = 10$ nm, and equilibration was continued for at least 0.5 ns. Statistics were gathered from production periods of 2 ns.

Next, simulations were performed to probe a variety of thermophysical properties of both homogeneously stretched liquid and bubbly water phase. The box lengths (L) studied were 4, 8, 16, and 32 nm for the mW water model (and these simulations are designated here as mW-4, mW-8, mW-16, and mW-32, respectively), containing up to 1,092,850 molecules in the largest system, and 4, 6 and 8 nm for TIP4P/2005 model (desginated as TIP4P/2005-4, TIP4P/2005-6, and TIP4P/2005-8, respectively). For each box size, densities ranging from 800 kg/m^3 to ρ_{sat} were explored. The densities and number of molecules for all the systems studied here are provided in Table SB1.

Two simulation protocols were utilized. In protocol 1, a homogeneous configuration was used as the starting point. The initial structure for the mW water model was generated randomly using LAMMPS. Then the potential energy was minimized by adjusting the coordinates of the molecules iteratively for a maximum of 1000 steps. For the TIP4P/2005 water model, the initial structure was generated using the Packmol program with a tolerance distance of 0.2 nm [92,93]. After that, for both water models, the system underwent a brief melting and a subsequent cooling stage. In protocol 2, the initial heterogeneous configuration was generated such that all molecules were placed outside of a sphere with its center located at the center of the simulation box and a radius ranging from 0.5 to 1.5 nm. For the single-site mW water model, this was achieved with built-in commands in LAMMPS, while Packmol was utilized for the TIP4P/2005 water model. Next, a spherical wall was created in LAMMPS, that interacts with the water molecules through a 12-6 LJ potential and prevents any molecules from entering the cavity during the pre-equlibration period. After a short melting and subsequent cooling stage, the wall was removed and the system was allowed to equilibrate.

For the mW-4, mW-8, and mW-16 systems, statistics were collected from a production period of 20 ns. For the mW-32 system, statistics were collected from five independent runs, each consisting of a production period of 4 ns. For the TIP4P/2005-4 and TIP4P/2005-6 systems, statistics were collected from a production period of 4 ns.

For the TIP4P/2005-8 system, statistics were collected from five independent runs, each consisting of 1.6 ns. For both water models, the independent runs were created by using different random number seeds for generating initial structures and velocity distributions. A detailed description of how the uncertainties were estimated for the various quantities is given in the following sections.

3.2.4 Bubble volume and sphericity

For bubbly water systems, the bubble volume for a given configuration along the trajectory was determined as follows. First, Stillinger's cluster criterion [94] was utilized to calculate the number of nearest neighbors of each water molecule where two molecules are considered nearest neighbors if their distance (oxygen–oxygen for TIP4P/2005 and site–site for mW) is less or equal to 0.33 nm; a distances that is approximately the position of the first minimum in the O–O radial distribution function for the saturated liquid phase of the TIP4P/2005 model. If a molecule has two or more nearest neighbors, it is then considered to be liquid-like, otherwise it is counted as vapor-like. Next, a three-dimensional grid with a cell mesh of $(0.2\ \text{nm})^3$ was imposed onto the simulation box. If a cell contains a liquid-like water molecule or its center falls within 0.33 nm of at least two liquid-like molecules, then it is considered to be liquid-like, and the remaining cells are defined as vapor-like. Finally, a cluster analysis was performed considering only the vapor-like cells, and two cells are part of the same cluster when they share at least one corner. The detected number of clusters is sometimes larger than one for a single configuration due to the existence of very small voids being formed by thermal fluctuations inside the liquid region. Only the largest cluster, whose size is always significantly larger than any of the transient liquid voids, is taken as the bubble and used to calculate the bubble volume, V_{cell}, from its number of cells. Since this analysis is based on individual configurations, the statistical uncertainty in the bubble volume fraction is reported here as the standard deviation obtained from the values for individual configurations. The cell method for bubble detection described here is computationally more efficient than the V- and M-methods proposed by González *et al.* [61], and a simpler method is justified because the current work deals with much larger systems (requiring a more efficient method) and much larger bubbles (being less sensitive to the method for bubble detection).

To characterize bubble sphericity, the gyration tensor was calculated as

$$S = \frac{1}{N_{\text{vap}}} \begin{bmatrix} \sum_i (x_i - x_{\text{C}})^2 & \sum_i (x_i - x_{\text{C}})(y_i - y_{\text{C}}) & \sum_i (x_i - x_{\text{C}})(z_i - z_{\text{C}}) \\ \sum_i (y_i - y_{\text{C}})(x_i - x_{\text{C}}) & \sum_i (x_i - x_{\text{C}})^2 & \sum_i (y_i - y_{\text{C}})(z_i - z_{\text{C}}) \\ \sum_i (z_i - z_{\text{C}})(x_i - x_{\text{C}}) & \sum_i (z_i - z_{\text{C}})(y_i - y_{\text{C}}) & \sum_i (z_i - z_{\text{C}})^2 \end{bmatrix} \tag{3.5}$$

where x_{C}, y_{C}, and z_{C} give the position of the center of mass of the bubble. The summation runs only over the vapor-like cells belonging to the bubble, where N_{vap} is the number of these cells. Then, the gyration tensor is diagonalized, and its three eigenvalues λ_1, λ_2, and λ_3 are obtained. The relative shape anisotropy is defined as [95]

$$\kappa^2 = 1 - 3\frac{\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_1\lambda_3}{(\lambda_1 + \lambda_2 + \lambda_3)^2} \tag{3.6}$$

κ^2 is bounded between 0 and 1 corresponding to a spherically symmetric shape and a linear shape, respectively.

3.2.5 Thermophysical properties

The pressure of the system was calculated based on the virial theorem of Clausius [96,97]

$$p = \rho k_{\text{B}} \langle T \rangle + \frac{1}{3V} \left\langle \sum_i \mathbf{r}_i \cdot \mathbf{f}_i \right\rangle \tag{3.7}$$

where k_{B} is the Boltzmann constant, ρ, V, and T are the number density of molecules, volume, and absolute temperature of the system, respectively. For the molecular dynamics simulations, T is calculated based on the kinetic energy of the system. $\mathbf{r}_i$ and $\mathbf{f}_i$ are the position vector and the force exerted on atom i. The angular brackets in the first and second terms denote time averages (or, only for the second term, an ensemble average for the Monte Carlo simulations) of temperature and the internal virial, respectively.

The isochoric heat capacity can be split into ideal and residual parts:

$$C_V(T, V) = C_{V,\text{ideal}}(T) + C_{V,\text{res}}(T, V) \tag{3.8}$$

where $C_{V,\text{ideal}}$ is a property of a collection of isolated molecules and can be determined from calorimetric measurements or the molecular partition function with input either

from spectroscopic data or quantum-mechanical calculations [98, 99]. $C_{V,\text{ideal}}$ includes the kinetic contributions. The residual part is determined from fluctuations of the intermolecular potential energy of the system. However, for the rigid TIP4P/2005 and the single-site mW models, there are no intramolecular contributions. Therefore, the potential energy (U) of the system can be directly used:

$$C_{V,\text{res}} = \frac{\langle U^2 \rangle - \langle U \rangle^2}{k_\text{B} T^2} \tag{3.9}$$

The shear viscosity was calculated using the Green-Kubo approach [100],

$$\eta = \frac{V}{k_\text{B} T} \int_0^\infty dt \langle P_{\alpha\beta}(0) P_{\alpha\beta}(t) \rangle \tag{3.10}$$

where $P_{\alpha\beta}(0)$ and $P_{\alpha\beta}(t)$ are the off-diagonal components of the pressure tensor at time 0 and t, and are averaged over three components P_{xy}, P_{yz}, and P_{xz}. The viscosity calculated this way for an inhomogeneous system is the averaged value over the entire box [101]. Since the auto-correlation function in the above equation often diverges due to accumulation of noise at long time, a time decomposition method was utilized in this work [102]. To make sure the viscosity value has converged, the integral was truncated at 1.0 ps for the mW model and 3.0 ps for the TIP4P/2005 model. The justification for these choices of the time cutoff is presented in Figure SB1.

For the estimation of the statistical uncertainties for pressure, potential energy, residual isochoric heat capacity, and shear viscosity, the total production period of each simulation was divided into 8 blocks and the block means were calculated. The uncertainties were estimated as the standard error of the mean at 95% confidence interval.

Finally, the diffusion coefficient was obtained through the mean squared displacement (MSD) of the center of mass of molecules and the Einstein relationship, [77]

$$D = \frac{1}{6t} \left\langle \frac{1}{N} \sum_{i=1}^{N} [\mathbf{r}_i(t) - \mathbf{r}_i(0)]^2 \right\rangle \tag{3.11}$$

where $\mathbf{r}_i(t)$ and $\mathbf{r}_i(0)$ are the position vectors of the center of mass of molecule i at time t and 0. To calculate the uncertainties for the diffusion coefficient, each trajectory was divided into five blocks. For the largest system, mW-32 and TIP4P/2005-8, each independent simulation was considered as one block. The lengths of each block (4 ns for

Table 3.1: Calculated and Experimental Orthobaric Liquid Density (ρ_{liq}), Vapor Density (ρ_{vap}), Saturated Vapor Pressure (P_{sat}), Enthalpy of Vaporization (ΔH_{vap}), Liquid–Vapor Surface Tension at Saturation (γ), and Their Uncertainties (u) at $T = 298$ K

Property		TIP4P/2005		mW		Expt. [103]
		Value	u	Value	u	
ρ_{liq}	[kg/m^3]	994.4	1.3	997.66	0.19	997.04
ρ_{vap}	[g/m^3]	5.39	0.13	0.3561	0.0051	22.881
P_{sat}	[kPa]	0.739	0.017	0.04853	0.00078	3.1
ΔH_{vap}	[kJ/mol]	50.164	0.042	44.432	0.062	44.01
γ	[mN/m]	65.9	1.7	65.47	0.24	71.995

Statistical uncertainties were estimated from the standard error of the mean at 95% confidence interval obtained from production periods of eight independent simulations.

the mW-4, mW-8, and mW-16 systems, 0.8 ns for the TIP4P/2005-4 and TIP4P/2005-6 systems, and 1.6 ns for the TIP4P/2005-8 system) were chosen to be sufficiently long such that, on average, molecules have diffused more than half of the box length. The diffusion coefficient for the mW-32 system was not calculated because, although a linear region is present in the MSD plot, the trajectory length of 4 ns is insufficient to reach an MSD value of (16 nm)^2. The diffusion coefficient for each block was obtained from the slope of the least-squares linear fitting line between MSD and time (where the first 10% and last 10% of data points were thrown out in the fitting to avoid the initial ballistic region and the relatively poor statistics at the end). An example MSD plot obtained from the five blocks for the TIP4P/2005-8 system is shown in Figure SB2. The statistical uncertainties provided are the standard error of the mean at 95% confidence interval obtained from the 5 blocks.

3.3 Results and discussion

3.3.1 Vapor–liquid equilibrium properties and surface tension

The predicted VLE properties and surface tension at $T = 298$ K are listed in Table 3.1. The liquid-phase density at saturation for both types of water models agrees well with experiment. Both water models predict a surface tension value which is approximately 7% smaller than the experimental value, and the values obtained here are consistent with previous simulation results [81, 104]. However, in terms of vapor pressure and vapor density at saturation, both water models underestimate by a significant amount.

The TIP4P/2005 water model yields predictions that are a factor of 4.2 smaller than the experimental values, which is due to an overestimation of the enthalpy of vaporization when a polarization correction is not applied [75, 105]. The mW water model underestimates them by a factor of 64. Factorovich et al. [106] also computed the vapor pressure for the mW water model at $T = 298$ K using both GEMC and grand canonical molecular dynamics (GCMD) simulations and reported values of 45.8 and 48.6 Pa, respectively. Their GCMD value is more consistent with the value obtained here from a GEMC simulation. Interestingly, the mW model yields a fairly accurate value for the enthalpy of vaporization [81] and, hence, its significant underestimation of the saturated vapor pressure is caused by an underestimation of the entropy gain upon vaporization. The fact that a model can yield a very accurate liquid density and a fairly accurate surface tension (and, for the mW model, also enthalpy of vaporization), but fail dramatically for the saturated vapor pressure, indicates that the former two properties are not sufficient for force field development.

3.3.2 Spinodal cavitation and bubble collapse points

An important observation for systems with $L \geq 6$ nm is that the transition point between the homogeneously stretched water phase and the bubbly water phase occurs at a lower density for the set of simulations starting from a homogeneous configuration, where the system has to overcome the nucleation free energy barrier for bubble formation, compared to those simulations starting with a pre-existing cavity, where no (or only a small) free energy barrier for bubble collapse is present. These two transition processes and their associated densities are denoted as spinodal cavitation (SC) and bubble collapse (BC), respectively. It should be noted that, for the finite-size systems investigated here, the BC process occurs with a downward jump in pressure and leads to a homogeneous phase with a negative pressure, i.e., it is distinct from the vapor–liquid binodal point that denotes the coexistence of two stable macroscopic phases with positive pressure. While analyzing the bubble volume fraction, we observed that when the SC density is crossed, there is always a significant jump of the volume fraction of the largest void/bubble from less than 0.2% due to small transient liquid voids in the homogeneously stretched water phase to more than 1% when a bubble is formed. Similarly, for simulations using protocol (2), we find that there is a certain minimum bubble

size that can be observed in the simulations. This minimum bubble size found in the simulations increases with system size and the bubble volume fraction is always significantly larger than 1%. We surmise that at the BC point, the free energy of the bubbly phase with this minimum bubble size is equal to that of the homogeneously stretched phase. Any reduction of the bubble size beyond this limit cannot be observed because this state would have a free energy higher than the homogeneously stretched phase at the same density; i.e., the system prefers homogeneous stretching (with a more negative pressure) over affording an interface (and less stretching of the liquid region). Since larger systems have to stretch by a smaller relative amount to consume a given bubble volume, the minimum observable bubble size increases with increasing system size.

Therefore in this work, a system is considered to be a bubbly water phase only when the bubble volume fraction is larger than 1%. Similar significant and discontinuous changes were also observed in other thermophysical properties near the same transition points and will be discussed in the following sections.

For the two smallest systems (mW-4 and TIP4P/2005-4) that contain about 2000 molecules, distinct SC and BC points were not observed with a density resolution of 5 kg/m^3. At the intermediate density (930 kg/m^3 for mW-4 and 880 kg/m^3 for TIP4P/2005-4) between homogeneously stretched and bubbly water phases, the systems were found to alternate between homogeneous and inhomogeneous configurations (as defined by the 1% threshold in the bubble volume fraction) causing large fluctuations also in other observables. Due to the large fluctuations arising from the bimodal distributions, thermophysical properties at these two state points are not reported here.

The Helmholtz free energies as function of the largest void/bubble volume for the mw-4 and TIP4P/2005-4 systems at these state points are shown in Figure SB3. The barrier separating the two states is found to be 10-12 kJ/mol (4-5 RT) and its location at $V \approx 0.6$ nm^3 is aligned with the 1% threshold used here to separate the two states. Menzl *et al.* [62] have used simulations in the isobaric-isothermal ensemble at $N = 2000$ and $T = 296.4$ K with enhanced sampling techniques to probe the cavitation of TIP4P/2005 water under tension. Our canonical ensemble simulations for TIP4P/2005-4 correspond to a similar system size, and our observation of the SC point falling into a pressure range from -190 to -200 MPa is consistent with the data of Menzl *et al.* that show nucleation free energy barriers of 30 and 20 RT at $P = -150$ and -165 MPa, respectively.

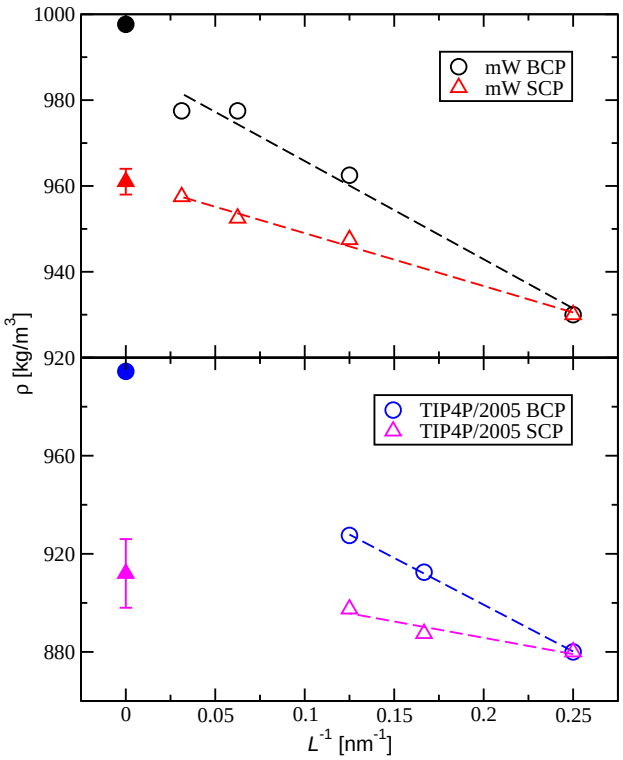


Figure 3.2: Specific densities of the SC and BC points for the mW and TIP4P/2005 systems as a function of inverse linear box dimension. The dashed lines indicate linear fits. The filled circles and triangles denote ρ_{liq} and the estimate of the SC density for the thermodynamic limit.

The specific densities of the SC and BC points are plotted as a function of inverse linear box dimension in Figure 3.2. Due to the finite resolution of the density grid (increments of 5 kg/m^3), the exact values of the spinodal and binodal points are not known and, hence, they are reported here as the average of the specific densities just below and just above of these points for $L \geq 6$ nm or as the density leading to large fluctuations in the bubble volume fraction for mW-4 and TIP4P/2005-4. González *et al.* [107] have previously used simulations in the canonical ensemble ($N = 500$) to estimate the vapor–liquid spinodal line for the TIP4P/2005 model and, from their data, a value of 850 kg/m^3 at $T = 298$ K can be deduced. In comparison, we find a value of 880 kg/m^3 for TIP4P/2005-4 that contains about 3.8 times more molecules (i.e., 1.6 times larger in linear dimension).

Another observation from Figure 3.2 is that, as system size increases, both SC and BC points shift to larger densities (except mW-16 and mW-32 yield the same binodal density due to the finite grid resolution). This behavior is consistent with the system size effects found by Binder *et al.* [58] for the LJ fluid and also with the upward shift

in the position of the hysteresis loop for adsorbents with larger mesopores [59]. Upon further increase in the system size, the BC point would move toward the saturated liquid density (i.e., the true binodal point), but the density domain for stable spherical bubbles also diminishes. In the thermodynamic limit, only the slab configuration with planar interfaces is stable at densities below the saturated liquid density (see Figure 3.1) [58]. From extrapolation of our data for the SC points of the finite systems, we estimate spinodal densities of 912 ± 14 and 961 ± 3 kg/m^3 for the TIP4P/2005 and mW models, respectively, in the thermodynamic limit.

For the systems with $L \geq 6$ nm, the simulations started from a homogeneous configuration at densities below the BC point but above the SC point are metastable. The definition of a metastable state inherently involves time [108]. In the present simulations, we found that all simulations in the metastable region remained in a homogeneously stretched state throughout the entire simulation trajectories, and the same holds for the five independent simulations for largest system size (mW-32 and TIP4P/2005-8). The production periods for the mW-8, mW-16, mW-32, TIP4P/2005-6, and TIP4P/2005-8 systems were 20, 20, 4, 4, and 1.6 ns, respectively, with the equilibration periods being of various lengths but generally of the same order of magnitude as the production periods. Clearly, the nucleation free energy barrier increases rapidly for densities just above the spinodal point, so that nucleation does not occur in our unbiased simulations. This is analogous to the observation that increasing the observation time does not lead to a significant upward shift in the pressure of the desorption onset for the type H2(a) hysteresis loop (see Figure 3.1) [59]. The Helmholtz free energies as function of the volume of the largest void/bubble for the mW-8 and TIP4P/2005-8 systems are shown in Figure SB4. The unbiased simulations used here only allow for observation of states with relative free energies less than 20 kJ/mol (8 RT). Extrapolating the free energy profiles for the metastable homogeneously stretched systems with the lowest density (i.e., closest to the SC point) yields a nucleation free energy barrier of 25 and 35 kJ/mol (10 and 14 RT) for TIP4P/2005-8 and mW-8, respectively. For both systems, the steepness of the free energy profiles increases rapidly with increasing system density. The estimate for the free energy barrier for TIP4P/2005-8 with $N = 15398$ and $P = -168$ MPa is considerably smaller than the 20 RT reported by Menzl *et al.* [62] for $N = 2000$ and $P = -165$ MPa; thus, finite-size effects are significant for the nucleation barrier where

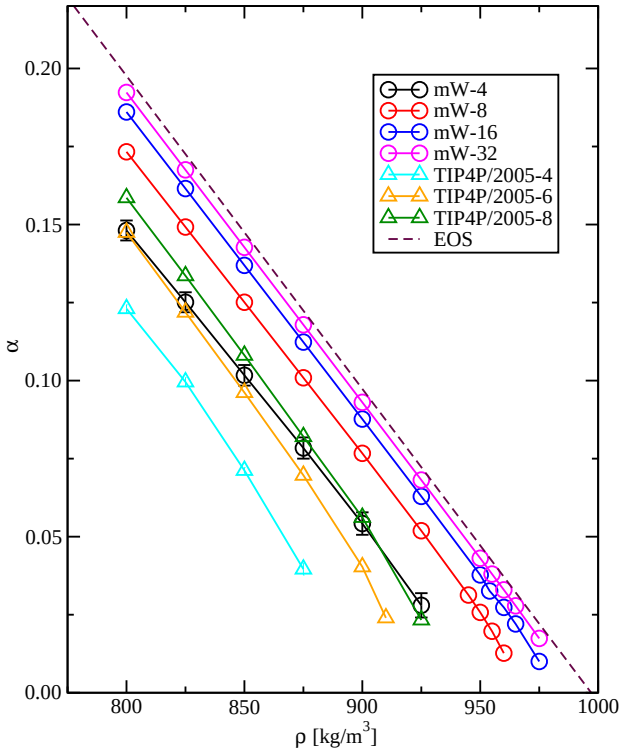


Figure 3.3: Bubble volume fraction (α) versus system density. The brown dashed line represents the analytic equation $\rho = (1-\alpha)\rho_{\text{liq}} + \alpha\rho_{\text{vap}}$, where ρ_{liq} and ρ_{vap} are taken from experiment. Not shown are data points for the homogeneously stretched region where $\alpha \ll 0.01$. For all but the mW-4 system, the statistical uncertainties are smaller than the symbol size.

smaller system sizes lead to larger barriers and a lower density for the SC point.

For the bubbly phases, the width of the free energy wells remains remarkably constant even as the BP density is approached; i.e., we do not observe a significant change in curvature of the free energy profile that would hint at the emergence of the homogenously stretched phase. On the other hand, the simulations using protocol (2) for the system density just above the BC point indicate that there is no thermodynamic barrier for the collapse.

3.3.3 Bubble volume and sphericity

The calculated bubble volume fraction (α) versus system density for all bubbly water phases is plotted in Figure 3.3 (numerical data are provided in Tables SB2–SB15). The line predicted by the simple analytic equation, $\rho_{\text{mix}} = (1-\alpha)\rho_{\text{liq}} + \alpha\rho_{\text{vap}}$ is also shown in Figure 3.3 for comparison, where ρ_{liq} and ρ_{vap} are the experimental liquid and vapor coexistence densities. This analytic equation, which ignores any surface and pressure

effects, is commonly utilized in predicting multi-phase flow in CFD [65].

As shown in Figure 3.3, α at a specific value of ρ is influenced by both system size and choice of water model. As the system size increases, the line predicted by the simulations for both the mW and TIP4P/2005 model is shifting to higher α values, i.e., approaching the prediction of the analytic equation. For the same system size, the α values for the TIP4P/2005 water model fall below those for the mW water model, and the lines for the TIP4P/2005 model have a slightly steeper slope. Given the magnitude of the differences between the simulation data and the analytic equation and also the presence of finite-size effects, it is clear that the deviations cannot be attributed to differences in ρ_{liq} for the bulk systems (see Table 3.1). Further simplifying the analytic equation by dropping the vapor term because ρ_{vap} is orders of magnitude smaller than ρ_{liq} (by factors of 4.4×10^4, 2.8×10^6, and 1.8×10^5 for experiment and the mW and TIP4P/2005 water models, respectively) and rearranging yields $\alpha = 1 - \rho/\rho_{\text{liq}}$ with a slope of $-1/\rho_{\text{liq}}$ and an x-intercept of ρ_{liq}. The simulation data in Figure 3.3 can then be described by replacing ρ_{liq} for the bulk system with a smaller $\rho_{\text{liq,eff}}$ for the bubbly phase to account for surface effects. The values of the x-intercepts indicate that $\rho_{\text{liq,eff}}$ decreases with decreasing system size and is smaller for the TIP4P/2005 model than the mW model. However, evaluating the data in Figure 3.3 with a finer lense indicates failure of this simplified analytic equation because (i) there is a noticeable deviation from the linear behavior at smaller α values particularly for the smaller system sizes and (ii) comparison of the data for the mW-4 and TIP4P/2005-8 systems shows a smaller slope but larger intercept for the mW-4 system. Thus, a more microscopic analysis is needed to determine the density of the liquid region for the bubbly water phases. Furthermore, surface effects neglected in the analytic equation may also play a role.

The liquid density in the bubbly water phase can be obtained by fitting the radial shell density profile $\rho(r)$ to a hyperbolic tangent function

$$\rho(r) = \frac{1}{2}\left(\rho_{\text{liq,fit}} + \rho_{\text{vap,fit}}\right) + \frac{1}{2}\left(\rho_{\text{liq,fit}} - \rho_{\text{vap,fit}}\right)\tanh\frac{2(r - r_0)}{d} \tag{3.12}$$

where r is the radial distance from the center of mass of the bubble; r_0 and d are the radial position and thickness of the interface, respectively [109]. This radial analysis assumes a predominantly spherical shape of the bubble, an issue that will be addressed

below. The location of the inflection point in the hyperbolic tangent function (r_0) can also be used as an estimate of the bubble radius. The shell density profiles for the mW-32 and TIP4P/2005-4 systems at $\rho = 800$ kg/m^3 are shown in Figure SB5. The radial shell density profiles also allow for an estimation of the fraction of molecules present within the interfacial region

$$\alpha_{\text{int}} = \frac{1}{N} \int_{r_0-d/2}^{r_0+d/2} 4\pi r^2 \rho(r) dr \tag{3.13}$$

The uncertainties for $\rho_{\text{liq,fit}}$, r_0, and α_{int} were estimated as the standard error of the mean at 95% confidence interval from five shell density profiles, each averaged over 200 frames. The numerical values are provided in Tables SB2–SB15.

Figure 3.4 depicts $\rho_{\text{liq,fit}}$ and α_{int} as a function of system density for all bubbly water phases. At fixed density, as the system size increases, $\rho_{\text{liq,fit}}$ is getting closer to the value of the bulk saturated liquid density. Interestingly, for a given system size and model, $\rho_{\text{liq,fit}}$ increases as ρ decreases, but decreasing ρ corresponds to heterogeneous phases with larger bubbles. The values of $\rho_{\text{liq,fit}}$ range from 927 kg/m^3 ($= 0.932\,\rho_{\text{liq}}$) for TIP4P/2005-4 at $\rho = 875$ kg/m^3 (smallest system size and ρ closest to the BC point) to 996.5 kg/m^3 ($= 0.9988\,\rho_{\text{liq}}$) for mW-32 at $\rho = 800$ kg/m^3 (largest system size and lowest value of ρ). The differences in $\rho_{\text{liq,fit}}$ for the bubbly phases between the two models at the same ρ and system size are much larger than the 3 kg/m^3 difference in ρ_{liq} for the bulk liquid. Since the differences in the bulk vapor–liquid surface tension are also very small, differences in the isothermal compressibility for the stretched liquid and in the ability to accommodate high-curvature interfaces must be responsible for the differences between the mW and TIP4P/2005 models. The fraction of molecules at the interface (see Figure 3.4) decreases with increasing system size at the same ρ (whereas α increases) because the number of molecules in the liquid region grows much faster than the interfacial area. At the same ρ and system size, the α_{int} values are quite similar (with the exception of the data point for the TIP4P/2005 model closest to the BC point) for the mW and TIP4P/2005 models. Therefore, the differences in $\rho_{\text{liq,fit}}$ are mostly responsible for the shift of the α versus ρ lines in Figure 3.3.

One important question is whether the nano-bubbles present in the simulations (ranging in radius from about 0.8 to 11 nm) deviate significantly from a spherical shape

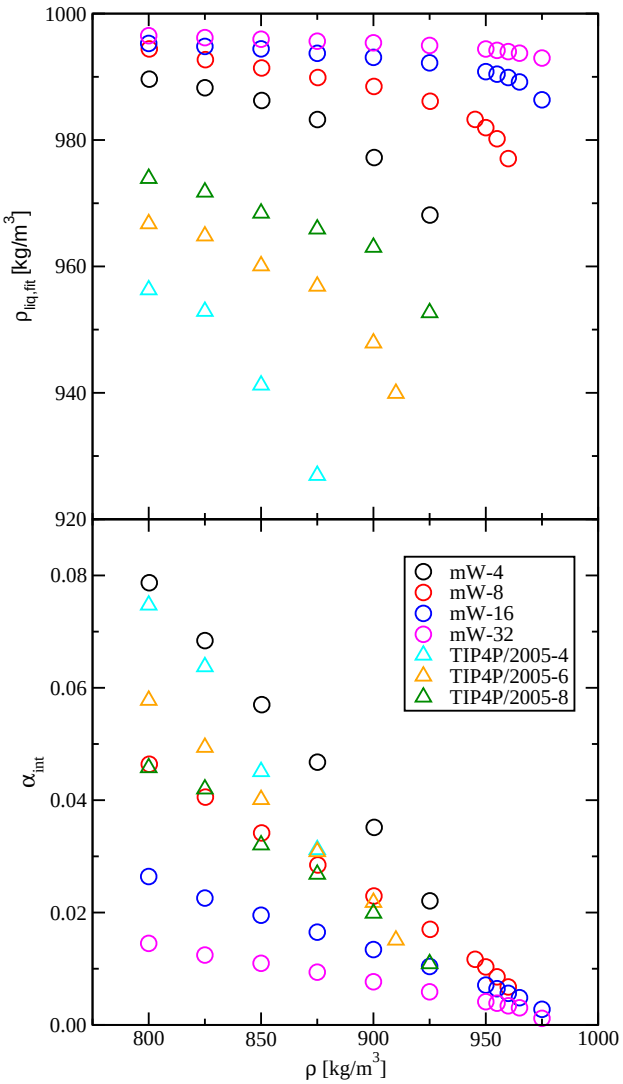


Figure 3.4: Density of liquid region ($\rho_{\mathrm{liq,fit}}$) (top) and fraction of interfacial molecules (α_{int}) (bottom) determined from the radial shell density profile as functions of system density ρ.

due to enhanced fluctuations for such small sizes (compared to bubbles observed in flow experiments, see Figure 3.1). The dependence of the shape anisotropy factor (κ^2) on system density and size is shown in Figure 3.5 (and numerical data are provided in Tables SB2–SB15). The TIP4P/2005-4 and mW-4 simulations yield the largest κ^2 values, but even here the $\kappa^2 < 0.06$, i.e., a very small shape anisotropy. The values for mW-4 and mW-8 at the same density are somewhat smaller than for TIP4P/2005-4 and TIP4P/2005-8, respectively, but these do not correspond to bubbles of exactly the same size due to the differences in $\rho_{\mathrm{liq,fit}}$ between the two models. As the linear dimension of the system increases by a factor of 2, κ^2 decreases by about a factor of 4 at the same density. Thus, in all cases, the bubbles are sufficiently spherical in shape that a sphere

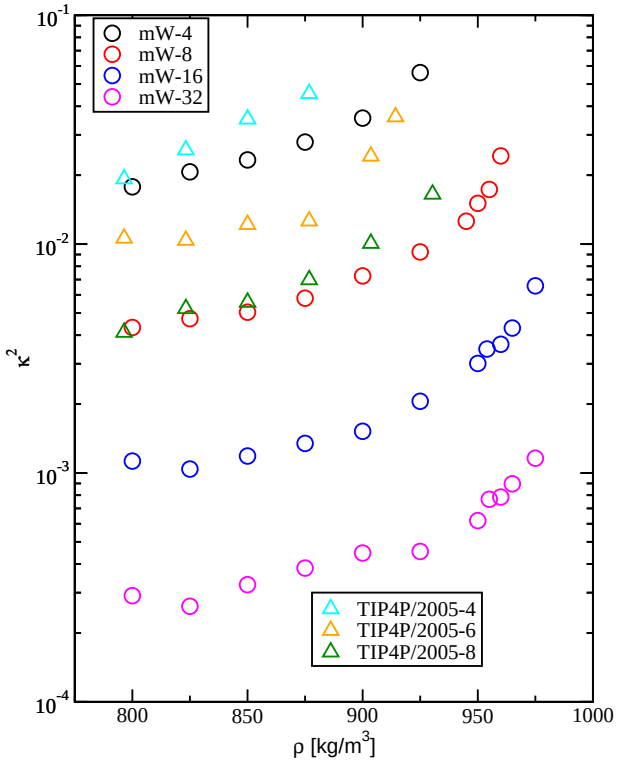


Figure 3.5: Shape anisotropy factor as function of system density.

radius can also be obtained from $r^* = (3V_{\text{cell}}/4\pi)^{1/3}$. The numerical values of r^* are provided in Tables SB2–SB15. The dependence of κ^2 on r^* is shown in Figure SB6. In this case, all data points fall on the same master curve; that is, the dependence on model is much smaller than on r^*, i.e., the bubble size governs the extend of the shape fluctuations.

Figure SB7 shows a comparison of r_0 values obtained from the shell density profiles to the r^* values calculated from the bubble volumes. These data can be well described by a linear fit, $r_0 = kr^* + d$. Irrespective of model and system size, regression coefficients are greater than 0.9999, indicating an exceptional correlation between r_0 and r^*. The slopes for all combinations of model and system size are unity with a maximum deviation of 1% for the TIP4P/2005-4 system. In terms of the offset d, the values for all mW system sizes fall between 0.08 and 0.09 nm. This consistent offset can be explained by the difference in the analysis approaches. The density profiles treat atoms as point particles without any spatial extent. In contrast, the cell method for determining the bubble volume correctly accounts for the volume taken up by the water molecules (excluded volume due to repulsive part of the interaction potential). Hence, we argue that r^* is a better descriptor of the true volume occupied by the bubble and available to vapor

or permanent gas molecules, whereas r_0 is a better descriptor of the thermodynamic location of the interface.

3.3.4 Pressure

The virial pressure (P) of the systems versus total density is displayed in Figure 3.6 (numerical data are provided in Tables SB2–SB15). The pressure is negative for all systems except for the saturated liquid. Please note that the finite number of molecules used in the canonical-ensemble simulations leads to small variations in the system density (see Tables SB2–SB15) that cause some scatter in P for $\rho \approx \rho_{\mathrm{liq}}$. González *et al.* [107] have also calculated the pressure of homogeneously stretched liquid water for the TIP4P/2005 water model using NVT simulations ($N = 500$, $T = 296.4$ K) and reported values of -187 and -207 MPa at 893 and 878 kg/m^3, respectively. In comparison, our values of -168 and -189 MPa for the TIP4P/2005-4 system at 900 and 885 kg/m^3, respectively, ($N = 1925$ and 1892, respectively) yield a slightly less negative pressure. Furthermore, our TIP4P/2005-4 simulation at 875 kg/m^3 spontaneously forms a bubbly phase (resulting in an upward jump in pressure), and the larger TIP4P/2005-8 system yields a bubbly phase already at 910 kg/m^3. However, considering differences in the simulation parameters (number of molecules, cut-off distance used for Lennard-Jones interactions, Ewald sum control parameters, and method to constrain the geometry of each molecule), these differences in pressure are not unexpected.

Among all combinations of water model and system size, the qualitative trends for the relationship of system pressure and density are consistent. Comparing the two simulation protocols (decreasing the density starting from a homogeneous phase versus increasing the density starting from a bubbly phase), the overall shape of the pressure versus density curve is very similar, but the path of decreasing density allows for a deeper penetration into the homogeneously stretched region and more significant (in magnitude) upward jump of the pressure as the SC point is passed than the downward jump in pressure as the BC point is passed. For convenience, the discussion here and also for the other thermophysical properties starts with the path of decreasing density from ρ_{sat} to 800 kg/m^3 (protocol 1), then turns to the path of increasing density from 800 kg/m^3 to ρ_{sat} (protocol 2).

The data in Figure 3.6 show that, as the system density is decreased from the satu-

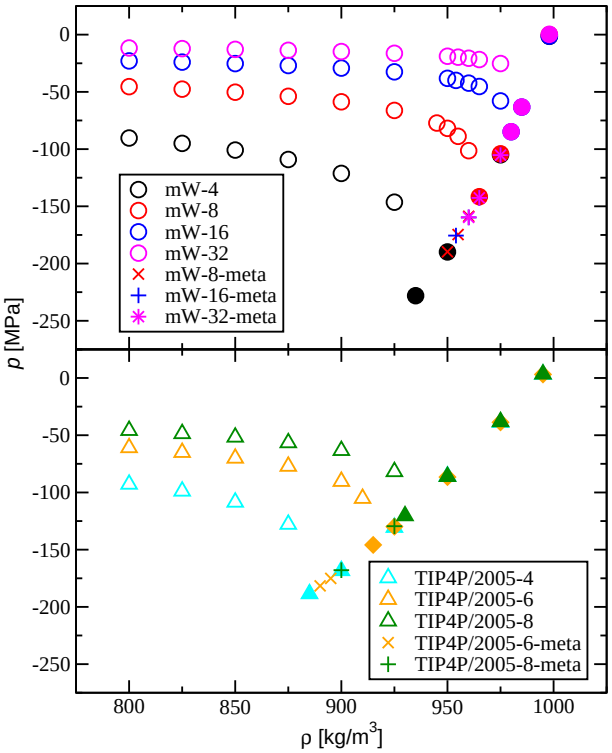


Figure 3.6: System pressure versus density. Open and filled symbols represent bubbly and homogeneously stretched water systems, respectively. The plusses, crosses, and stars show data for metastable homogeneous systems. Uncertainties are all smaller than the symbol size.

rated liquid density, the pressure initially decreases approximately linearly to very large negative values. Comparing the behavior of the pressure for the mW and TIP4P/2005 systems at the same system size, it is apparent that the rate at which P decreases with decreasing ρ in the homogeneously stretched region, the SC density, and the maximum tension that the system can sustain before cavitation are all three smaller in magnitude for the TIP4P/2005 model than for the mW model. After passing the SC point, P jumps upward to a much less negative value, and then enters a region where it slowly increases with decreasing density (increasing bubble volume), but always remains negative. The negative pressure signals that the system could lower its energy through a decrease in the system volume ($P = -(\partial U/\partial V)_T$) that would lead to disappearance of the bubble and less stretching for the homogeneous phase but, of course, the volume of the simulation box is constrained in the canonical ensemble (as is also the cell volume in most CFD simulations).

While the values of P in the homogeneously stretched state at a given ρ do not depend significantly on system size, there is a clear system-size dependence for P in

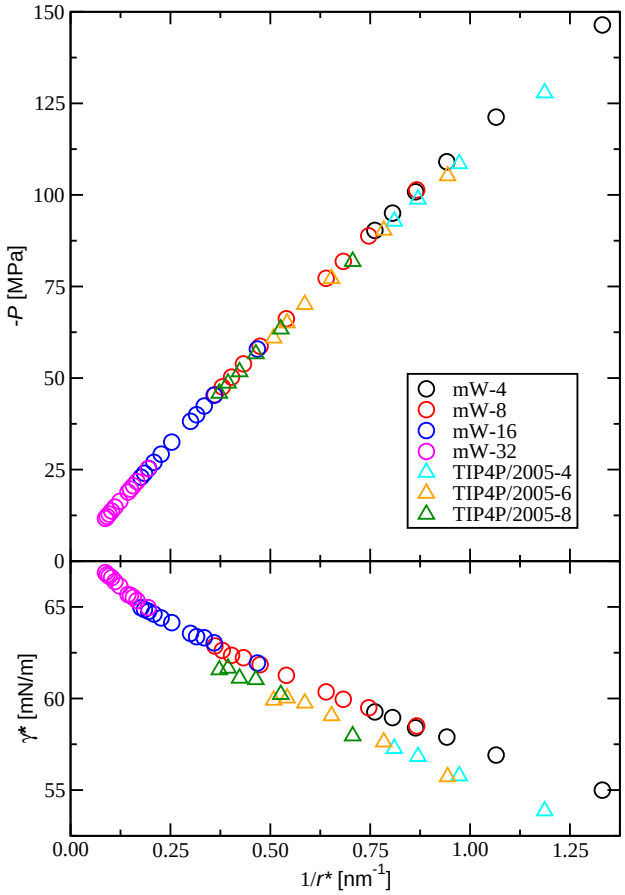


Figure 3.7: Negative system pressure (top) and bubble surface tension (bottom) versus inverse bubble radius. Data points are collected from all bubbly water phases for all combinations of water models and system sizes. Statistical uncertainties are smaller than the symbol size.

the bubbly water region where larger systems yield less negative values of P at the same ρ (see Figure 3.6). However, the density of the liquid region of the bubbly phase itself depends on system size (see Figure 3.4). When P is plotted as function of $\rho_{\mathrm{liq,fit}}$ for the bubbly systems instead of ρ, then the data for the homogeneously stretched phases and the bubbly systems fall onto the same curve (see Figure SB8). Based on the current molecular simulation data (see Figure 3.6), we surmise that use of two separate equations of state covering homogeneously stretched and bubbly phases should be explored for homogeneous mixture models.

It should be noted that, at the same ρ, a larger system contains a larger bubble than a smaller system when both are in the bubbly region. To further investigate the system size dependence, the simulation data were analyzed based on the Young–Laplace equation, which describes the relationship between pressure inside the bubble (P_{in}) and

pressure outside the bubble (P_{out}) for two-phase fluids based on mechanical equilibrium:

$$P_{\text{in}} - P_{\text{out}} = \frac{2\gamma^*}{R} \tag{3.14}$$

where γ^* and R are the surface tension and the bubble radius, respectively, that is taken here as r^*. Since the saturated vapor pressure predicted by both water models is a very small positive number at $T = 298$ K, we assume that P_{in} is negligible and that P_{out} can be set equal to the system pressure ($-P \approx 2\gamma^*/r^*$). For the bubbly region, the negative of the system pressure P is plotted against the reciprocal of the bubble radius r^* in Figure 3.7. As an aside, we again note that the radius of the smallest bubble oberved for the system density just below the BC point is found to increase with increasing system size. Regardless of system size, the correlation between $-P$ and $1/r^*$ is close to linear for both water models with a correlation coefficient of R^2 larger than 0.995. Therefore, we can conclude that even at the sub-10-nm scale, the Young–Laplace equation is still applicable within some small tolerance, which is consistent with the findings of some previous studies. [110, 111] However, when the surface tension value γ^* is estimated by applying the Young-Laplace equation to each individual data point, then a trend emerges where γ^* increases nearly linearly with decreasing inverse bubble radius for both the mW and TIP4P/2005 models (see Figure 3.7). Thus, formation of the convex liquid-vapor surface results in a smaller free energy penalty than for the planar surface. The slope of γ^* versus $1/r^*$, a metric for the curvature effect, is more negative for the TIP4P/2005 model than for the mW model. The smaller curvature effect for the mW model may explain why this model can sustain a larger maximum tension before cavitation becomes spontaneous.

Using an unweighted linear fit to estimate the value of γ^* at zero curvature ($1/r^* = 0$) yields 65.3 ± 0.2 and 66.9 ± 0.2 mN/m for the TIP4P/2005 and mW models, respectively. For the TIP4P/2005 model, this value agrees within uncertainties with the separate calculation of γ for the planar interface (see Table 3.1), whereas the extrapolation for the mW model yields a value that is about 2% larger than γ.

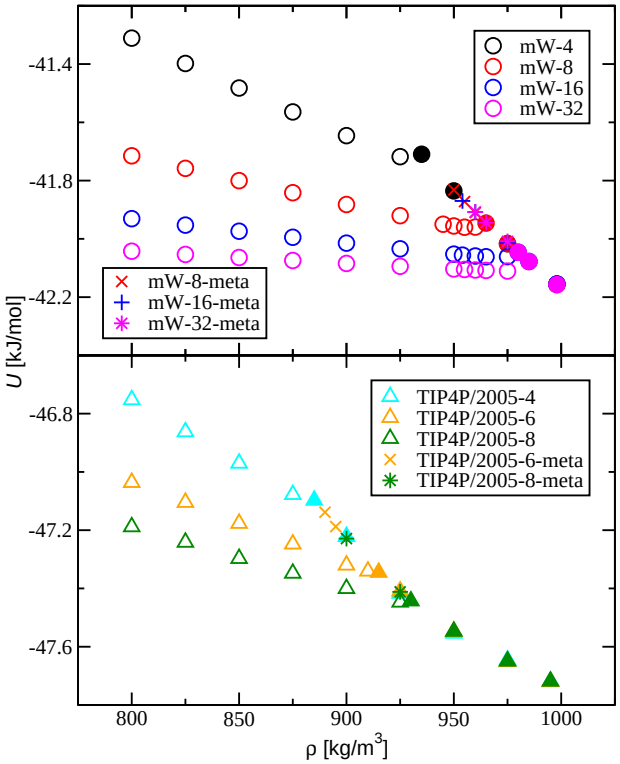


Figure 3.8: System potential energy versus density. Open and filled symbols represent bubbly and homogeneously stretched water systems, respectively. The plusses, crosses, and stars show data for metastable homogeneous systems. Uncertainties are all smaller than the symbol size.

3.3.5 Potential energy

The system potential energy U is plotted as function of density in Figure 3.8 (numerical data are provided in Tables SB2–SB15). The TIP4P/2005 model yields an overestimation of ΔH_{vap} and, hence, U is shifted downward compared to the mW model. In the homogeneously stretched region, the potential energy increases with decreasing density (consistent with the negative pressure, see Figure 3.6) and the magnitude of the slope is greater for the mW model (consistent with its more negative value of P compared to the TIP4P/2005 model at the same ρ and system size). For both water models, U exhibits a downward jump when cavitation occurs from the metastable phase because the favorable decrease of the potential energy of the less-stretched liquid region outweighs the energetic cost of forming a highly curved interface. The slower increase in U as density decreases results from a competition due to the favorable increase of the density of the liquid region and the unfavorable increase in the fraction of interfacial molecules (see Figure 3.4), and it is consistent with the smaller magnitude of the negative pressure.

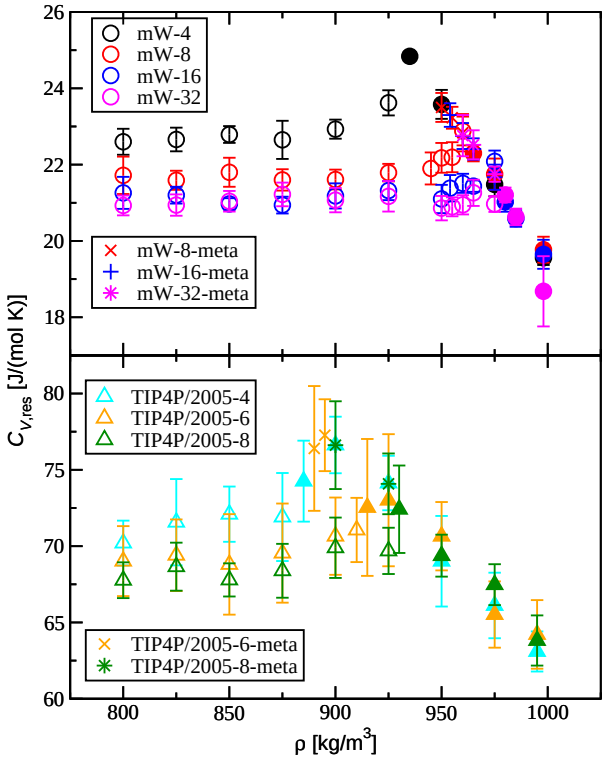


Figure 3.9: System residual isochoric heat capacity versus density. Open and filled symbols represent bubbly and homogeneously stretched water systems, respectively. The plusses, crosses, and stars show data for metastable homogeneous systems.

3.3.6 Residual isochoric heat capacity

The density dependence of the residual isochoric heat capacity $C_{V,\mathrm{res}}$ is shown in Figure 3.9 (numerical data are provided in Tables SB2–SB15). $C_{V,\mathrm{ideal}}$ for water at 300 K is around 27.3 J/(K·mol) as estimated from the molecular partition function in previous work [112]. The experimental C_V for bulk liquid water at ambient conditions is 74.5 J/(mol·K) [113]. After incorporating the contribution from $C_{V,\mathrm{ideal}}$, the TIP4P/2005 water model overestimates C_V by around 20%, while the mW water model underestimates C_V by around 35%.

Among all combinations of model and system size, the qualitative changes in $C_{V,\mathrm{res}}$ with ρ are consistent. $C_{V,\mathrm{res}}$ initially increases when approaching the SC point from the saturated liquid phase. When cavitation occurs, $C_{V,\mathrm{res}}$ drops and then enters a flat region where the change does not exceed the statistical uncertainties. However, in all cases, $C_{V,\mathrm{res}}$ in the two-phase region remains larger than that of the saturated liquid phase. The observation that $C_{V,\mathrm{res}}$ initially increases when approaching the transition point is consistent with previous findings [114, 115], indicating growing energy fluctua-

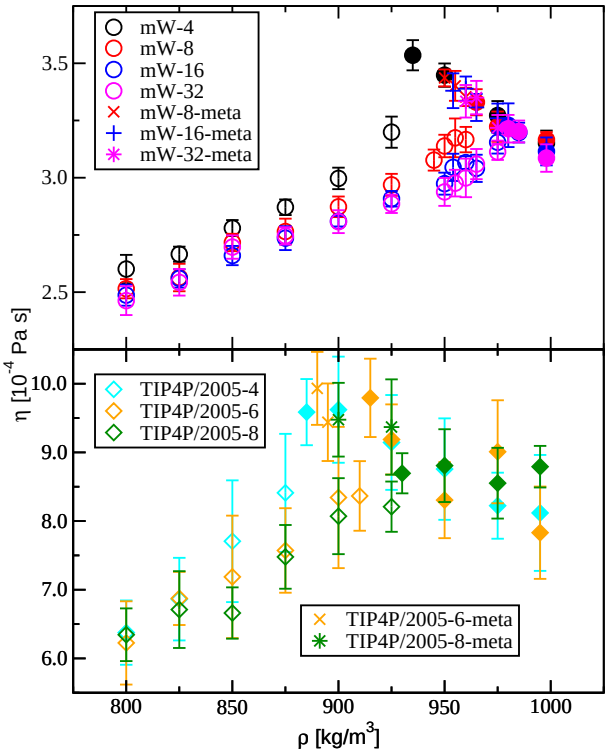


Figure 3.10: Shear viscosity versus density. Open and filled symbols represent bubbly and homogeneously stretched water systems, respectively. The plusses, crosses, and stars show data for metastable homogeneous systems.

tions as the system is being stretched. The decrease of $C_{V,\mathrm{res}}$ when cavitation occurs indicates that the system undergoes a relaxation. The relatively larger value of $C_{V,\mathrm{res}}$ for the bubbly region compared with the saturated liquid phase indicates that the presence of an interface introduces additional energy fluctuations into the system. However, the density of the liquid region $\rho_{\mathrm{liq,fit}}$ in the bubbly systems is smaller than ρ_{liq} at saturation (i.e., the liquid region is stretched), and this certainly also contributes to the larger value of $C_{V,\mathrm{res}}$.

$C_{V,\mathrm{res}}$ values for the bubbly water systems exhibit a system size dependence as well. Although going from mW-32 to mW-16 and TIP4P/2005-8 to TIP4P/2005-6 at a given ρ, the $C_{V,\mathrm{res}}$ values overlap within uncertainties but, with one exception, the values for the smaller system are always larger. A further decrease of box length yields a larger increase in $C_{V,\mathrm{res}}$. These findings indicate that fluctuations are more pronounced for smaller system sizes (larger values of α_{int}).

3.3.7 Shear viscosity

The shear viscosity η as a function of system density is shown in Figure 3.10 (numerical data are provided in Tables SB2–SB15). The shear viscosities predicted by the mW and TIP4P/2005 water models for the saturated liquid phase agree well with previous simulations [105,116], and the corresponding experimental value is 8.90×10^{-4} Pa·s. [113] In terms of accuracy, the TIP4P/2005 model underestimates η by less than 10%, whereas the mW model underestimates η by a factor of 2.9. This significant underestimation can likely be attributed to a faster momentum transfer through long-range collisions without the need to reorient dipole–dipole vectors and break hydrogen bonds for the case of the mW model.

For the mW model, the shear viscosity is found to initially increase when approaching the SC point. The value at the SC point is approximately 12% larger than that of the saturated liquid phase for the mW-4 system. When cavitation occurs, η jumps down and then enters a region where it decreases approximately linearly with decreasing system density. Dhabal *et al.* [116] have previously studied liquid-state anomalies of water using the mW model, and they also observed an increase of η when ρ decreases from bulk liquid density to 960 kg/m^3 at $T = 300$ K. Due to the larger relative uncertainites for the TIP4P/2005 systems, the initial increase in η as the density decreases in the homogeneously stretched region is less evident. For the TIP4P/2005-4 system, η at $\rho = 885$ kg/m^3 is about 20% larger than for the saturated liquid. De Hijes *et al.* [117] have also calculated the shear viscosity of liquid water for the TIP4P/2005 model at $T = 300$ K and $N = 216$, and they reported an initial increase and then decrease when ρ decreases from bulk liquid density to 800 kg/m^3, with a local maximum near 880 kg/m^3. The initial increase is consistent with our data, but since our smallest system size ($N \approx 2000$) only allows for water to stay homogeneous with density values larger than 875 kg/m^3, the local maximum was not detected clearly in this work.

For the two-phase region, the slope in η versus ρ is larger for the TIP4P/2005 model than for the mW model, and there appears to be a system size dependence for η as well. However, predicted values for the two largest systems for both water models agree with each other within uncertainties. Since an analytic equation that correlates η and ρ could potentially be useful for predicting multi-phase flow in CFD, and the TIP4P/2005 model is much more accurate than the mW model for predicting

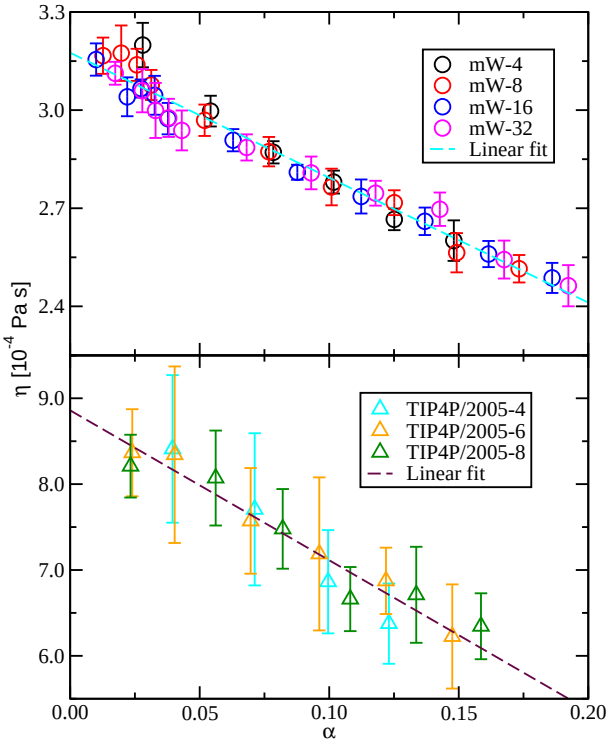


Figure 3.11: Shear viscosity versus bubble volume fraction.

shear viscosity, the data for the TIP4P/2005-8 system were fitted to a linear equation: $\eta\,/\,(\mathrm{Pa}\cdot\mathrm{s}) = (0.0163 \times 10^{-4}\,\mathrm{m}^3/\mathrm{kg})\,\rho - 6.7732 \times 10^{-4}$, with a correlation coefficient of $R^2 = 0.93$. Some of the analytic equations used in current multi-flow CFD simulations [70] correlate the viscosity of the bubbly water phase as function of the bubble volume fraction (which itself is a linear function of ρ, see Figure 3.3) and, hence, the shear viscosity is plotted versus bubble volume fraction in Figure 3.11. In this case, all data points fall onto a straight line [118] for a given water model. Thus, the dependence of η on system size for a given ρ is caused by the dependence of the bubble volume fraction on system size. Using the data from the TIP4P/2005 systems, the linear equation $\eta\,/\,(\mathrm{Pa}\cdot\mathrm{s}) = -17.463 \times 10^{-4}\,\alpha + 8.8596 \times 10^{-4}$ (with $R^2 = 0.93$) is obtained, where the intercept agrees well with the value of $(8.82 \pm 0.38) \times 10^{-4}$ Pa·s found for the TIP4P/2005-8 system at saturation.

Beyond the simple linear relation, a more complex relation between two-phase viscosity and bubble volume fraction was proposed by Beattie and Whalley: [119] $\eta = \alpha\eta_{\mathrm{vapor}} + (1-\alpha)(1+2.5\alpha)\eta_{\mathrm{liq}}$, where the term $(1+2.5\alpha)$ follows theoretical arguments. [120] The quadratic dependence yields an initial increase in η with increasing α for small α values. The data for bubbly phases presented in Figure 3.11 do not show an

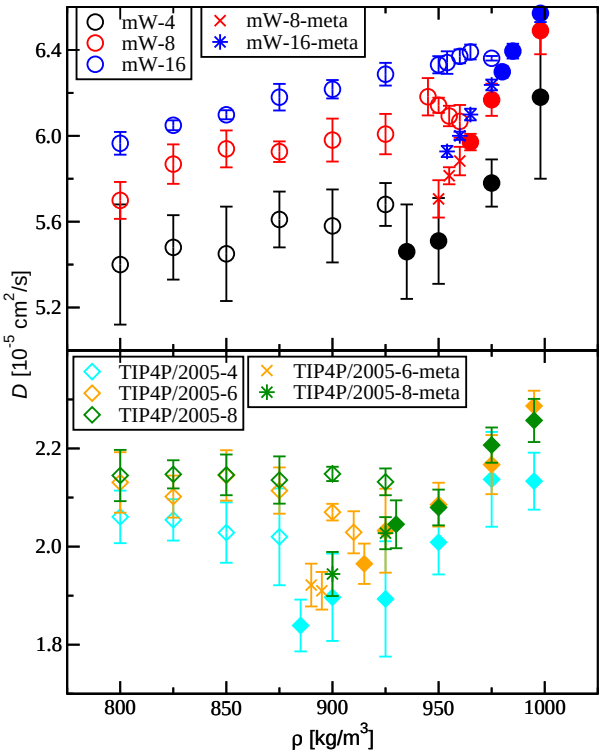


Figure 3.12: Self-diffusion coefficient versus density. Open and filled symbols represent bubbly and homogeneously stretched water systems, respectively. The plusses, crosses, and stars show data for metastable homogeneous systems.

initial increase, but we also do not observe infinitely small bubbles due to the transition to the homogenously stretched phase at the BC density. On the other hand, the data including the homogeneously stretched phase show an initial increase with decreasing density (see Figure 3.10). Indeed the size of the transient voids in the homogeneously stretched regime is of molecular dimensions (< 0.1 nm^3). Based on the current molecular simulation data (see Figure 3.10), we surmise that use of two separate linear regimes covering homogeneously stretched and bubbly phases should be explored to describe the dependence of the viscosity on density and direction of the density change.

3.3.8 Self-Diffusion

The self-diffusion coefficient D_{self} as a function of density is shown in Figure 3.12 (numerical data are provided in Tables SB2–SB15). Previous studies have shown that diffusion coefficients calculated from molecular dynamics simulations depend on system size, and corrections have been proposed to estimate the thermodynamic limit of D_{self}. [121, 122] However, in practice, many studies prefer to use a sufficiently large system instead of a

correction, and it has been shown to mitigate the system size effect of water at ambient conditions. [123, 124] Since system size effects are of interest for the current paper, a correction for the diffusion coefficient was not used here. The diffusion coefficient for the saturated liquid phase obtained for the mW-16 system is $(6.57 \pm 0.04) \times 10^{-5}$ cm^2/s, which is in excellent agreement with 6.5×10^{-5} cm^2/s from a previous study. [81] The calculated value for the saturated liquid phase of the TIP4P/2005-8 system ($N = 17023$) is $(2.26 \pm 0.04) \times 10^{-5}$ cm^2/s, which is significantly larger than the value of 2.06×10^{-5} cm^2/s reported previously for $N = 360$. [105,125] The experimental diffusion coefficient for liquid water at 298 K and saturation is 2.30×10^{-5} cm^2/s [126]. Our calculated value from TIP4P/2005-8 simulation underestimates the experimental value by less than 2% and finite-size corrections or simulations for larger systems may further reduce this deviation. In contrast, the mW-16 simulations yield D_{self} that is too large by a factor of 2.9, which is consistent with the predicted shear viscosity of saturated liquid water assuming the Stokes-Einstein relation.

Regardless of the system size and water model, the diffusion coefficient initially decreases when approaching the spinodal point from the saturated liquid phase. The reason for this anomalous behavior of water is that, as the density decreases, a larger fraction of the water molecules populates low-density states that are dynamically less mobile because a less-strained tetrahedral network leads to a higher activation barrier for transport. [117] When cavitation occurs, D_{self} jumps to a larger value. A further decrease in the density leads to a statistically significant decrease in D_{self} for the mW-8 and mW-16 systems, but no significant change was observed for the mW-4 and all TIP4P/2005 systems. Although the relative uncertainties are larger for the TIP4P/2005 systems (likely because of its slower, but more accurate diffusion), there is no indication that the data supports a decrease in D_{self} with decreasing ρ in the bubbly region. As indicated in Figure 3.12, D_{self} increases with system size irrespective of the system being homogeneously stretched or in the two-phase region, but the relative increase appears to be larger for the two-phase region. For the homogeneously stretched systems, the diffusion coefficients seem to reach a satisfactory level of convergence in terms of system size going from mW-8 to mW-16 and TIP4P/2005-6 and TIP4P/2005-8, while convergence in terms of system size is not achieved for the two-phase systems.

For both water models, D_{self} for the bubbly region is smaller than for the saturated

liquid phase. A likely explanation for this behavior is that the activation barrier for evaporation is much larger than the activation barrier for diffusion and, hence, the molecules in the liquid region must follow a more tortuous diffusion pathway around the bubble. [127] Furthermore, the fact that η is also found to be smaller for the bubbly region indicates that the Stokes-Einstein relation does not hold for the bubbly region. Figure SB9 shows D_{self} as function of η^{-1}; within statistical uncertainties, the behavior of the homogeneously stretched systems (including the metastable states) follows the Stokes-Einstein relation with D_{self} increasing linearly with increasing η^{-1}. However, the Stokes-Einstein relation breaks down for the bubbly region. For the mW model, D_{self} is found to decrease with increasing η^{-1}, and D_{self} appears to be independent of η^{-1} for the TIP4P/2005 model. The deviations from the Stokes-Einstein relation in the bubbly region are much more pronounced than the previously observed deviations for the homogeneously stretched region. [117]

3.4 Conclusion

For all combinations of water models and system sizes studied in this work, the canonical ensemble simulations at $T = 298$ K yield a homogeneously stretched phase for $\rho \geq 980$ kg/m^3 and a bubbly state (with a spherical vapor bubble surrounded by a liquid region) for $\rho \leq 875$ kg/m^3. For the intermediate density region, the outcome of the simulations depends on the simulation protocol and system size. For simulations started from a homogeneous configuration, the homogeneously stretched state persists to lower densities until reaching a spinodal cavitation (SC) point where bubble formation is spontaneous (or requires passing a readily surmounted free energy barrier). In contrast, for simulations started from a state with a pre-formed spherical cavity, the bubbly state persists to densities higher than the SC point until a bubble collapse (BC) point is reached when the system spontaneously converts to a homogeneously stretched state. This behavior is akin to the hysteresis loop well-known for adsorption-desorption isotherms in mesoporous materials. As the system size is increased, both SC and BC points occur at larger densities. Extrapolation versus inverse simulation box length yields infinite-size estimates of 912 ± 14 and 961 ± 3 kg/m^3 for the spinodal densities of the TIP4P/2005 and mW models, respectively.

When the water system is still in the homogeneously stretched state (including the metastable part of the hysteresis loop), then system size effects are not significant with the exception of the diffusion coefficient. For the inhomogeneous bubbly water region, system size effects (comparing properties for the same ρ and T, but different N and V) are observed for all properties. The main reason for these pronounced system size effects is that, at the same ρ and T, a larger system contains a larger bubble than the smaller system, and some properties (e.g., pressure, density of the liquid region, and shape anisotropy of the bubble) are described better as function of bubble radius instead of bubble volume fraction. On the other hand, the viscosity of the two-phase system is well described as function of the bubble volume fraction, the isochoric heat capacity for bubbly systems does not show a strong dependence on bubble radius nor volume fraction.

Qualitatively similar trends are observed for the thermophysical properties obtained for the TIP4P/2005 and mW water models with the exception of D_{self} in the bubbly region. As the density is decreased from the saturated liquid density, the pressure decreases to a large negative value ($P < -100$ MPa near the SC point for both models), the magnitude of the potential energy decreases by only about 1%, whereas the residual isochoric heat capacity and viscosity increase by more than 10%, and the self-diffusion coefficient decreases by more than 10%. Upon cavitation, the system undergoes a relaxation as signaled by steep and sudden changes in these properties. A further decrease in the density (increase in bubble volume fraction and radius) leads to a relatively slower increase in P and U, a decrease in η, and a decrease in D_{self} for the mW model, but not the TIP4P/2005 model, whereas $C_{V,\text{res}}$ exhibit no significant changes. Upon bubble collapse, the discontinuous changes in properties follow the reverse trend (i.e., opposite sign) as those encountered at the SC point, but are much smaller in magnitude.

Given that the mW and TIP4P/2005 water models yield the same qualitative trends for the various thermophysical properties (except for the diffusion coefficient in the bubbly region) studied in this work, we surmise that the mW water model, albeit coarse-grained and much less accurate than the TIP4P/2005 model, remains a good choice for the qualitative study of bubbly water systems at ambient conditions and allows to access much larger system sizes and longer time scales. However, when quantitative predictions are of concern, then the TIP4P/2005 water model is much preferable but,

without polarization correction, there are significant deviations for the saturated vapor pressure and heat of vaporization.

The pressure of bubbly water systems is well described by the Young–Laplace equation, but both water models exhibit a decrease in the effective surface tension with increasing curvature. Interestingly, the simulations show that many properties in the two-phase region cannot be obtained from an interpolation of the properties of the saturated liquid and vapor phases. The value of the residual isochoric heat capacity, which depends on fluctuations of the system potential energy, is larger for the bubbly water phase than for the saturated liquid phase (and of course also the saturated vapor phase), whereas the self-diffusion coefficient is smaller for the bubbly region than for the saturated liquid phase (and of course also the saturated vapor phase) due to molecules having to diffuse around the bubble. The viscosity is largest for the homogeneously stretched phase near the SC point, and the Stokes-Einstein relation is not obeyed in the bubbly region.

The results of this work can inform multi-phase flow CFD studies. Most importantly, although a negative pressure is not considered possible in most present CFD simulations because of the use of an EOS valid only in the thermodynamic limit, a negative pressure is observed here for both homogeneously stretched liquid water and bubbly water. The Young–Laplace equation can be utilized to estimate the negative pressure of bubbly water systems from the bubble radius that, in turn, can be deduced from the CFD cell density and length. Furthermore, some recent experimental studies have also investigated the EOS for water under negative pressure [128, 129]. Using extensive simulation data for the TIP4P/2005 water model, Biddle *et al.* [130] have proposed an extension of the two-structure EOS for water that covers a pressure range from large negative to large positive values and incorporates the spinodal. In general, we have to keep in mind that under a turbulent cavitating flow condition, both the pressure range and temperature range might be large inside the flow domain. Thus, a proper choice of EOS to describe liquid water is already posing challenge, not to mention the challenge of describing the two-phase region. In the homogeneously stretched and bubbly regions, as discussed above, values of some of the thermophysical properties fall outside the range spanned by the corresponding values for the saturated vapor and liquid properties; an observation that precludes estimating these properties based on interpolation between

the thermodynamically stable phases. In addition, due to the pronounced hysteresis loop, some properties are path dependent. Thus, changes in the analytic equations underlying CFD simulations of multi-phase flow may need to be considered.

Chapter 4

Large-scale molecular dynamics simulations of bubble collapse in water[a]

4.1 Introduction

The prediction of multi-phase, or cavitating flows is inherently complex due to the large span in spatial and temporal scales of the distinct physical phenomena involved, ranging from nucleation at molecular scale to sheet or cloud cavitation at macroscopic scale [131, 132]. Among them the collapse of bubble(s) is of particular interest as it is known to cause erosion in hydro turbines through high-pressure pulses, liquid jet, or local high temperatures generated during collapse [133, 134]. A lot of experiments have been performed to explore bubble collapse. However, the time and spatial resolution is restricted by the imaging instrument [135], and the measurements of local extreme conditions (e.g. temperature) within bubble are often indirect [136]. In addition to experiments, a lot of theoretical and computational studies have been done on bubble collapse as well. The Rayleigh-Plesset (RP) equation [137,138] is an ordinary differential

[a]This chapter is adapted from my publication in preparation: Chen, J. L.; Sun, Y.; Bhatt, M.; Knight, C.; Mahesh, K.; Siepmann, J. I. Large-scale molecular dynamics simulations of bubble collapse in water: effects of system size, water model, and nitrogen. My contribution includes running all the simulations, analyzing all the data, and writing up the manuscript. Yangzesheng Sun was responsible for compiling LAMMPS and submitting jobs at Argonne Leadership Computing Facility. Dr. Mrugank Bhatt provided an initial version of the MATLAB code for solving the Rayleigh-Plesset equation. Dr. Christopher Knight helped benchmark the performance of LAMMPS at Argonne Leadership Computing Facility and suggested an efficient way to initialize the system.

equation describing spherical cavity dynamics in an incompressible infinite medium. It is a simple yet powerful tool for deriving many basic results regarding bubble dynamics [139]. Meanwhile, computational fluid dynamics (CFD) simulations can provide more details and are capable of probing more complex systems. However, various assumptions are introduced to account for the multi-phase nature of the system, such as the source term describing the vapor production rate in a local volume, leading to different number and/or forms of the transport equations being solved [140–142].

Molecular dynamics (MD) simulations directly apply statistical mechanics or classical mechanics to generate trajectories of molecular (particle) systems where the particle interactions are computed either from quantum mechanics or are represented through force fields. The time and length scale approachable by MD simulations are typically orders of magnitude smaller than continuum simulations, making it only suitable for probing molecular-scale phenomena such as nucleation [143]. However, with increasing computing power, recent MD simulations are capable of probing much larger system sizes [144, 145]. In this way system size effects can be carefully estimated and MD simulation results can potentially serve as a validation for the underlying physical models utilized in continuum simulations. Previous MD simulations have looked into bubble collapse in both Lennard-Jones (LJ) fluids and molecular fluids without or with a nearby surface or external field [145, 145–150]. However, to our knowledge, no previous MD simulations on bubble collapse have done a comprehensive study on system size effects, or have discussed about the different collapse behavior predicted by coarse-grained and atomistic water models.

In this work, we report MD simulation results for the collapse of a bubble in water using both coarse-grained single-site mW [151] and the TIP4P/2005 water model [152]. System size effects were studied for both water models. Results obtained from MD simulations were compared with numerical solutions from the RP equation, which further explains the different collapse dynamics predicted by the two types of water models. To study nitrogen effects on bubble collapse, a single-site nitrogen model was developed and used together with the mW water model, while the TraPPE nitrogen model [153] was used with the TIP4P/2005 water model. Local conditions such as density and temperature were analyzed and visualized with heat maps.

4.2 Simulation methods

4.2.1 Force fields

The mW water model [151] includes both two-body and three-body interaction terms

$$U = \sum_i \sum_{j>i} \phi_2(r_{ij}) + \sum_i \sum_{j\neq i} \sum_{k>j} \phi_3(r_{ij}, r_{ik}, \theta_{ijk}) \tag{4.1}$$

$$\phi_2(r_{ij}) = A\varepsilon \left[B \left(\frac{\sigma}{r_{ij}} \right)^m - \left(\frac{\sigma}{r_{ij}} \right)^n \right] \exp\left(\frac{\sigma}{r_{ij} - a\sigma} \right) \tag{4.2}$$

$$\phi_3(r_{ij}, r_{ik}, \theta) = \lambda\varepsilon(\cos\theta - \cos\theta_0)^2 \exp\left(\frac{\gamma\sigma}{r_{ij} - a\sigma} \right) \exp\left(\frac{\gamma\sigma}{r_{ik} - a\sigma} \right) \tag{4.3}$$

where $A = 7.049556277$, $B = 0.6022245584$, $m = 4$, $n = 0$, $\gamma = 1.2$, $a = 1.8$, $\theta_0 = 109.47°$, $\lambda = 23.15$, $\epsilon/k_B = 3114.5$ K, and $\sigma = 2.3925$ Å. Both two-body and three-body interactions are truncated at a distance R_c of $a\sigma$ (4.3065 Å). The nitrogen model used together with the mW water model is also a single-site coarse-grained model. N_2–N_2 and water–N_2 interactions utilizes pair-wise shifted-force LJ 12-6 potential

$$U_{sf}(r_{ij}) = U(r_{ij}) - U(R_c) - (r_{ij} - R_c) \frac{dU(r_{ij})}{dr_{ij}}|_{r_{ij}=R_c} \tag{4.4}$$

where

$$U(r_{ij}) = 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] \tag{4.5}$$

$\epsilon_{N_2-N_2}$ and $\sigma_{N_2-N_2}$ are fitted to the critical temperature and density of N_2. $\epsilon_{water-N_2}$ and $\sigma_{water-N_2}$ are fitted to the solubility of nitrogen in liquid water at two chosen conditions with experimental data. The resulting parameters are $\epsilon_{N_2-N_2/k_B} = 135$ K, $\sigma_{N_2-N_2} = 3.62$ Å, $\epsilon_{water-N_2/k_B} = 123$ K, and $\sigma_{water-N_2} = 3.22$ Å. For both N_2–N_2 and water–N_2 interactions, the truncation distance of potential R_c is 9 Å. Details regarding the fitting of N_2–N_2 and water–N_2 interaction parameters are provided in supporting information (SI).

Both the TIP4P/2005 water [152] and TraPPE nitrogen model [153] are rigid, non-polarizable, atomistic models. The pair-wise potential is described by a combination of LJ and electrostatic interactions

$$U(r_{ij}) = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \tag{4.6}$$

Electrostatic interaction is computed using the particle-particle particle-mesh Ewald summation technique [154] with a relative force error of 10^{-5}. Both LJ interaction and the real-space contribution from the Ewald summation are truncated at a distance of 14 Å. To test the performance of the combination of TIP4P/2005 and TraPPE nitrogen model, the solubility of N_2 in water at several conditions is computed and the agreement with experimental data is overall satisfactory (see Table SC1).

4.2.2 Molecular dynamics simulation details

MD simulations were performed using the LAMMPS software package [88]. All simulations were carried out with periodic boundary condition. The procedure for simulating bubble collapse can be divided into three steps. First, water is equilibrated in the NVT ensemble at $T = 298$ K with a spherical LJ wall of radius R_w centered at the center of a cubic simulation box. Due to the low vapor pressure and density predicted by both water models at 298 K [155], the estimated number of vapor molecules within the largest bubble is less than 5. Therefore all water molecules are placed outside the LJ wall and the bubble is essentially a void. Next, the LJ wall is removed and system undergoes a short relaxation for 0.1 ps. Finally, simulation is switched to the NVE ensemble and the timer for bubble collapse starts. For simulations with coarse-grained models, the time step size for equilibration and relaxation is 5 fs, and is reduced by half near the collapse point due to presence of large atomic velocities. For simulations with atomistic models, the time step size is 1 fs. Simulations for water/nitrogen system adopt a similar three-step procedure, except that the initial configuration is obtained by combining previously equilibrated pure water system with N_2 molecules placed randomly inside the LJ wall.

For all systems studied in this work, the overall liquid water density was fixed at a value close to the predicted saturated liquid density at $T = 298$ K (mW: 998 kg/m^3; TIP4P/2005: 995 kg/m^3) [155]. For the mW water model, effect of outside liquid pressure was explored by performing two series of simulations with $R_w/L = 0.15$ and 0.075, with initial bubble volume fractions around 0.014 and 0.0018, respectively. The

corresponding initial liquid pressures are around 795 and 110 atm, respectively. For the TIP4P/2005 water model, simulations were performed with $R_w/L = 0.15$, and the initial liquid pressure is around 355 atm. For pure water systems, simulations were performed under various L to study system size effects. For convenience, pure water systems in this work are distinguished by a combination of water model, L in units of nm, and R_w/L. Simulations for water/nitrogen system were performed for mW-128-0.15 and TIP4P/2005-64-0.15 systems, and are denoted as mW-128-0.15-N_2 and TIP4P/2005-64- 0.15-N_2, respectively. The nitrogen mole fraction χ_{N_2} is around 1.2×10^{-5}, and the resulting initial nitrogen pressure inside bubble estimated using ideal gas equation of state is approximately 1 atm. The details of all studied systems are provided in Table SC2.

4.2.3 Data analysis

The calculation of bubble volume follows a similar protocol with our previous work [155], where the bubble is defined as the largest vapor cluster of meshes of size $0.2 \times 0.2 \times 0.2$ nm^3. For water/nitrogen systems, only the water molecules are considered in calculating bubble volume. To analyze bubble volume efficiently for systems with large number of molecules, a paralleled neighbor search algorithm with domain decomposition was applied [156]. For consistency, the bubble collapse point is defined as when the bubble volume decays below a volume cut-off of 0.08 nm^3 for the neat water systems, and as the first volume minimum for the water/nitrogen mixture systems. Bubble radius is calculated from bubble volume assuming a spherical shape, which holds reasonably well even for voids with 1 nm radius [155]. Wall velocity $\frac{dR}{dt}(t)$ is obtained numerically from $R(t)$, with a center difference scheme for middle points and one-sided difference scheme for boundary points. The time resolutions near the collapse point for the mW and TIP4P/2005 systems are 0.125 ps and 0.5 ps, respectively. The values of initial bubble radius R_0, initial wall velocity $\frac{dR}{dt}_0$, bubble collapse time t_c, and most negative wall velocity captured preceding initial collapse point $\frac{dR}{dt}_m$ for all simulated systems are given in Table SC3. Note that there is a slight difference between R_w and R_0 since only R_w is the directly controllable parameter. The values of $\frac{dR}{dt}_0$ also deviate from 0, but the deviation generally decreases as bubble size increases.

4.2.4 The Rayleigh-Plesset equation

The RP equation takes the form

$$R\frac{d^2R}{dt^2} + \frac{3}{2}\left(\frac{dR}{dt}\right)^2 + \frac{4\nu_L}{R}\frac{dR}{dt} + \frac{2\gamma}{\rho_L R} + \frac{\Delta P(t)}{\rho_L} = 0 \tag{4.7}$$

where ν_L and ρ_L represent the kinematic viscosity and density of the surrounding liquid. γ represents the surface tension of the bubble interface. $\Delta P(t)$ is defined as $P_\infty - P_B$, which is the difference between external pressure infinitely far from bubble and the pressure within bubble. Without the viscosity and surface tension term, the RP equation is named the Rayleigh equation. To map MD simulation conditions into input parameters for solving the RP equation, P_∞ is taken as the outside liquid pressure, which is approximately the system pressure P_{sys}. P_B is assumed to be 0. Both ρ_L and P_{sys} are time-dependent, where ρ_L is estimated assuming an infinitely sharp interface. The trajectories of ρ_L and P_{sys} are elongated by repeating the last data point when solving the RP equation. Since the change in ν_L with respect to ρ_L at $T = 298$ K is minimal within the density range of interest (see Figure SC4), it is assumed to be constant and the values are chosen to be approximately the simulated bulk liquid values at $T = 298$ K ($\nu_{L,\ \text{mW}} = 3 \times 10^{-7}$ m^2/s, $\nu_{L,\ \text{TIP4P/2005}} = 8 \times 10^{-7}$ m^2/s) [155]. Based on previous work [155], γ only changes by around 15% when the interface changes from being flat to being roughly spherical with a radius of around 1 nm. Thus it is also assumed to be constant and takes the value predicted for the flat interface at $T = 298$ K ($\gamma_{\text{mW and TIP4P/2005}} = 0.065$ N/m) [155]. To test the sensitivity of the solutions to the viscosity term, the value of ν_L is varied. Initial conditions (R_0 and $\frac{dR}{dt}_0$) are taken from Table SC2. The numerical solutions are obtained using the MATLAB ode45 package [157, 158].

4.3 Results and discussion

4.3.1 Bubble dynamics predicted by the mW water model

Time evolution of bubble volume V, bubble radius R, and wall velocity $\frac{dR}{dt}$ for pure water systems predicted by the mW water model are presented in Figure 4.1. With the increase of system size and initial outside liquid pressure, the most negative wall

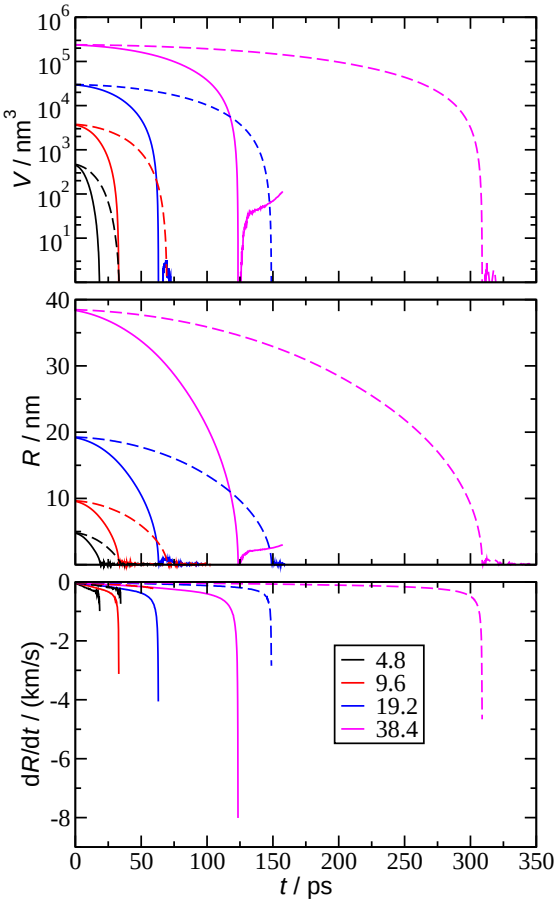


Figure 4.1: Time evolution of bubble volume (top), bubble radius (middle), and wall velocity (bottom) for all mW water systems. Different colors represent systems with different initial bubble radius. Solid lines represent data for mW-0.15 systems, and dashed lines represent data for mW-0.075 systems. For better visualization purpose, data for wall velocity is only shown up to the most negative value preceding the initial collapse point.

velocity captured preceding the collapse point $\frac{dR}{dt}_m$ becomes more negative (see Table SC2), indicating more violent collapse. The most violent collapse is observed for mW-256-0.15 system, and the corresponding $\frac{dR}{dt}_m$ is near -8 km/s, or Mach 3.5 (the speed of sound predicted by the mW water model for bulk liquid water at $T = 298$ K is around 2.3 km/s using the calculation method described in reference [159]). Concomitantly, a clear rebound after the initial collapse point is also observed.

To explore system size effects, time evolution of bubble volume and radius are plotted in reduced unit in Figure 4.2, where bubble volume and radius are normalized by their initial values and time is normalized by the corresponding bubble collapse time t_c. Note that this type of nondimensionalization process is a common practice in fluid mechanics

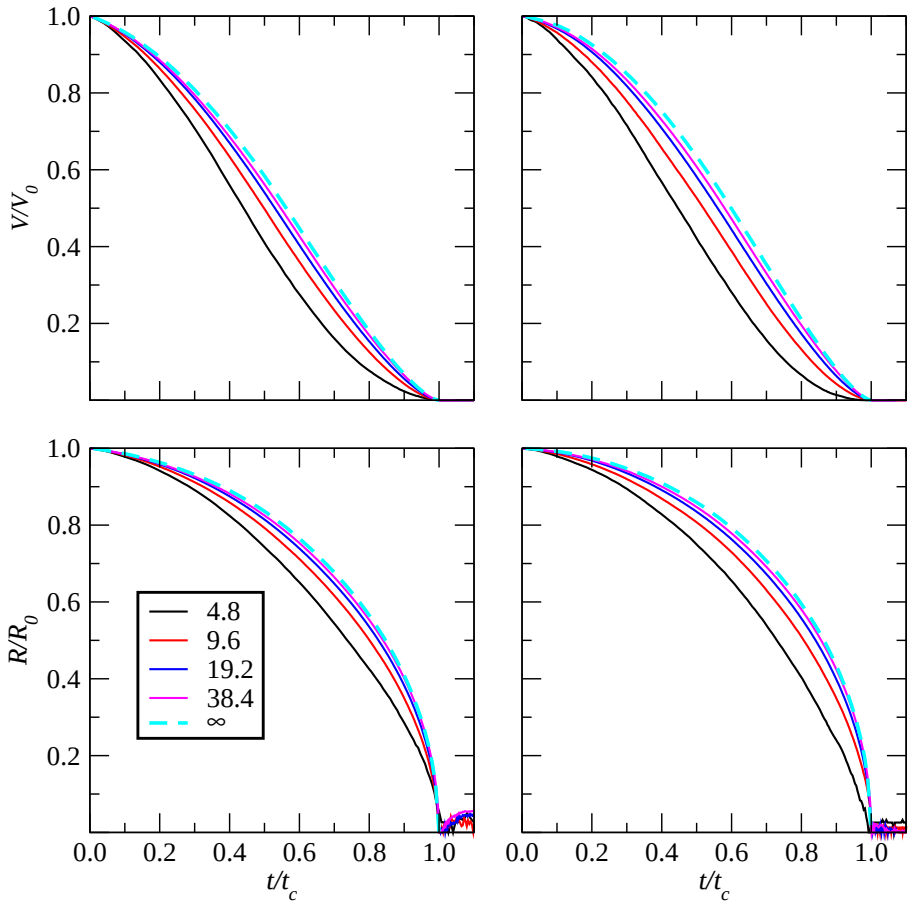


Figure 4.2: Reduced bubble volume (top) and bubble radius (bottom) as a function of reduced time for all mW water systems. Left column presents data for mW-0.15 systems, and right column presents data for mW-0.075 systems. Different colors represent systems with different initial bubble radius. Cyan curve was obtained by extrapolation.

[160], and bubble collapse results solved in larger scales are often presented in this way as well [161, 162]. For both mW-0.15 and mW-0.075 systems, a converging behavior is observed as system size increases for the initial collapse process. Note that in this case the RP equation is not capable of predicting a rebound after the initial collapse point. To estimate the limiting behavior ($L = \infty$), extrapolation based on system size was performed. First, reduced time range (0, 1) was divided into intervals of length 0.01. For each time point, a linear fit was performed between the corresponding V/V_0 and $1/L$, and the curve under $L = \infty$ is composed of all the y-intercepts of the linearly fitted lines, part of which are shown in Figure 4.3. The correlation coefficient squared R^2 for all the linearly fitted lines are also given in Figure 4.3. Overall, the fittings are satisfactory with more than 90% of them having a R^2 close to 1. A relatively larger deviation from linearity was observed at the beginning and near the end of collapse,

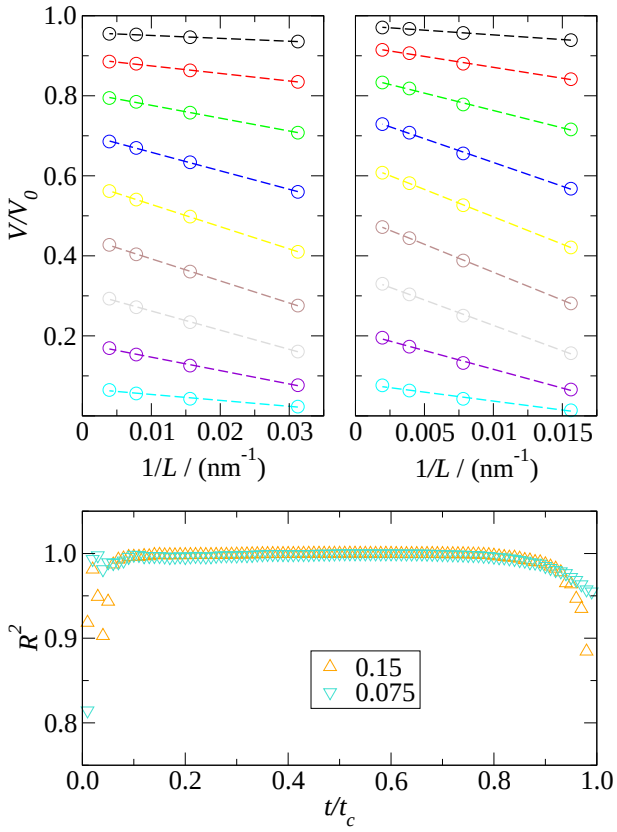


Figure 4.3: V/V_0 as a function of $1/L$ obtained for mW-0.15 (top left) and mW-0.075 (top right) systems at t/t_c = 0.1 (black), 0.2, 0.3...0.9 (cyan). Circular dots represent simulation data and dashed lines represent linearly fitted lines. Bottom graph shows R^2 as a function of t/t_c for mW-0.15 and mW-0.075 systems.

which could be explained by their slightly different initial conditions (see Table SC2) and less accurate determination of bubble size when its small near the collapse point. For both mW-0.15 and mW-0.075 systems, the behavior of the largest bubble is very close to the limiting behavior, indicating that the collapse behavior of a bubble of initial radius around 40 nm is already a good estimation for the collapse of macroscopic bubbles with radius at least on the order of microns.

4.3.2 Water model difference explained with the Rayleigh-Plesset equation

Time evolution of bubble volume V, bubble radius R, and wall velocity $\frac{dR}{dt}$ for pure water systems predicted by the TIP4P/2005 water model are presented in Figure 4.4. With an initial liquid pressure value in between those for mW-0.15 and mW-0.075 systems and a similar simulation set-up, one would expect a similar collapse behavior. However,

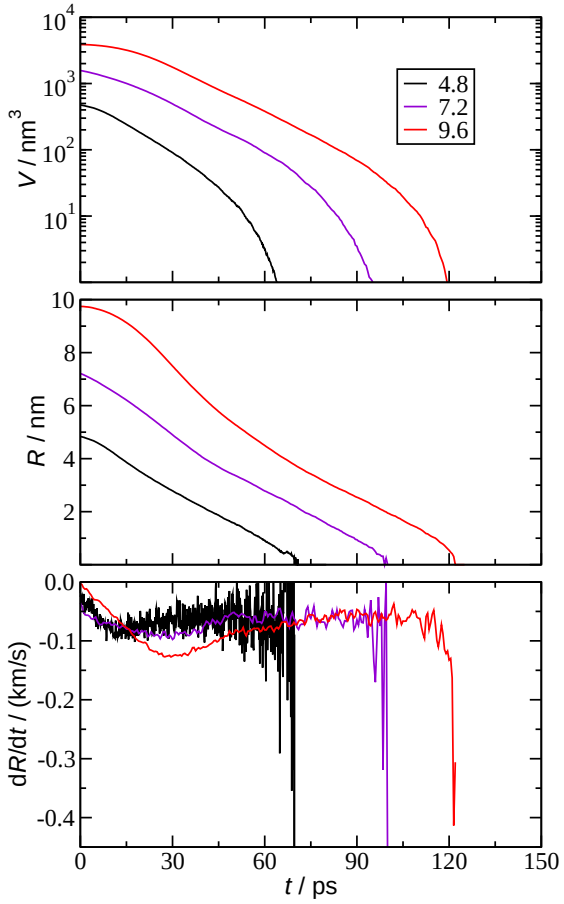


Figure 4.4: Time evolution of bubble volume (top), bubble radius (middle), and wall velocity (bottom) for all TIP4P/2005 water systems. Different colors represent systems with different initial bubble radius. For better visualization, data for wall velocity is only shown up to the most negative value preceding the collapse point.

comparison between Figure 4.4 and Figure 4.1 indicates a qualitatively different behavior in $R(t)$ curve. The TIP4P/2005 water model predicts a local minimum in $\frac{dR}{dt}(t)$ at the early stage of collapse. Near the end of collapse, a local maximum, though less obvious, is also observed in $\frac{dR}{dt}(t)$. These two local extrema corresponds to two inflection points in $R(t)$ where the sign of curvature changes, whereas the mW water model predicts a continuously decreasing $\frac{dR}{dt}$ before the initial collapse point. Because of this, bubble collapse predicted by the TIP4P/2005 water model is overall less violent.

To understand the water model difference, we turn to the RP equation. Simply by observing the signs of different terms on the left-hand side of the RP equation, it is possible to reach an inflection point in $R(t)$ where $\frac{d^2R}{dt^2}$ is 0, as long as the negative viscosity term balances out the other positive terms. Figure 4.5 presents MD simulation results and numerical solutions of the RP equation corresponding to mW-256-0.15,

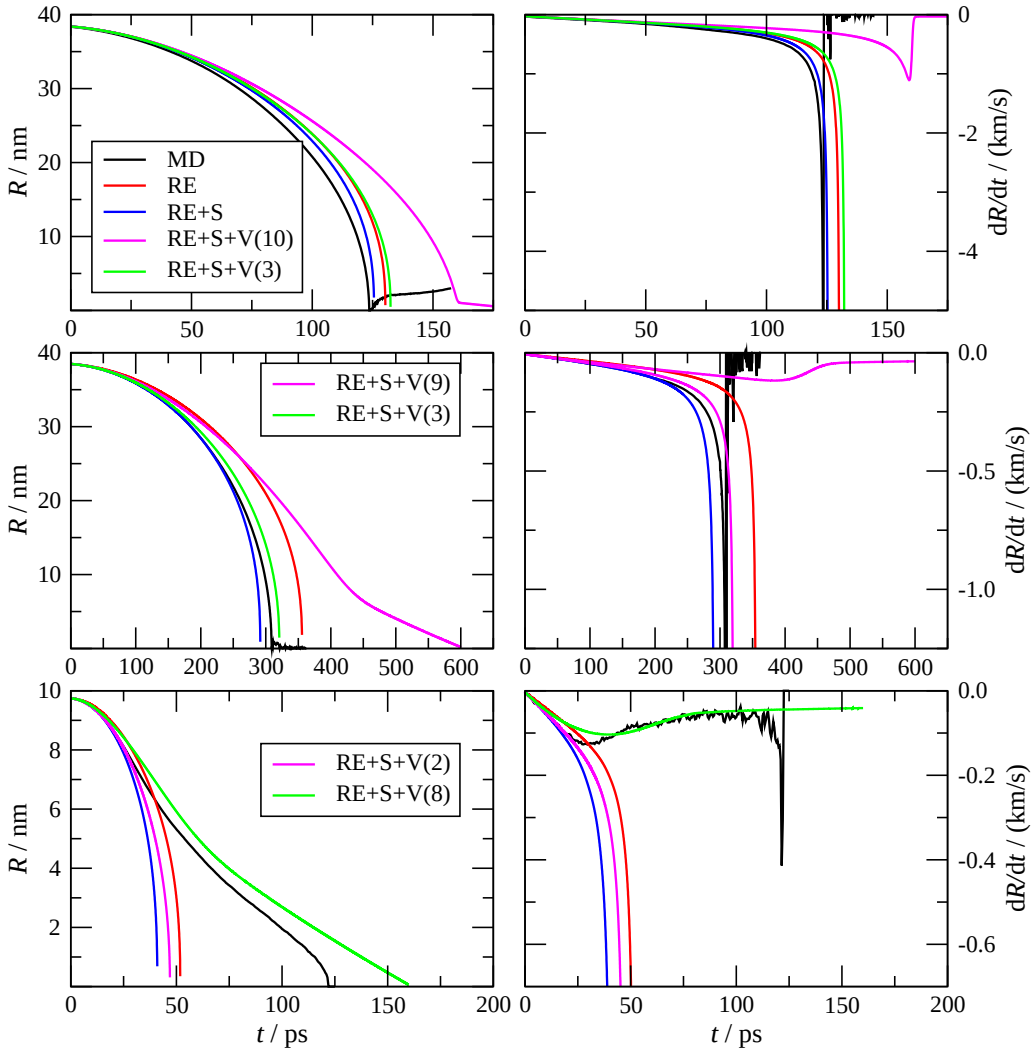


Figure 4.5: MD simulation results and numerical solutions of the RP equation for mW-256-0.15 (top), mW-512-0.075 (middle), and TIP4P/2005-64-0.15 (bottom) systems. Left column presents data for $R(t)$ and right column presents data for $\frac{dR}{dt}$. RE represents numerical solutions of the Rayleigh equation. RE+S represents numerical solutions of the Rayleigh equation plus the surface tension term. RE+S+V represents numerical solutions of the RP equation. The value in parenthesis denotes the value of ν_L in units of 10^{-7} m^2/s.

mW-512-0.075, and TIP4P/2005-64-0.15 systems. Overall, the agreement between MD simulation results (black) and the RP equation solutions (green) is satisfactory, which is also consistent with some previous studies [147, 150, 163]. The inclusion of the surface tension term leads to a faster collapse, whereas an opposite effect is observed for the viscosity term. This physically makes sense as the existence of surface increases system free energy and viscosity resists flow. For the mW-256-0.15 system where most violent collapse is observed, the solution is only slightly affected by the addition of surface tension and viscosity term due to a large contribution from the pressure term. However,

for the other two systems, the agreement is significantly better when all terms are taken into account, especially for the TIP4P/2005 system since the initial inflection point is only observed when the viscosity term is included. However, RP equation wasn't capable of predicting a clear second inflection point near the end of collapse, which could be due to a less accurate estimation of the surface tension, or the equation simply fails when bubble radius is below 1 nm. The sensitivity test shows that the solutions are indeed sensitive to the value of ν_L. Increasing ν_L not only leads to a longer collapse but also can change the shape of $R(t)$, as for the mW systems where an inflection point occurs near end of collapse when ν_L is artificially increased. This type of change is also observed in previous study [147]. Similarly, a decrease in ν_L leads to faster collapse and the inflection point in $R(t)$ can correspondingly disappear as well. To summarize for the water model difference, the fact that TIP4P/2005 water model predicts a larger viscosity value for water allows the viscosity to dominate over other contributions at the early stage of collapse, leading to the first inflection point in $R(t)$. Near the end of the collapse, the viscosity term loses its dominance and a second inflection point occurs. Since viscosity dominates a significant fraction of the collapse process, a less violent collapse is observed.

Bubble dynamics predicted by the TIP4P/2005 water model were also examined in reduced unit (see Figure SC5) and the same system size extrapolation was performed (see Figure SC6). Unlike the mW systems where an overall nice linear correlation was observed, the correlation coefficient squared R^2 is well below 1 when $t/t_c < 0.3$ and $t/t_c > 0.9$. One possible explanation is that the slight difference in $\frac{dR}{dt}_0$ among the TIP4P/2005-0.15 systems (see Table SC2) plays a more significant role compared with the mW-0.15 and mW-0.075 systems due to much less violent collapse. To test if the difference is due to randomness, another independent run was performed for the TIP4P/2005-48-0.15 system. However, the result is still similar (see Figure SC5). Other possible explanations are that the size threshold for observing a converging behavior is much larger for TIP4P/2005 systems, or inherently when $R(t)$ exhibits this type of shape, the linear correlation breaks down. However, to test this hypothesis, simulations needs to be performed with larger system sizes possibly up to $L = 128$ nm with simulation length up to 250 ps, which is not feasible with current available computing resources.

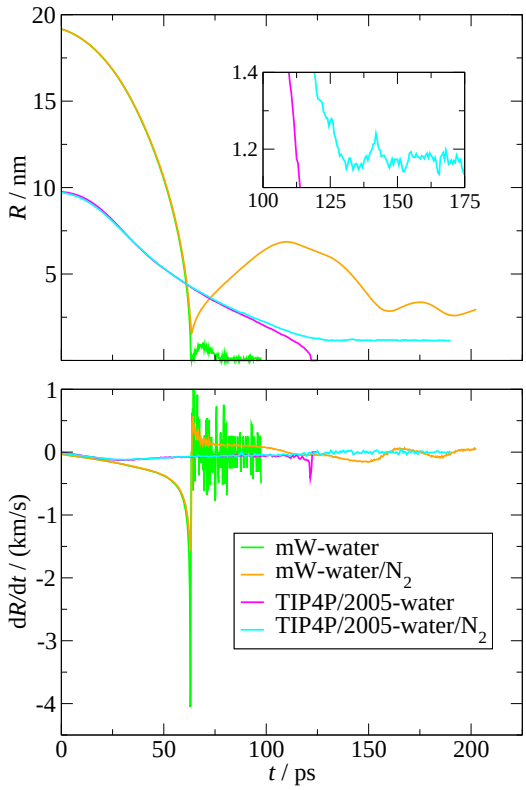


Figure 4.6: Time evolution of bubble radius (top) and wall velocity (bottom) for mW-128-0.15-N_2 and TIP4P/2005-64-0.15-N_2 systems and the corresponding pure water systems.

4.3.3 Bubble dynamics in water/N_2 systems

Figure 4.6 presents bubble dynamics results for the water/nitrogen mixture systems studied in this work. For comparison, results for the corresponding pure water systems are also presented. With the presence of nitrogen within bubble, bubble collapse is not complete for both mW/N_2 and TIP4P/2005/N_2 systems, with a minimum bubble radius around 1.5 nm and 1.2 nm, respectively. A strong oscillation in bubble size is observed after the initial collapse point for the mW/N_2 system, while for the TIP4P/2005/N_2 system the oscillation is much weaker and barely observable. Comparison with the pure water system shows that N_2 only starts to affect bubbly dynamics near the end of the initial collapse point, leading to a less violent collapse indicated by a less negative value for $\frac{dR}{dt}_m$. The partial collapse of a gas bubble and a mitigating effect of gas content within bubble was also found in previous MD simulations for bubble collapse [145].

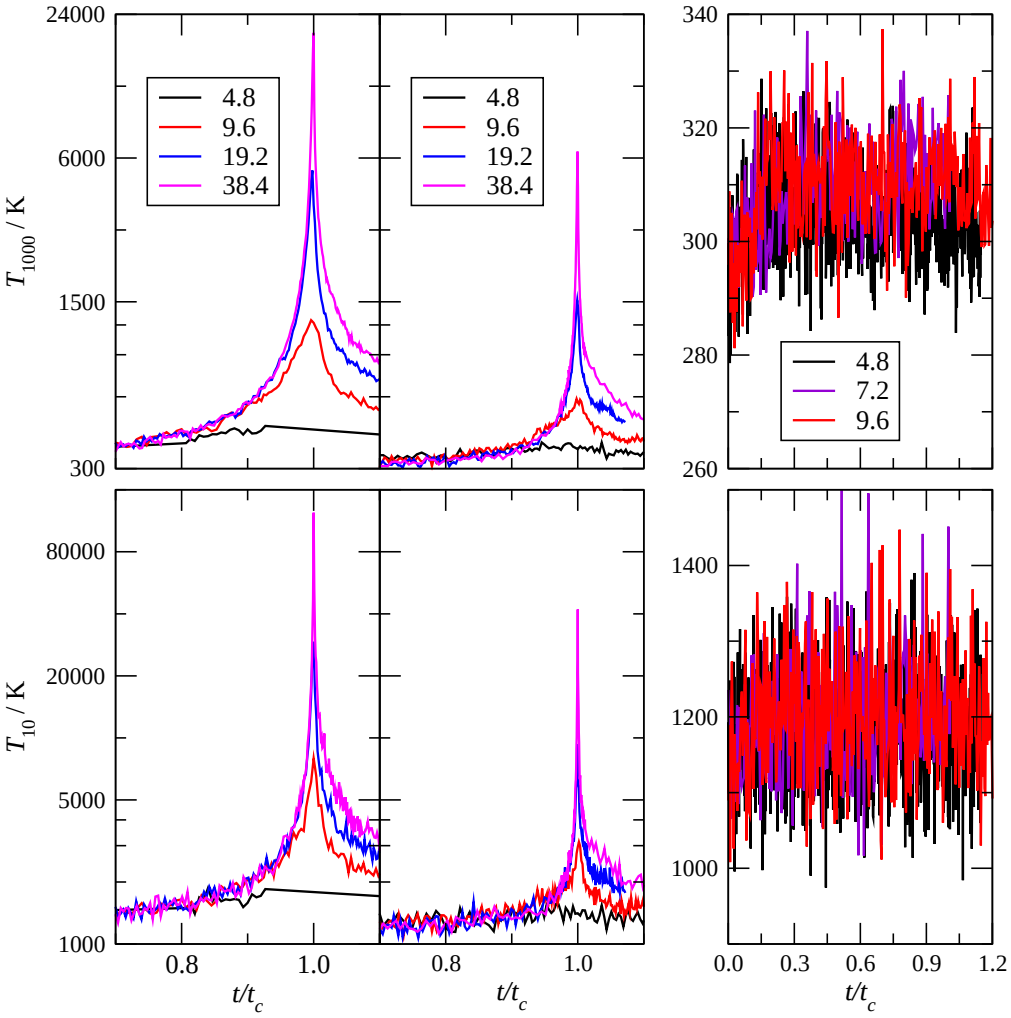


Figure 4.7: Time evolution of the temperature in the innermost shell averaged over approximately 1000 molecules (top) and of the threshold temperature for the top 1% of molecules (bottom). From left to right, data are presented for mW-0.15, mW-0.075, and TIP4P/2005-0.15 systems. Different colors represent systems with different initial bubble radius.

4.3.4 Local extreme conditions

To quantify the temperature evolution, the system is divided into spherical shells of approximately 1000 molecules. The highest temperature is almost always found for the innermost shell (see Figure SC7). Thus time evolution of the average temperature of the innermost shell (T_{1000}) for all pure water systems are presented in Figure 4.7. In addition to the averaged temperature, the top 1% molecular temperature in this shell (T_{10}) was also analyzed because these molecules are most likely to cause materials damage and cavitation erosion. In general, time evolution of T_{1000} and T_{10} are qualitatively similar for the same system. For the mW water systems, a sharp peak is observed in both T_{1000}

and T_{10} near the collapse point. With the increase of system size and initial outside liquid pressure, a higher peak temperature in T_{1000} and T_{10} was observed. Comparison between Figure 4.1 and 4.7 leads to the conclusion that a more violent collapse corresponds to a stronger heating effect for the pure water systems. For the largest bubble size studied in this work, T_{1000} peaks at a value exceeding 20000 K and 6000 K for mW-0.15 and mW-0.075 systems, respectively, which is approximately 67 and 20 times higher than the initial value (around 300 K). The corresponding values for T_{10} are much higher and reach approximately 120000 k and 40000 K, respectively. Compared with the mW systems, a qualitatively different heating behavior was observed for the TIP4P/2005 systems. First, system size does not seem to have an effect on the time evolution of both T_{1000} and T_{10}. Second, both T_{1000} and T_{10} increases slightly at the beginning of collapse, and then flattens out. At the transition point where t/t_c is roughly 0.2, the corresponding absolute time is around 15, 20, and 25 ps for R_0 of 4.8, 7.2, and 9.6 nm, respectively. On Figure 4.4, they approximately correspond to the initial inflection points in $R(t)$. In other words, after passing the first inflection point when viscosity effect begins to dominate, heating is no longer observed.

For water/N_2 systems, various properties of N_2 within bubble were calculated and presented in Figure 4.8. For the mW/N_2 mixture system, α_{N_2} (fraction of nitrogen within bubble) decreases from 1 to below 0.5 as the initial collapse point is approached, indicating movement of N_2 to the surrounding liquid. Meanwhile, ρ_{N_2}, T_{N_2}, and p_{N_2} increases by a large extent, with maximum density around 1000 kg/m^3, temperature exceeding 5000 K, and pressure more than 10000 bar. Previous hybrid RP-MD studies also found extreme heating exceeding 10000 K near the collapse point for violent bubble collapse with inert gas content [164,165]. After the initial collapse point, α_{N_2} jumps back to above 0.8 during the first expansion, indicating that most N_2 molecules previously moved to the liquid region were only trapped near the bubble interface. Along with the expansion, T_{N_2} decreases to around room temperature. For TIP4P/2005/N_2 system, during the initial collapse process, the decrease of α_{N_2} and increase of ρ_{N_2} and p_{N_2} are similar with what was observed for the mW/N_2 system, but with a less extent. However, T_{N_2} did not increase monotonically during the initial collapse process, instead it increases during the initial 25 ps and then flattens out, consistent with what was observed in $T_{1000}(t)$ and $T_{10}(t)$ for the pure TIP4P/2005 water system.

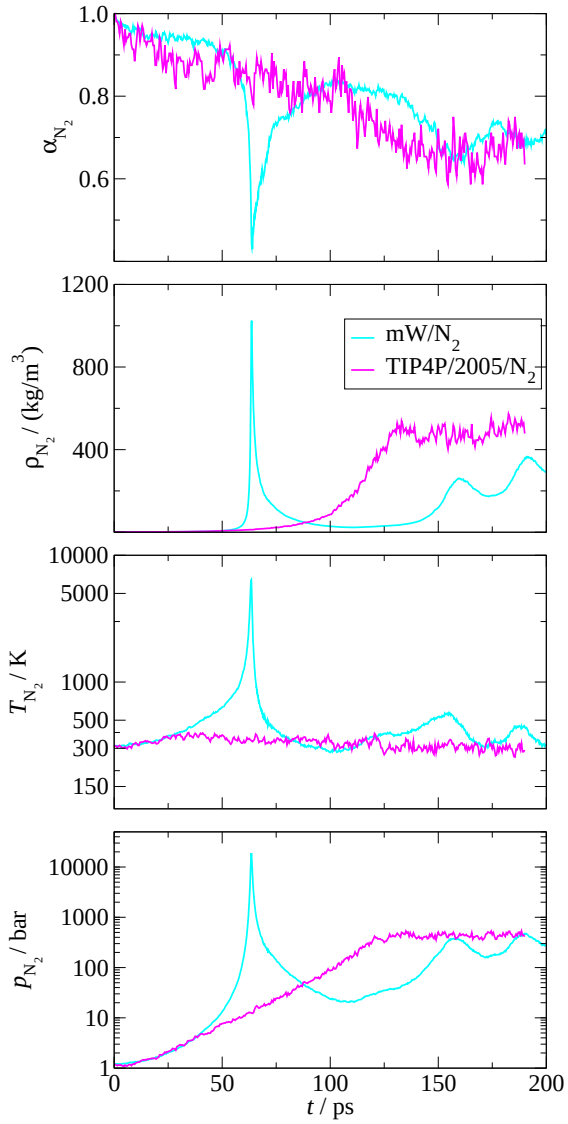


Figure 4.8: Time evolution of fraction (top), density (top middle), temperature (bottom middle), and pressure (bottom) of N_2 within bubble for mW-128-0.15-N_2 and TIP4P/2005-64-0.15-N_2 systems. Pressure is estimated using ideal gas equation of state.

Figure 4.9 presents space-time evolution of density and temperature in the vicinity of the initial collapse point for mW-128-0.15 and TIP4P/2005-64-0.15 water systems. Within the 2 ps preceding the collapse point, bubble radius decreases by around 4 nm for the mW system, while a decrease of around 0.8 nm is observed for the TIP4P/2005 system, indicating a much faster collapse predicted by the mW water model. Along with the fast collapse, local high temperature is observed near the interface. For both water models, the moving interfacial region has an intermediate density and a width of around 0.5 nm. As the collapse point is approached, a slight decrease in the width is observed. At the collapse point, extreme high local density exceeding 1.6 g/cm^3 and

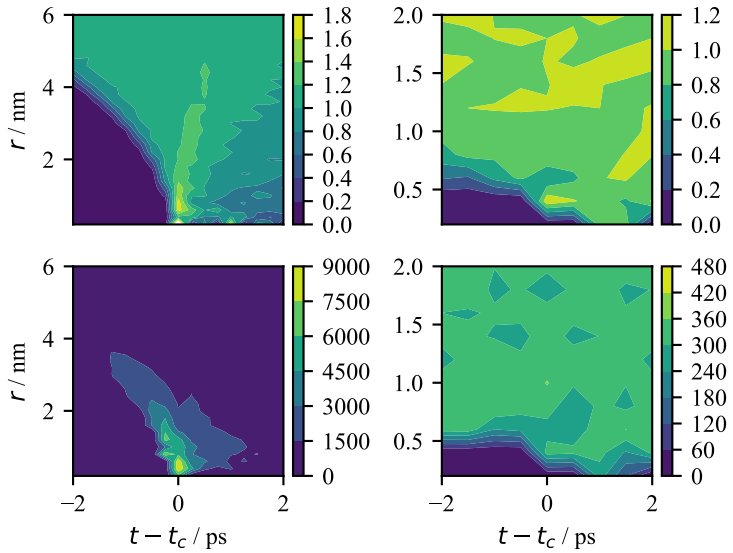


Figure 4.9: Space-time evolution of specific density in units of g/cm^3 (top) and kinetic temperature in units of Kelvin (bottom). Properties are computed on shells of width 0.2 nm in the radial dimension. Left column represents data for mW-128-0.15 system and right column represents data for TIP4P/2005-64-0.15 system.

temperature as high as 9000 K are observed at bubble center for the mW system, while no extreme compression or heating is observed for the TIP4P/2005 system. For the mW system, after the collapse point, high density shells moving away from bubble center are observed within the first 0.5 ps, and the growing inner region with depleted density is also consistent with the rebound observed in Figure 4.1 .

Similarly, Figure 4.10 presents space-time evolution of water density, nitrogen density, and temperature near the initial collapse point for the mixture systems. The V-shaped interface observed in the water density heat map for the mW/N_2 system indicates a clear rebound, consistent with the bubble dynamic result presented in Figure 4.6. Preceding the collapse, the interfacial width remains at around 0.5 nm, while it grows during the rebound. In contrast, the radius of the bubble barely changes within the ± 2 ps near the collapse point for the TIP4P/2005/N_2 system, consistent with a relatively flat $R(t)$ and hardly observable rebound shown in Figure 4.6. For both mW/N_2 and TIP4P/2005/N_2 systems, an incomplete mixing is observed near the collapse point, where water molecules stay outside of the bubble and nitrogen molecules are highly-compressed within bubble. This results from the difference in the time scales between the slow dissolution of nitrogen into the surrounding liquid water and the fast collapsing water front. With respect to heating, the presence of nitrogen leads to a more signifi-

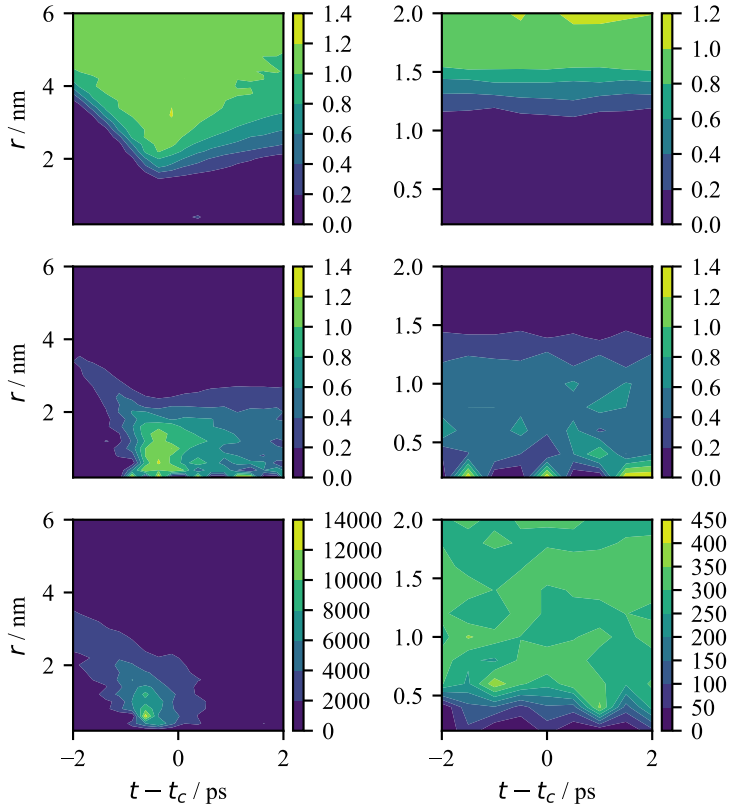


Figure 4.10: Space-time evolution of specific density in units of g/cm^3 for water (top) and nitrogen (middle) and kinetic temperature in units of Kelvin (bottom). Properties are computed on shells of width 0.2 nm in the radial dimension. Left column represents data for mW-128-0.15-N_2 system and right column represents data for TIP4P/2005-64-0.15-N_2 system.

cant heating for the mW/N_2 system, while no significant difference is observed for the TIP4P/2005/N_2 system.

For visualization purpose, Figure SC8 and SC9 presents cross-sectional heat maps of density, temperature, and voxel velocity analyzed on 3D meshes in the vicinity of the collapse point for mW-128-0.15 and the corresponding mixture system. Overall, the conclusion is consistent with what is observed in Figure 4.9 and 4.10.

4.4 Conclusions

Bubble dynamics results obtained from MD simulations for the mW pure water systems indicate a strong system size effect, where violent collapse with supersonic speed and subsequent rebound are only observed for large bubble collapse. For bubbles with same size, a higher initial liquid pressure leads to more violent collapse. A converging behavior

is observed when bubble dynamics results are plotted in reduced unit, and the limiting behavior is obtained by extrapolating the curves linearly to $L = \infty$. The resulting limiting behavior is very close to the collapse behavior of the largest bubble predicted by the MD simulations.

$R(t)$ curves predicted by the TIP4P/2005 water model are qualitatively different with those predicted by the mW water model, where inflection points corresponding to local extrema in $\frac{dR}{dt}(t)$ are observed. To understand the difference, MD simulation results are compared with the numerical solutions of the RP equation, and the overall agreement is satisfactory as long as all terms in the RP equations are accounted for properly. It is found that the inclusion of viscosity term in the RP equation is necessary for observing the initial inflection point in $R(t)$, at which the viscosity effect starts to dominate. The fact that TIP4P/2005 water model predicts a higher viscosity for liquid water explains why there is an initial inflect point in $R(t)$ predicted by the TIP4P/2005 water model, but not the mW water model. Sensitivity test with respect to the value of kinematic viscosity ν_L shows that an improper value of ν_L can not only affect collapse time, but also lead to a qualitatively different $R(t)$ curve.

Temperature analysis is performed using approximately 1000 molecules in the inner most shell for pure water systems. Both the average temperature T_{1000} and the top 1% molecular temperature T_{10} are analyzed along the trajectory. For the mW water systems, with the increase of system size and initial liquid pressure, a higher peak is observed in both T_{1000} and T_{10} near the collapse point, indicating an enhanced heating effect. Extreme heating is observed for large bubble collapse, with maximum peak value of T_{1000} exceeding 20000 K and 5000 K for the mW-0.15 and mW-0.075 systems, respectively. Consistent with the fact that TIP4P/2005 water model predicts a different collapse behavior, the heating behavior is also different between mW and TIP4P/2005 water systems. Instead of peaking near the collapse point, both T_{1000} and T_{10} slightly increases at the beginning of collapse, and then flattens out. The transition point coincides with the initial inflection point in $R(t)$, indicating that when viscosity effect starts to dominate, heating is no longer observed.

For both mW/N_2 and TIP4P/2005/N_2 systems, bubble collapse is incomplete and less violent compared with the corresponding pure water systems. Near the collapse point, an incomplete mixing is observed between water and nitrogen molecules where

nitrogen within bubble is compressed to liquid-like density. This can be explained by a much slower dissolution of N_2 molecules into surrounding liquid water compared with the motion of the collapsing water front. The presence of nitrogen in the mW system leads to a more significant heating effect, while no significant difference is observed for the TIP4P/2005 system.

Chapter 5

Simulating Vapor–liquid equilibria of PH_3, AsH_3, and SbH_3 from first principles[a]

5.1 Introduction

Phosphine (PH_3), arsine (AsH_3), and Stibine (SbH_3) are group V hydrides that are widely used as dopants in the semiconductor industry and precursors in the chemical vapor deposition processes [166, 167]. PH_3 is also used as fumigant [168]. However, all of them are colorless, highly flammable, and highly toxic [169–172]. In addition, SbH_3 also decomposes slowly at room temperature to give metallic antimony and hydrogen [173, 174]. Owing to these hazards, experimental data on their vapor–liquid equilibria (VLE) properties, which is essential for thermodynamic modelling, are relatively sparse, and are available mostly at relatively low temperatures [175–179].

With the increase of computing power and development of efficient simulation algorithms, it is now feasible to predict VLE properties of molecular systems from first principles using Kohn-Sham density functional theory (KS-DFT) [180]. McGrath *et al.* computed the vapor–liquid coexistence curves (VLCCs) of water [181, 182], methanol [183],

[a]This chapter is adapted from my publication accepted by *J. Phys. Chem. C* **2021**: Chen, J. L.; Siepmann, J. I. Simulating Vapor–liquid equilibria of PH_3, AsH_3, and SbH_3 from first principles. The work described in this chapter is a side project that I have completed during my PhD studies. Unlike previous work, the system of interest here is not an aqueous interfacial system, and the energies of the system are not described by molecular-mechanical force fields, but Kohn-Sham density functional theory.

and methane using KS-DFT. More recently, Schienbein and Marx also investigated the VLE of water [184], and Goel *et al*. explored VLE for hydrofluoromethanes [185], carbon tetrafluoride [186], 1,1,1,2-tetrafluoroethane [186], carbon dioxide [187], sulphur dioxde [187], and argon [188]. Most recently, first principles simulations have also been used to investigate the adsorption of nitrogen, carbon dioxide, water, and their mixtures in the Mg-MOF-74 metal–organic framework [189], and the absorption of carbon dioxide in an ionic liquid [190]. Many of these studies investigated the sensitivity of the first-principles predictions with respect to various parameters, such as exchange-correlation functional, system size, basis set, and plane wave cut-off, and found that phase equilibria are quite sensitive to the first-principles description and parameters used for the system, but there is consensus that dispersion corrections are needed for generalized gradient approximation functionals.

In this work, we report first-princiles Monte Carlo simulations for the VLCCs of PH_3, AsH_3, and SbH_3. Since these are rather expensive simulations, system size and exchange-correlation functional dependence are only investigated for PH_3. The third-generation (D3) dispersion correction of Grimme *et al*. [191] is used because dispersion interactions are likely to play a major role for the cohesive energy of these group V hydrides, whereas hydrogen bonding is expected to be weak. Radial distribution functions (RDFs) is used to probe whether directional interactions play a major role.

5.2 Simulation details

Isochoric–isothermal (NVT) Gibbs ensemble Monte Carlo (GEMC) [192] simulations were performed at different temperatures for PH_3, AsH_3, and SbH_3 using the CP2K simulation package [193, 194]. The Quickstep module provides fast and accurate density functional calculations using a mixed Gaussian and plane waves approach [195]. Simulations for all three hydrides were performed using a system size of 48 molecules, the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional [196], and D3 correction [191] with a cut-off distance at 32 Å. Additional simulations were performed using 96 molecules for PH_3 to investigate system size effects. The functional dependence was also investigated for PH_3 using the Becke-Lee-Yang-Parr (BLYP) exchange-correlation functional [197, 198]. The plane wave cut-off was set at 800 Ry. The Goedecker-Teter-

Hutter (GTH) pseudopotentials [199,200] were utilized to account for the core electrons. The molecularly optimized double-ζ valance polarization basis sets were utilized for P and H atoms (DZVP-MOLOPT-GTH) [201], and the shorter-ranged version were utilized for As and Sb atoms (DZVP-MOLOPT-SR-GTH) to reduce computational cost.

For convenience, systems in this work are distinguished by a combination of the chemical's name, exchange-correlation functional, and system size: PH_3-PBE-48, PH_3-PBE-96, PH_3-BLYP-48, AsH_3-PBE-48, and SbH_3-PBE-48. For each of the five systems, GEMC simulations were carried out for five to seven temperatures with the total volume adjusted to yield two simulation boxes of approximately equal size representing the vapor and liquid phases (see Table SD1). It should be noted that for the small system sizes amenable to first principles simulations, temperatures above 0.9 T_c (whereT_c is the critical temperature) are difficult to access due to the large density fluctuations in the near-critical region. Simulation trajectories in terms of specific densities for the two simulation boxes for all five systems at the lowest and highest temperature are shown in Figure SD1 and SD2. Each simulation trajectory is divided into equilibration and production period. The equilibration period generally consists of 200 Monte Carlo cycles (MCCs; one MCC consists of N randomly selected moves, where N is the number of molecules in the system). The classical-mechanical configuration space was sampled via molecular and atomic translations, molecular rotations, and molecule transfers and volume exchanges between the two simulation boxes. To improve sampling efficiency, the maximum displacements for translational, rotational, and volume moves were adjusted during equilibration, and presampling sequences using a simple molecular-mechanical force field were employed [181, 202]. The presampling sequences consisted of 640 trial configurations for swap moves and 8 Monte Carlo moves for translations, rotations, and volume exchanges. Details of the molecular mechanics force fields are provided in the Supporting Information.

For most systems and temperatures, the production period was taken as the subsequent part of the simulation trajectory, ranging from 800 to 1200 MCCs (see Figure SD1 and SD2). To estimate statistical uncertainties, the production period was divided into four blocks, and the standard error of the mean at 95% confidence interval obtained from the block averages was calculated. For the PH_3-PBE-48 system at $T = 290$ K (second highest temperature), large density fluctuations were encountered. Thus a sec-

ond independent simulation was also carried out, and both simulation trajectories were extended to approximately 3000 MCCs. Parts of these trajectories that showed at least one of the phases with large deviations from the mean values for the orthobaric densities were subsequently excluded from the analysis (see Figure SD3). The remaining parts from the two trajectories were combined to constitute a total production period of approximately 4000 MCCs and, again, divided into four blocks for the uncertainty estimation.

Critical points for all systems were estimated using the law of rectilinear diameters and the scaling law for the density [203–205] with a critical exponent of $\beta = 0.326$ and including all temperatures in the analysis. The saturated vapor pressures are not directly available from the first principles Monte Carlo simulations, thus a non-molecule-specific correlation for the compressibility factor as function of reduced temperature was used to convert vapor-phase number densities into pressures [206, 207]. Although use of such a correlation is not completely satisfactory, and we have not been able to confirm this correlation for group V hydrides with experimental saturated vapor pressure and density data, we would like to note that relative deviations of experimental data for water and methane from this correlation are less than 10% (i.e. much smaller than the statistical error in the vapor density for current first-principles simulations).

5.3 Results and discussion

5.3.1 Vapor–liquid equilibria

Experimental [173, 174, 208–211] and simulation data for the VLCCs of PH_3, AsH_3, and SbH_3 are shown in Figure 5.1. Predicted critical points are summarized in Table 5.1. Numerical data for the simulations are provided in Tables SD1. The simulated orthobaric densities and the estimated critical points are consistent within statistical uncertainties between PH_3-PBE-48 and PH_3-PBE-96 systems; that is, a system size of 48 molecules is a reasonable choice for first principles simulations of these compounds. The critical temperature for PH_3 predicted by the PBE functional is 342 ± 9 K, an overestimation of 6% compared to the experimental value (324.5 K [208]). In contrast, the BLYP functional yields T_c = 289 ± 12 K, an underestimation of 11%. To our knowledge, experimental data are not available for the critical density of PH_3. The

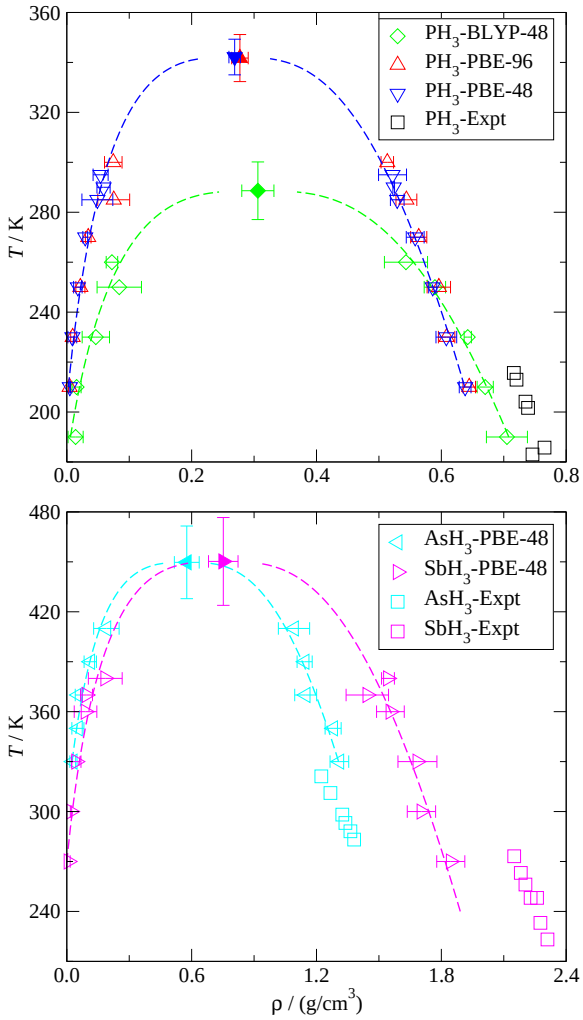


Figure 5.1: Vapor–liquid coexistence curves for PH_3 (top) and AsH_3 and SbH_3 (bottom). Dashed lines and filled symbols represent fits to the scaling law and the calculated critical points. Experimental liquid densities are taken from Ref. [173, 174, 208–211]

critical densities predicted by the PBE and BLYP functionals are 273 ± 13 kg/m^3 (averaged over the two system sizes) and 306 ± 26 kg/m^3, respectively. Both the PBE and BLYP functionals appear to underestimate liquid densities [208, 209] with deviations of about 11% and 7%, respectively, at $T \approx 210$ K. The BLYP functional also yields better agreement for the slope of the liquid density with increasing temperature. An under prediction of the liquid density for the generalized gradient approximation functionals was previously observed for water [181, 182] whereas the opposite was found for methane and methanol [183] Accounting for the deviations in the liquid densities and critical temperature, we cautiously suggest 290 ± 10 kg/m^3 as a reasonable estimate for the critical density of PH_3.

Figures 5.2 and 5.3 show a comparison of the experimental [175–179, 208, 210, 212]

Table 5.1: Critical parameters obtained from first principles simulations

Molecule	Functional	N	T_c / (K)	ρ_c / (kg/m^3)
PH_3	PBE	48	342±7	268±9
PH_3	PBE	96	342±9	277±13
PH_3	BLYP	48	289±12	306±26
AsH_3	PBE	48	450±22	577±60
SbH_3	PBE	48	450±26	752±71

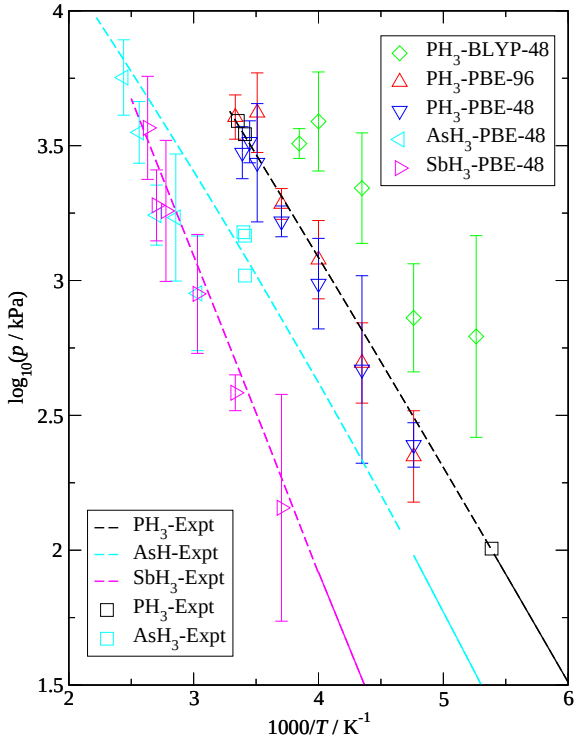


Figure 5.2: Clausius-Clapeyron plot for all systems. Experimental vapor pressures for PH_3, AsH_3 , and SbH_3 are taken from Refs [175–179, 208, 210]. Solid and dashed lines represent Antoine fits below and extrapolations above, respectively, the normal boiling point.

and calculated vapor pressures and heats of vaporization. For PH_3, the simulations with the PBE functional slightly under predict p_{vap} at low reduced temperatures compared to the experimental data [175, 176] but yield excellent agreement at the higher temperatures. Similarly, extrapolation of H_{vap} obtained with the PBE functional yields again excellent agreement with the experimental data [212]. In contrast, the BLYP functional significantly over predicts p_{vap} and under predicts H_{vap}. Overall, the PBE exchange-correlation functional together with the Grimme D3 correction and the DZVP-MOLOPT-GTH basis set allows for a fairly accurate prediction of the VLCC for PH_3.

Although the experimental data [179, 213] for stibine are very limited because it

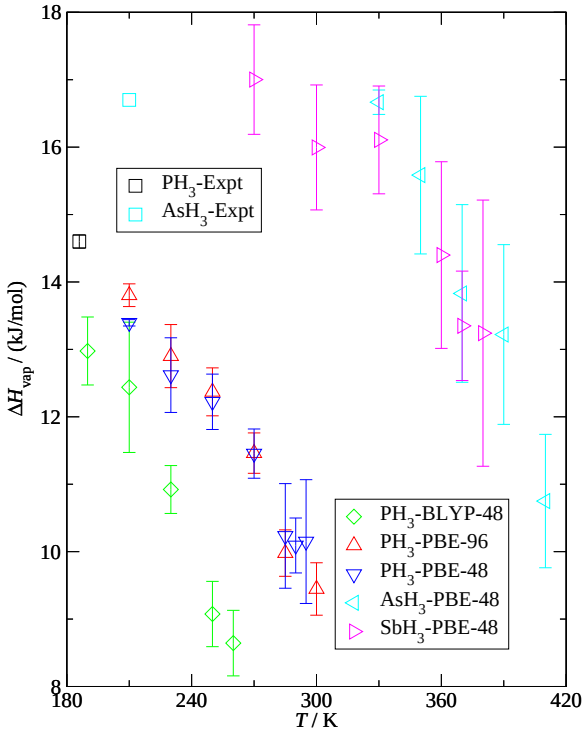


Figure 5.3: Heat of vaporization as a function of temperature for all systems. Experimental data are taken from Refs [177, 212].

already slowly decomposes at room temperature, a comparison between the predictions obtained with the PBE exchange-correlation functional and experiment yields trends similar to those observed for PH_3. The liquid density of SbH_3 is underpredicted by 14% at 270 K, and the slope of ρ_{liq} versus T is also slightly underestimated (see Figure 5.1). On the other hand, the predicted p_{vap} data agree very well with an extrapolation of the experimental data46 (see Figure 5.2). Thus, there is some confidence in the predicted critical temperature of 450 ± 27 K for molecular SbH_3. Here it should be noted that it is the critical point of molecular SbH_3 that is needed as input for an equation of state, i.e., chemical decomposition should not be included in such predictions [214].

In contrast to the fairly accurate predictions found here for PH_3 and SbH_3, the predictions obtained with the PBE exchange-correlation functional for arsine do not follow the same trends. The simulations predict a critical temperature of 450 ± 22 K for AsH_3 that essentially matches the value for SbH_3. For non-hydrogen-bonding compounds, such as the group V hydrides investigated here (see structural analysis), the expectation is that the critical temperature should increase as one progresses downward in the periodic table. Beyond not reproducing the expected trend in the PH_3, AsH_3, and SbH_3,

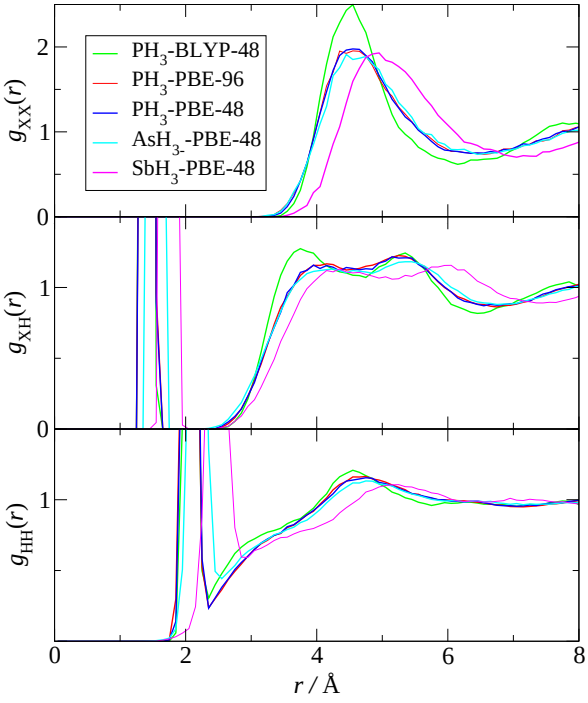


Figure 5.4: Radial distribution functions for X-X (top), X-H (middle), and H-H (bottom) pairs in liquid phase for all systems at $T/T_c \approx 0.73$. A bin width is 0.1 Å was used for this analysis.

the predicted critical temperature for AsH_3 is also $\approx 20\%$ higher than the experimental value of 373.1 K [210]. At 330 K, the predicted liquid density at 330 K is about 7% too high and the vapor pressure is underestimated by a factor of about two [210, 211]. The heat of vaporization also appears to be significantly overestimated because the predicted H_{vap} at 330 K is similar to the experimental value at 210 K (see Figure 5.3). Overall, it is clear that the PBE functional with the Grimme D3 correction does not yield an accurate description of the interactions for AsH_3.

5.3.2 Structural properties

Temperature-dependent data for the intramolecular structure (bond length and bending angle) of PH_3, AsH_3, and SbH_3 obtained for the liquid phase are illustrated in Figure SD4. In general, neither of these properties shows a significant change with increasing temperature. Notably, the H-X-H bend angles for these higher group V hydrides are closer to 90° than to the tetrahedral angle found for ammonia and also water [215]. For PH_3, the PBE and BLYP functionals yield P-H bond lengths of 1.432

and 1.428 Å, respectively, that are slightly longer than the experimental gas-phase value of 1.419 Å [215], whereas the H-P-H angles of 92.5 and 93.1° for the PBE and BLYP functionals, respectively, are slightly smaller than the experimental gas-phase value of 93.7° [215]. Similar trends are also observed for AsH_3, but the deviation in the As-H bond length of 1.550 Åpredicted by the PBE functional is somewhat larger compared to the experimental gas-phase value of 1.523 Å [215], whereas the differences in the H-As-H bending angle are similar (AsH_3: 90.2° versus 91.34° [215];SbH_3: 90.7° versus 91.42° [216]).

The X-X, X-H, and H-H radial distribution functions (RDFs) for all five XH_3 systems at a reduced temperature of approximately 0.73 are shown in Figure 5.4. Comparing PH_3, AsH_3, and SbH_3, the RDFs exhibit qualitatively consistent shapes with a fairly pronounced first peak for the X-X RDF, a weak double peak for the intermolecular region of the X-H RDF, and a shoulder followed by a weak peak for the intermolecular part of the H-H RDF. The RDFs for PH_3-PBE-48 and PH_3-PBE-96 systems are almost indistinguishable from each other, supporting again that system size effects are negligible. On the other hand, the simulations with the BLYP functional for PH_3 yield substantially higher first peaks and slight shifts to shorter separations for X-X and intermolecular X-H RDFs and shoulder and first peak for the intermolecular H-H RDF than for the corresponding RDFs obtained with the PBE functional. These changes agree with the higher liquid density found for the BLYP functional (see Figure 5.1). Somewhat surprisingly, the X-X and intermolecular parts of the X-H and H-H RDFs for AsH_3-PBE-48 are very close to those for PH_3, whereas the bonded X-H and H-H peaks are shifted outward for AsH_3 (see discussion of molecular structure). These structural features also point to a significant deficiency for the PBE-D3 description of AsH_3, as was also concluded from the VLCC.

The number integrals for the first solvation shell of the X atoms (as defined by the first minimum of g_{XX}) yield values of 11.2, 10.6, 10.7, and 11.2 for PH_3-BLYP-48, PH_3-PBE, AsH_3, and SbH_3, respectively, but it should be noted that the location of the first minimum is not very precises due to the noise in the RDFs. Coordination numbers near a value of 12 are indicative of the liquid-phase packing of noble gases and methane. In marked contrast to the tetrahedral coordination in liquid water, neutron diffraction experiments and first principles molecular dynamics simulations [217–219] also found

a coordination number of 12 to 13 for NH_3. Compared to the RDFs for NH_3 [217], the RDFs for the higher group V hydrides do not show a shoulder in the X-X RDF and also not a pre-peak for the H-H RDF. These observations also indicate a lack of hydrogen bonding for PH_3, AsH_3, and SbH_3 that is consistent with electronic structure calculations for their dimers and spectroscopic data that do not indicate a significant preference for self-association in their liquid phases [215].

5.4 Conclusions

The vapor–liquid equilibria of PH_3, AsH_3, and SbH_3 are investigated using first-principles Gibbs ensemble Monte Carlo simulations where the interactions are described by KS-DFT with generalized gradiant approximation functionals and the Grimme D3 dispersion correction. For PH_3 and SbH_3, the PBE exchange-correlation functional allows for a fairly accurate description of the saturated vapor pressure, the critical temperature, and the heat of vaporization (only for PH_3 because experimental data for the latter two properties are lacking for SbH_3). However, the saturated liquid densities for PH_3 and SbH_3 are somewhat underestimated with the PBE functional. For PH3, the BLYP functional yields a slightly improved liquid density at the cost of a significant over prediction of the saturated vapor pressure and under prediction of the heat of vaporization. The predictions of the VLCC properties for AsH_3 with the PBE functional do not follow the same trends as for PH_3 and SbH_3, and the location of the first peak in the As-As RDF does not fall in between those for PH_3 and SbH_3. Thus, it appears likely that there are some deficiencies in the PBE-D3 description of AsH_3. The liquid-phase RDFs of the three group V hydrides are indicative of the packing of spherical, non-polar molecules, and there is no evidence for hydrogen bonding.

Bibliography

[1] M. S. Kamal, I. A. Hussein, and A. S. Sultan, Review on surfactant flooding: phase behavior, retention, IFT, and field applications, *Energy & Fuels*, **2017**, *31*, 7701–7720.

[2] C. Negin, S. Ali, and Q. Xie, Most common surfactants employed in chemical enhanced oil recovery, *Petroleum*, **2017**, *3*, 197–211.

[3] T. J. Young, C. W. Nelson, C. Lyons-Bell, and J. D. Moore, High temperature and pressure water/oil interfacial tension benchmark data for the ninth industrial fluid properties simulation challenge, *Fluid Ph. Equilibria*, **2018**, *476*, 6–8.

[4] G. Wiegand and E. U. Franck, Interfacial tension between water and non-polar fluids up to 473 K and 2800 bar, *Ber. Bunsen. Phys. Chem.*, **1994**, *98*, 809–817.

[5] B. Y. Cai, J. T. Yang, and T. M. Guo, Interfacial tension of hydrocarbon + water/brine systems under high pressure, *J. Chem. Eng. Data*, **1996**, *41*, 493–496.

[6] S. Zeppieri, J. Rodríguez, and A. L. López de Ramos, Interfacial tension of alkane + water systems, *J. Chem. Eng. Data*, **2001**, *46*, 1086–1088.

[7] H. Y. Jennings Jr, The effect of temperature and pressure on the interfacial tension of benzene-water and normal decane-water, *J. Colloid Interface Sci.*, **1967**, *24*, 323–329.

[8] A. S. Michaels and E. A. Hauser, Interfacial tension at elevated pressure and temperature. II interfacial properties of hydrocarbon-water systems, *J. Phys. Chem.*, **1951**, *55*, 408–421.

[9] A. Georgiadis, G. Maitland, J. P. M. Trusler, and A. Bismarck, Interfacial tension measurements of the (H_2O + *n*-decane + CO_2) ternary system at elevated pressures and temperatures, *J. Chem. Eng. Data*, **2011**, *56*, 4900–4908.

[10] R. Pecha and B. Gompf, Microimplosions: cavitation collapse and shock wave emission on a nanosecond time scale, *Phys. Rev. Lett.*, **2000**, *84*, 1328.

[11] A. Philipp and W. Lauterborn, Cavitation erosion by single laser-produced bubbles, *J. Fluid Mech.*, **1998**, *361*, 75–116.

[12] M. Kjeldsen, R. E. A. Arndt, and M. Effertz, Spectral characteristics of sheet/cloud cavitation, *J. Fluids Eng.*, **2000**, *122*(3), 481–487.

[13] https://en.wikipedia.org/wiki/Cavitation (accessed Dec 7 2020).

[14] M. Dular, T. Požar, J. Zevnik, et al., High speed observation of damage created by a collapse of a single cavitation bubble, *Wear*, **2019**, *418*, 13–23.

[15] Industrial Fluid Property Simulation Collective; http://fluidproperties.org.

[16] M. Bavière, *Basic Concepts in Enhanced Oil Recovery Processes;* Springer: Netherland, 1991.

[17] S. S. Chadwick, *Ullmann's Encyclopedia of Industrial Chemistry;* Wiley: New York, 2006.

[18] W. Sachs and V. Meyn, Pressure and temperature dependence of the surface tension in the system natural gas/water principles of investigation and the first precise experimental data for pure methane/water at 25 °C up to 46.8 MPa, *Colloid Surface A*, **1995**, *94*, 291–301.

[19] A. R. van Buuren, S. J. Marrink, and H. J. C. Berendsen, A molecular dynamics study of the decane/water interface, *J. Phys. Chem.*, **1993**, *97*, 9206–9212.

[20] D. Michael and I. Benjamin, Solute orientational dynamics and surface roughness of water/hydrocarbon interfaces, *J. Phys. Chem.*, **1995**, *99*, 1530–1536.

[21] J. L. Rivera, C. McCabe, and P. T. Cummings, Molecular simulations of liquid–liquid interfacial properties: water–*n*-alkane and water–methanol– *n*-alkane systems, *Phys. Rev. E.*, **2003**, *67*, 011603.

[22] J. P. Nicolas and N. R. De Souza, Molecular dynamics study of the *n*-hexane–water interface: towards a better understanding of the liquid–liquid interfacial broadening, *J. Chem. Phys.*, **2004**, *120*, 2464–2469.

[23] F. Biscay, A. Ghoufi, V. Lachet, and P. Malfreyt, Monte Carlo calculation of the methane-water interfacial tension at high pressures, *J. Chem. Phys.*, **2009**, *131*, 124707.

[24] G. C. Maitland, M. Rigby, and E. B. Smith, *Intermolecular Forces: Their Origin and Determination;* Pergamon Press, Oxford, 1987.

[25] J. L. F. Abascal and C. Vega, A general purpose model for the condensed phase of water: TIP4P/2005, *J. Chem. Phys.*, **2005**, *123*, 234505.

[26] C. Vega and E. De Miguel, Surface tension of the most popular models of water by using the test-area simulation method, *J. Chem. Phys*, **2007**, *126*, 154707.

[27] M. G. Martin and J. I. Siepmann, Transferable potentials for phase equilibria. 1. united-atom description of *n*-alkanes, *J. Phys. Chem. B*, **1998**, *102*, 2569–2577.

[28] B. Chen, J. I. Siepmann, K. J. Oh, and M. L. Klein, Simulating vapor–liquid nucleation of *n*-alkanes, *J. Chem. Phys*, **2002**, *116*, 4317–4329.

[29] N. Rai and J. I. Siepmann, Transferable potentials for phase equilibria. 9. explicit-hydrogen description of benzene and 5-membered and 6-membered heterocyclic aromatic compounds, *J. Phys. Chem. B*, **2007**, *111*, 10790–10799.

[30] C. D. Wick, J. I. Siepmann, W. L. Klotz, and M. R. Schure, Temperature effects on the retention of *n*-alkanes and arenes in helium–squalane gas–liquid chromatography: experiment and molecular simulation, *J. Chromatogr. A*, **2002**, *954*, 181–190.

[31] J. G Kirkwood and F. P Buff, The statistical mechanical theory of surface tension, *J. Chem. Phys.*, **1949**, *17*, 338–343.

[32] A. Ghoufi, P. Malfreyt, and D. J. Tildesley, Computer modelling of the surface tension of the gas-liquid and liquid-liquid interface, *Chem. Soc. Rev.*, **2016**, *45*, 1387–1409.

[33] J. Janeček, Long range corrections in inhomogeneous simulations, *J. Phys. Chem. B*, **2006**, *110*, 6264–6269.

[34] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids;* Oxford university press: Oxford, 1989.

[35] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, Equation of state calculation by fast computing machines, *J. Chem. Phys.*, **1953**, *21*, 1087–1092.

[36] J. A. Barker and R. O. Watts, Structure of water: a Monte Carlo calculation, *Chem. Phys. Lett.*, **1969**, *3*, 144–145.

[37] J. I. Siepmann and D. Frenkel, Configurational bias Monte Carlo: a new sampling scheme for flexible chains, *Mol. Phys.*, **1992**, *75*, 59–70.

[38] T. J. H. Vlugt, M. G. Martin, B. Smit, J. I. Siepmann, and R. Krishna, Improving the efficiency of the configurational-bias Monte Carlo algorithm, *Mol. Phys.*, **1998**, *94*, 727–733.

[39] M. G. Martin and J. I. Siepmann, Novel configurational-bias Monte Carlo method for branched molecules. transferable potentials for phase equilibria.2. united-atom description of branched alkanes, *J. Phys. Chem. B*, **1999**, *103*, 4508–4517.

[40] G. C. A. M. Mooij, D. Frenkel, and B. Smit, Direct simulation of phase equilibria of chain molecules, *J. Phys.: Condens. Matter*, **1992**, *4*, L255–L259.

[41] K. Esselink and L. D. J. C. Loyens, Parallel Monte Carlo simulations, *Phys. Rev. E*, **1995**, *51*, 1560–1568.

[42] A. Z. Panagiotopoulos, N. Quirke, M. Stapleton, and D. J. Tildesley, Phase equilibria by simulation in the Gibbs ensemble: Alternative derivation, generalization and application to mixture and membrane equilibria, *Mol. Phys.*, **1988**, *63*, 527–545.

[43] A. Z. Panagiotopoulos, Direct determination of phase coexistence properties of fluids by Monte Carlo simulation in a new ensemble, *Mol. Phys.*, **1987**, *61*, 813–826.

[44] J. J. De Pablo and J. M. Prausnitz, Phase equilibria for fluid mixture from Monte-Carlo simulation, *Fluid Phase Equilib*, **1989**, *53*, 177–189.

[45] J. I. Siepmann and I. R. McDonald, Monte Carlo simulations of mixed monolayers, *Mol. Phys.*, **1992**, *75*, 255–259.

[46] M. G. Martin and J. I. Siepmann, Predicting multicomponent phase equilibria and free energies of transfer for alkanes by molecular simulation, *J. Am. Chem. Soc.*, **1997**, *119*, 8921–8924.

[47] P. Bai and J. I. Siepmann, Selective adsorption from dilute solutions: Gibbs ensemble Monte Carlo simulations, *Fluid Phase Equilib.*, **2013**, *351*, 1–6.

[48] M. G. Martin and J. I. Siepmann, Calculating Gibbs free energies of transfer from Gibbs ensemble Monte Carlo simulations, *Theor. Chem. Acc.*, **1998**, *99*(5), 347–350.

[49] L. R. Pratt and D. Chandler, Theory of the hydrophobic effect, *J. Chem. Phys.*, **1977**, *67*, 3683–3704.

[50] A. R. van Buuren, S. J. Marrink, and H. J. C. Berendsen, A molecular dynamics study of the decane/water interface, *J. Phys. Chem.*, **1993**, *97*, 9206–9212.

[51] E. Johansson, K. Bolton, D. N. Theodorou, and P. Ahlström, Monte Carlo simulations of equilibrium solubilities and structure of water in *n*-alkanes and polyethylene, *J. Chem. Phys.*, **2007**, *126*, 224902.

[52] H. Docherty, A. Galindo, C. Vega, and E. Sanz, A potential model for methane in water describing correctly the solubility of the gas and the properties of the methane hydrate, *J. Chem. Phys.*, **2006**, *125*, 074510.

[53] D. Ballal, P. Venkataraman, W. A. Fouad, K. R. Cox, and W. G. Chapmann, Isolating the non-polar contribution to the intermolecular potential for water–alkane interactions, *J. Chem. Phys.*, **2014**, *141*, 064905.

[54] K. S. Suslick, Sonochemistry, *Science*, **1990**, *247*, 1439–1445.

[55] C. T. Crowe, *Multiphase Flow Handbook;* CRC press, 2005.

[56] G. Yadigaroglu and G. F. Hewitt, *Introduction to Multiphase Flow: Basic Concepts, Applications and Modelling;* Springer, 2017.

[57] P. S. Russell, D. R. Giosio, J. A. Venning, B. W. Pearce, P. A. Brandner, and S. Ceccio; Microbubble generation from condensation and turbulent breakup of sheet cavitation; In *the 31st Symposium on Naval Hydrodynamics*, pages 11–16, 2016.

[58] K. Binder, B. J. Block, P. Virnau, and A. Tröster, Beyond the van der waals loop: What can be learned from simulating lennard-jones fluids inside the region of phase coexistence, *Am. J. Phys.*, **2012**, *80*, 1099–1109.

[59] M. Thommes, K. Kaneko, A. V. Neimark, J. P. Olivier, F. Rodriguez-Reinoso, J. Rouquerol, and K. S. W. Sing, Physisorption of gases, with special reference to the evaluation of surface area and pore size distribution (IUPAC technical report), *Pure Appl. Chem.*, **2015**, *87*, 1051–1069.

[60] J. L. F. Abascal, M. A. González, J. L. Aragones, and C. Valeriani, Homogeneous bubble nucleation in water at negative pressure: A voronoi polyhedra analysis, *J. Chem. Phys.*, **2013**, *138*, 084508.

[61] M. A. González, G. Menzl, J. L. Aragones, P. Geiger, F. Caupin, J. L. F. Abascal, C. Dellago, and C. Valeriani, Detecting vapour bubbles in simulations of metastable water, *J. Chem. Phys.*, **2014**, *141*, 18C511.

[62] G. Menzl, M. A. Gonzalez, P. Geiger, F. Caupin, J. L. F. Abascal, C. Valeriani, and C. Dellago, Molecular mechanism for cavitation in water under tension, *Proc. Natl. Acad. Sci.*, **2016**, *113*, 13582–13587.

[63] A. Niedźwiedzka, G. H. Schnerr, and W. Sobieski, Review of numerical models of cavitating flows with the use of the homogeneous approach, *Arch. Thermodyn.*, **2016**, *37*, 71–88.

[64] E. Goncalves and R. F. Patella, Numerical simulation of cavitating flows with homogeneous models, *Comput. Fluids*, **2009**, *38*, 1682–1696.

[65] A. Gnanaskandan and K. Mahesh, A numerical method to simulate turbulent cavitating flows, *Int. J. Multiph. Flow*, **2015**, *70*, 22–34.

[66] B. R. Shin, Y. Iwata, and T. Ikohagi, Numerical simulation of unsteady cavitating flows using a homogeneous equilibrium model, *Comput. Mech.*, **2003**, *30*, 388–395.

[67] S. W. Kieffer, Sound speed in liquid-gas mixtures: Water-air and water-steam, *J. Geophys. Res.*, **1977**, *82*, 2895–2904.

[68] B. G. Pokusaev, E. A. Tairov, and S. A. Vasilyev, Low frequency pressure waves in a vapor-liquid medium with a fixed layer of spherical particles, *Acoust. Phys.*, **2010**, *56*, 306–312.

[69] M. M. Awad and Y. S. Muzychka, Effective property models for homogeneous two-phase flows, *Exp. Therm. Fluid Sci.*, **2008**, *33*, 106–113.

[70] N. Z. Aung and T. Yuwono, Evaluation of mixture viscosity models in the prediction of two-phase flow pressure drops, *ASEAN J. Sci. Technol. Dev.*, **2012**, *29*, 115–128.

[71] G. Ruocco, M. Sampoli, A. Torcini, and R. Vallauri, Molecular dynamics results for stretched water, *J. Chem. Phys.*, **1993**, *99*, 8095–8104.

[72] P. A. Netz, F. W. Starr, H. E. Stanley, and M. C. Barbosa, Static and dynamic properties of stretched water, *J. Chem. Phys.*, **2001**, *115*, 344–348.

[73] P. A. Netz, F. Starr, M. C. Barbosa, and H. E. Stanley, Computer simulation of dynamical anomalies in stretched water, *Braz. J. Phys.*, **2004**, *34*, 24–31.

[74] H. E. Stanley, M. C. Barbosa, S. Mossa, P. A. Netz, F. Sciortino, F. W. Starr, and M. Yamada, Statistical physics and liquid water at negative pressures, *Physica A*, **2002**, *315*, 281–289.

[75] J. L. F. Abascal and C. Vega, A general purpose model for the condensed phase of water: TIP4P/2005, *J. Chem. Phys.*, **2005**, *123*, 234505.

[76] P. Ewald, Die berechnung optischer und elektrostatischer gitterpotentiale, *Ann. Phys.*, **1921**, *64*, 253–287.

[77] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids;* Oxford University Press: Oxford, UK, 1987.

[78] R. W. Hockney and W. E. James, *Computer Simulation Using Particles;* CRC Press, 1988.

[79] R. J. Sadus, *Molecular Simulation of Fluids;* Elsevier Science: Amsterdam, 1999.

[80] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithm to Applications;* Elsevier Science, 2001.

[81] V. Molinero and E. B. Moore, Water modeled as an intermediate element between carbon and silicon, *J. Phys. Chem. B*, **2009**, *113*, 4008–4016.

[82] A. Z. Panagiotopoulos, Direct determination of phase coexistence properties of fluids by Monte Carlo simulation in a new ensemble, *Mol. Phys*, **1987**, *61*, 813–826.

[83] A. Z. Panagiotopoulos, N. Quirke, M. Stapleton, and D. J. Tildesley, Phase equilibria by simulation in the Gibbs ensemble: Alternative derivation, generalization and application to mixture and membrane equilibria, *Mol. Phys.*, **1988**, *63*, 527–545.

[84] J. I. Siepmann, M. G. Martin, B. Chen, C. D. Wick, J. M. Stubbs, J. J. Potoff, B. L. Eggimann, M. J. McGrath, X. S. Zhao, K. E. Anderson, J. L. Rafferty, N. Rai, K. A. Maerzke, S. J. Keasler, P. Bai, E. O. Fetisov, M. S. Shah, Q. P. Chen, R. F. DeJaco, J. L. Chen, and X. Bai, Monte Carlo for Complex Chemical Systems — Minnesota, Versions 16.1; Minneapolis; MN (2016).

[85] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, Equation of state calculations by fast computing machines, *J. Chem. Phys.*, **1953**, *21*, 1087–1092.

[86] J. A. Barker and R. O. Watts, Structure of water: A Monte Carlo calculation, *Chem. Phys. Lett.*, **1969**, *3*, 144–145.

[87] T. J. H. Vlugt, M. G. Martin, B. Smit, J. I. Siepmann, and R. Krishna, Improving the efficiency of the configurational-bias Monte Carlo algorithm, *Mol. Phys*, **1998**, *94*, 727–733.

[88] S. Plimpton, Fast parallel algorithms for short-range molecular dynamics, *J. Comput. Phys.*, **1995**, *117*, 1–19.

[89] W. G. Hoover, Canonical dynamics: Equilibrium phase-space distributions, *Phys. Rev. A*, **1985**, *31*, 1695.

[90] J.-P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of *n*-alkanes, *J. Comput. Phys.*, **1977**, *23*, 327–341.

[91] J. G. Kirkwood and F. P. Buff, The statistical mechanical theory of surface tension, *J. Chem. Phys*, **1949**, *17*, 338–343.

[92] L. Martínez, R. Andrade, E. G. Birgin, and J. Martínez, Packmol: a package for building initial configurations for molecular dynamics simulations, *J. Comput. Chem*, **2009**, *30*, 2157–2164.

[93] J. Martínez and L. Martínez, Packing optimization for automated generation of complex system's initial configurations for molecular dynamics and docking, *J. Comput. Chem*, **2003**, *24*, 819–825.

[94] F. H. Stillinger Jr, Rigorous basis of the frenkel-band theory of association equilibrium, *J. Chem. Phys.*, **1963**, *38*, 1486–1494.

[95] H. Arkın and W. Janke, Gyration tensor based analysis of the shapes of polymer chains in an attractive spherical cage, *J. Chem. Phys*, **2013**, *138*, 054904.

[96] D. H. Tsai, The virial theorem and stress calculation in molecular dynamics, *J. Chem. Phys.*, **1979**, *70*, 1375–1382.

[97] A. P. Thompson, S. J. Plimpton, and W. Mattson, General formulation of pressure and stress tensor for arbitrary many-body interaction potentials under periodic boundary conditions, *J. Chem. Phys.*, **2009**, *131*, 154107.

[98] J. M. Smith, H. C. Van Ness, and M. M. Abbott, *Introduction to chemical engineering thermodynamics;* McGraw-Hill: Boston, 2005.

[99] D. A. McQuarrie, *Statistical Mechanics;* Sausalito, Calif.: University Science Books, 2004.

[100] R. Kubo, Statistical-mechanical theory of irreversible processes. i. general theory and simple applications to magnetic and conduction problems, *J. Phys. Soc. Jpn*, **1957**, *12*, 570–586.

[101] J. Petravic and P. Harrowell, Linear response theory for thermal conductivity and viscosity in terms of boundary fluctuations, *Phys. Rev. E*, **2005**, *71*, 061201.

[102] Y. Zhang, A. Otani, and E. J. Maginn, Reliable viscosity calculation from equilibrium molecular dynamics simulations: A time decomposition method, *J. Chem. Theory Comput.*, **2015**, *11*, 3537–3546.

[103] Nist chemistry webbook: Water.

[104] C. Vega and E. De Miguel, Surface tension of the most popular models of water by using the test-area simulation method, *J. Chem. Phys.*, **2007**, *126*, 154707.

[105] C. Vega and J. L. F. Abascal, Simulating water with rigid non-polarizable models: A general perspective, *Phys. Chem. Chem. Phys*, **2011**, *13*, 19663–19688.

[106] M. H. Factorovich, V. Molinero, and D. Scherlis, A simple grand canonical approach to compute the vapor pressure of bulk and finite size systems, *J. Chem. Phys.*, **2014**, *140*, 064111.

[107] M. A. González, C. Valeriani, F. Caupin, and J. L. F. Abascal, A comprehensive scenario of the thermodynamic anomalies of water using the tip4p/2005 model, *J. Chem. Phys*, **2016**, *145*, 054505.

[108] P. G. Debenedetti, *Metastable liquids: concepts and principles;* Vol. 1; Princeton university press, 1996.

[109] J. W. Cahn and J. E. Hilliard, Free energy of a nonuniform system. i. interfacial free energy, *J. Chem. Phys.*, **1958**, *28*, 258–267.

[110] M. Matsumoto and K. Tanaka, Nano bubble—size dependence of surface tension and inside pressure, *Fluid Dyn. Res.*, **2008**, *40*, 546.

[111] H. Yan, J. Wei, S. Cui, S. Xu, Z. Sun, and R. Zhu, On the applicability of young–laplace equation for nanoscale liquid drops, *Russ. J. Phys. Chem. A*, **2016**, *90*, 635–640.

[112] T. Furtenbacher, T. Szidarovszky, J. Hrubỳ, A. A. Kyuberis, N. F. Zobov, O. L. Polyansky, J. Tennyson, and A. G. Császár, Definitive ideal-gas thermochemical functions of the h_2 ^{16}o molecule, *J. Phys. Chem. Ref. Data*, **2016**, *45*, 043104.

[113] W. M. Haynes, *CRC Handbook of Chemistry and Physics;* CRC Press, 2014.

[114] V. G. Baidakov, S. P. Protsenko, and Z. R. Kozlova, The isochoric heat capacity of a metastable lennard-jones fluid, *Chem. Phys. Lett.*, **2007**, *447*, 236–240.

[115] Y. E. Altabet, R. S. Singh, F. H. Stillinger, and P. G. Debenedetti, Thermodynamic anomalies in stretched water, *Langmuir*, **2017**, *33*, 11771–11778.

[116] D. Dhabal, C. Chakravarty, V. Molinero, and H. K. Kashyap, Comparison of liquid-state anomalies in stillinger-weber models of water, silicon, and germanium, *J. Chem. Phys.*, **2016**, *145*, 214502.

[117] P. M. de Hijes, E. Sanz, L. Joly, C. Valeriani, and F. Caupin, Viscosity and self-diffusion of supercooled and stretched water from molecular dynamics simulations, *arXiv:1805.11957*, **2018**.

[118] A. E. Dukler, M. Wicks III, and R. G. Cleveland, Frictional pressure drop in two-phase flow: B. an approach through similarity analysis, *AIChE Journal*, **1964**, *10*, 44–51.

[119] D. R. H. Beattie and P. B. Whalley, Simple two-phase frictional pressure drop calculation method, *Int. J. Multiphase Flow;(United Kingdom)*, **1982**, *8*.

[120] Albert Einstein, Eine neue bestimmung der molekuldimensionen, *Ann. Phys.*, **1911**, *34*, 591–592.

[121] I. C. Yeh and G. Hummer, System-size dependence of diffusion coefficients and viscosities from molecular dynamics simulations with periodic boundary conditions, *J. Phys. Chem. B*, **2004**, *108*, 15873–15879.

[122] D. M. Heyes, M. J. Cass, J. G. Powles, and W. A. B. Evans, Self-diffusion coefficient of the hard-sphere fluid: System size dependence and empirical correlations, *J. Phys. Chem. B*, **2007**, *111*, 1455–1464.

[123] R. E. Zeebe, On the molecular diffusion coefficients of dissolved co_2, hco_3^-, and co_3^{2-} and their dependence on isotopic mass, *Geochim. Cosmochim. Acta*, **2011**, *75*, 2483–2498.

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

[125] C. Vega, J. L. F. Abascal, M. M. Conde, and J. L. Aragones, What ice can teach us about water interactions: a critical comparison of the performance of different water models, *Faraday Discuss.*, **2009**, *141*, 251–276.

[126] K. Krynicki, C. D. Green, and D. W. Sawyer, Pressure and temperature dependence of self-diffusion in water, *Faraday Discuss. Chem. Soc.*, **1978**, *66*, 199–208.

[127] P. Bai, E. Haldoupis, P. J. Dauenhauer, M. Tsapatsis, and J. I. Siepmann, Understanding diffusion in hierarchical zeolites with house-of-cards nanosheets, *ACS nano*, **2016**, *10*, 7612–7618.

[128] K. Davitt, E. Rolley, F. Caupin, A. Arvengas, and S. Balibar, Equation of state of water under negative pressure, *J. Chem. Phys.*, **2010**, *133*, 174507.

[129] F. Caupin, A. Arvengas, K. Davitt, M. E. M. Azouzi, K. I. Shmulovich, C. Ramboz, D. A. Sessoms, and A. D. Stroock, Exploring water and other liquids at negative pressure, *J. Phys. Condens. Matter*, **2012**, *24*, 284110.

[130] J. W. Biddle, R. S. Singh, E. M. Sparano, F. Ricci, M. A. González, C. Valeriani, J. L. F. Abascal, P. G. Debenedetti, M. A. Anisimov, and F. Caupin, Two-structure thermodynamics for the tip4p/2005 model of water covering supercooled and deeply stretched regions, *J. Chem. Phys.*, **2017**, *146*, 034502.

[131] C. F. Delale, K. Okita, and Y. Matsumoto, Steady-state cavitating nozzle flows with nucleation, *J. Fluids Eng.*, **2005**, *127*, 770–777.

[132] G. Wang, I. Senocak, W. Shyy, T. Ikohagi, and S. Cao, Dynamics of attached turbulent cavitating flows, *Prog. Aerosp. Sci.*, **2001**, *37*, 551–581.

[133] P. Kumar and R. P. Saini, Study of cavitation in hydro turbines—a review, *Renewable Sustainable Energy Rev.*, **2010**, *14*, 374–383.

[134] B. K. Sreedhar, S. K. Albert, , and A. B. Pandit, Cavitation damage: Theory and measurements–a review, *Wear*, **2017**, *372*, 177–196.

[135] W. Lauterborn and T. Kurz; The bubble challenge for high-speed photography; In *The Micro-World Observed by Ultra High-Speed Cameras*, pages 19–47. Springer, 2018.

[136] M. P. Brenner, S. Hilgenfeldt, and D. Lohse, Single-bubble sonoluminescence, *Rev. Mod. Phys.*, **2002**, *74*, 425.

[137] L. R. OMFRS, On the pressure developed in a liquid during the collapse of a spherical cavity, *Philos. Mag.*, **1917**, *6*, 94–98.

[138] M. S. Plesset, The dynamics of cavitation bubbles, *J. Appl. Mech.*, **1949**, *16*, 277–282.

[139] J.P. Franc; In *Fluid dynamics of cavitation and cavitating turbopumps*, pages 1–41. Springer, 2007.

[140] A. Gnanaskandan and K. Mahesh, A numerical method to simulate turbulent cavitating flows, *Int. J. Multiph. Flow*, **2015**, *70*, 22–34.

[141] R. K. Shukla, C. Pantano, and J. B. Freund, An interface capturing method for the simulation of multi-phase compressible flows, *J. Comput. Phys.*, **2010**, *229*, 7411–7439.

[142] M. Pelanti and K. Shyue, A mixture-energy-consistent six-equation two-phase numerical model for fluids with interfaces, cavitation and evaporation waves, *J. Comput. Phys*, **2014**, *259*, 331–357.

[143] J. L. F. Abascal, M. A. Gonzalez, J. L. Aragones, and C. Valeriani, Homogeneous bubble nucleation in water at negative pressure: A voronoi polyhedra analysis, *J. Chem. Phys*, **2013**, *138*, 084508.

[144] J. Diemand, R. Angélil, K. K. Tanaka, and H. Tanaka, Large scale molecular dynamics simulations of homogeneous nucleation, *J. Chem. Phys.*, **2013**, *139*, 074309.

[145] A. Shekhar, K. Nomura, R. K. Kalia, A. Nakano, and P. Vashishta, Nanobubble collapse on a silica surface in water: Billion-atom reactive molecular dynamics simulations, *Phys. Rev. Lett.*, **2013**, *111*, 184503.

[146] C. Xiao, D. M. Heyes, and J. G. Powles, The collapsing bubble in a liquid by molecular dynamics simulations, *Mol. Phys.*, **2002**, *100*, 3451–3468.

[147] R. Hołyst, M. Litniewski, and P. Garstecki, Large-scale molecular dynamics verification of the rayleigh-plesset approximation for collapse of nanobubbles, *Phys. Rev. E*, **2010**, *82*, 066309.

[148] F. Lugli, S. Höfinger, and F. Zerbetto, The collapse of nanobubbles in water, *J. Am. Chem. Soc.*, **2005**, *127*, 8020–8021.

[149] H. Fu, J. Comer, W. Cai, and C. Chipot, Sonoporation at small and large length scales: effect of cavitation bubble collapse on membranes, *J. Phys. Chem. Lett.*, **2015**, *6*, 413–418.

[150] V. H. Man, M. S. Li, P. Derreumaux, and P. H. Nguyen, Rayleigh-plesset equation of the bubble stable cavitation in water: A nonequilibrium all-atom molecular dynamics simulation study, *J. Chem. Phys.*, **2018**, *148*, 094505.

[151] V. Molinero and E. B. Moore, Water modeled as an intermediate element between carbon and silicon, *J. Phys. Chem. B*, **2009**, *113*, 4008–4016.

[152] J. L. F Abascal and C. Vega, A general purpose model for the condensed phases of water: Tip4p/2005, *J. Chem. Phys.*, **2005**, *123*, 234505.

[153] J. J. Potoff and J. I. Siepmann, Vapor–liquid equilibria of mixtures containing alkanes, carbon dioxide, and nitrogen, *AIChE journal*, **2001**, *47*, 1676–1682.

[154] R.W. Hockney and J.W. Eastwood, *Computer simulation using particles;* crc Press, 1988.

[155] J. L. Chen, B. Xue, K. Mahesh, and J. I. Siepmann, Molecular simulations probing the thermophysical properties of homogeneously stretched and bubbly water systems, *J. Chem. Eng. Data*, **2019**, *64*, 3755–3771.

[156] D. Brown, J. Clarke, M. Okuda, and T. Yamazaki, A domain decomposition parallelization strategy for molecular dynamics simulations on distributed memory machines, *Comput. Phys. Commun*, **1993**, *74*, 67–80.

[157] J. R. Dormand and P.J. Prince, A family of embedded runge-kutta formulae, *J. Comput. Appl. Math*, **1980**, *6*, 19–26.

[158] L. F. Shampine and M. W. Reichelt, The matlab ode suite, *SIAM J. Sci. Comput.*, **1997**, *18*, 1–22.

[159] R. Lustig, Direct molecular nvt simulation of the isobaric heat capacity, speed of sound and joule–thomson coefficient, *Mol. Simul.*, **2011**, *37*, 457–465.

[160] B. R. Munson, T. H. Okiishi, W. W. Huebsch, and A. P. Rothmayer, *Fluid mechanics;* Wiley Singapore, 2013.

[161] D. Obreschkow, M. Bruderer, and M. Farhat, Analytical approximations for the collapse of an empty spherical bubble, *Phys. Rev. E*, **2012**, *85*, 066303.

[162] E. Johnsen and T. Colonius, Numerical simulations of non-spherical bubble collapse, *J. Fluid Mech*, **2009**, *629*, 231–262.

[163] H. Okumura and N. Ito, Nonequilibrium molecular dynamics simulations of a bubble, *Phys. Rev. E*, **2003**, *67*, 045301.

[164] D. Schanz, B. Metten, T. Kurz, and W. Lauterborn, Molecular dynamics simulations of cavitation bubble collapse and sonoluminescence, *New J. Phys.*, **2012**, *14*, 113019.

[165] A. Bass, S. J. Ruuth, C. Camara, B. Merriman, and S. Putterman, Molecular dynamics of extreme mass segregation in a rapidly collapsing bubble, *Phys. Rev. Lett.*, **2008**, *101*, 234301.

[166] H. O. Pierson, *Handbook of chemical vapor deposition: principles, technology and applications;* William Andrew, 1999.

[167] K. D. Weeks, J. Tolle, M. G. Goodman, and S. Mehta; Doped semiconductor films and processing, **2015;** US Patent 9,099,423.

[168] C. H. Bell, Fumigation in the 21st century, *Crop Prot.*, **2000**, *19*, 563–569.

[169] L. D. Freedman, G. Doak, G. G. Long, T. Mahmood, and C. B Lindhal, Antimony compounds, *Kirk-Othmer Encyclopedia of Chemical Technology*, **2000**.

[170] W. A. Rickelton, Phosphine and its derivatives, *Kirk-Othmer Encyclopedia of Chemical Technology*, **2000**.

[171] D. Pakulska and S. Czerczak, Hazardous effects of arsine: a short review, *Int J Occup Med Environ Health*, **2006**, *19*, 36–44.

[172] P. Andrewes, K. T. Kitchin, and K. Wallace, Plasmid dna damage caused by stibine and trimethylstibine, *Toxicol. Appl. Pharmacol.*, **2004**, *194*, 41–48.

[173] S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, Q. Li, B. A. Shoemaker, P. A. Thiessen, B. Yu, et al., Pubchem 2019 update: improved access to chemical data, *Nucleic Acids Res.*, **2019**, *47*, D1102–D1109.

[174] Pubchem identifier: Cid 9359.

[175] C. C. Stephenson and W. F. Giauque, A test of the third law of thermodynamics by means of two crystalline forms of phosphine. the heat capacity, heat of vaporization and vapor pressure of phosphine. entropy of the gas, *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*, Eds. P.J. Linstrom and W.G. Mallard, National Institute of Standards and Technology, Gaithersburg MD, 20899 (retrieved December 1, 2020).

[176] D. R. Stull, Vapor pressure of pure substances. organic and inorganic compounds, *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*, Eds. P.J. Linstrom and W.G. Mallard, National Institute of Standards and Technology, Gaithersburg MD, 20899 (retrieved December 1, 2020).

[177] R. H. Sherman and W. F. Giauque, Arsine. vapor pressure, heat capacity, heats of transition, fusion and vaporization. the entropy from calorimetric and from molecular data1, *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*, Eds. P.J. Linstrom and W.G. Mallard, National Institute of Standards and Technology, Gaithersburg MD, 20899 (retrieved December 1, 2020).

[178] D. R. Stull, Vapor pressure of pure substances. organic and inorganic compounds, *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*, Eds. P.J. Linstrom and W.G. Mallard, National Institute of Standards and Technology, Gaithersburg MD, 20899 (retrieved December 1, 2020).

[179] L. Berka, T. Briggs, M. Millard, and W. Jolly, The preparation of stibine and the measurement of its vapour pressure, *J Inorg Nucl Chem*, **1960**, *14*, 190–194.

[180] W. Kohn and L. J. Sham, Self-consistent equations including exchange and correlation effects, *Phys. Rev.*, **1965**, *140*, A1133.

[181] M. J. McGrath, J. I. Siepmann, I. W. Kuo, C. J. Mundy, J. VandeVondele, J. Hutter, F. Mohamed, and M. Krack, Simulating fluid-phase equilibria of water from first principles, *J. Phys. Chem. A*, **2006**, *110*, 640–646.

[182] M. J. McGrath, J. I. Siepmann, I. W. Kuo, and C. J. Mundy, Vapor–liquid equilibria of water from first principles: comparison of density functionals and basis sets, *Mol. Phys.*, **2006**, *104*, 3619–3626.

[183] M. J. McGrath, I. W. Kuo, J. N. Ghogomu, C. J. Mundy, and J. I. Siepmann, Vapor–liquid coexistence curves for methanol and methane using dispersion-corrected density functional theory, *J. Phys. Chem. B*, **2011**, *115*, 11688–11692.

[184] P. Schienbein and D. Marx, Liquid–vapor phase diagram of RPBE-D3 water: electronic properties along the coexistence curve and in the supercritical phase, *J. Phys. Chem. B*, **2017**, *122*, 3318–3329.

[185] H. Goel, C. L. Butler, Z. W. Windom, and N. Rai, Vapor liquid equilibria of hydrofluorocarbons using dispersion-corrected and nonlocal density functionals, *J. Chem. Theory Comput.*, **2016**, *12*, 3295–3304.

[186] H. Goel, Z. W. Windom, C. L. Butler, and N. Rai, Phase equilibria and condensed phase properties of fluorinated alkanes via first principles simulations, *Chemistry Select*, **2017**, *2*.

[187] H. Goel, Z. W. Windom, A. A. Jackson, and N. Rai, Performance of density functionals for modeling vapor liquid equilibria of CO_2 and SO_2, *J. Comput. Chem*, **2018**, *39*, 397–406.

[188] H. Goel, S. Ling, B. N. Ellis, A. Taconi, B. Slater, and N. Rai, Predicting vapor liquid equilibria using density functional theory: A case study of argon, *J. Chem. Phys*, **2018**, *148*, 224501.

[189] E. O. Fetisov, M. S. Shah, J. R. Long, M. Tsapatsis, and J. I. Siepmann, First principles Monte Carlo simulations of unary and binary adsorption: CO_2, N_2, and H_2O in Mg-MOF-74, *Chem. Commun.*, **2018**, *54*, 10816–10819.

[190] H. Goel, Z. W. Windom, A. A. Jackson, and N. Rai, CO_2 sorption in triethyl (butyl) phosphonium 2-cyanopyrrolide ionic liquid via first principles simulations, *J. Molec. Liq.*, **2019**, *292*, 111323.

[191] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (dft-d) for the 94 elements h-pu, *J. Chem. Phys*, **2010**, *132*, 154104.

[192] A. Z. Panagiotopoulos, Direct determination of fluid phase equilibria by simulation in the Gibbs ensemble: a review, *Mol Simul*, **1992**, *9*, 1–23.

[193] T. D. Kühne, M. Iannuzzi, M. Del Ben, V. V. Rybkin, P. Seewald, F. Stein, T. Laino, R. Z. Khaliullin, O. Schütt, F. Schiffmann, et al., Cp2k: An electronic structure and molecular dynamics software package-quickstep: Efficient and accurate electronic structure calculations, *J. Chem. Phys.*, **2020**, *152*, 194103.

[194] M. J. McGrath, J. I. Siepmann, I. W. Kuo, C. J. Mundy, J. VandeVondele, M. Sprik, J. Hutter, F. Mohamed, M. Krack, and M. Parrinello, Toward a Monte Carlo program for simulating vapor–liquid phase equilibria from first principles, *Comput. Phys. Commun*, **2005**, *169*, 289–294.

[195] J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, and J. Hutter, Quickstep: Fast and accurate density functional calculations using a mixed gaussian and plane waves approach, *Comput. Phys. Commun*, **2005**, *167*, 103–128.

[196] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized gradient approximation made simple, *Phys. Rev. Lett.*, **1996**, *77*, 3865.

[197] A. D. Becke, Density-functional exchange-energy approximation with correct asymptotic behavior, *Phys. Rev. A*, **1988**, *38*, 3098.

[198] C. Lee, W. Yang, and R. G. Parr, Development of the colle-salvetti correlation-energy formula into a functional of the electron density, *Phys. Rev. B*, **1988**, *37*, 785.

[199] S. Goedecker, M. Teter, and J. Hutter, Separable dual-space gaussian pseudopotentials, *Phys. Rev. B*, **1996**, *54*, 1703.

[200] C. Hartwigsen, S. Gœdecker, and J. Hutter, Relativistic separable dual-space gaussian pseudopotentials from h to rn, *Phys. Rev. B*, **1998**, *58*(7), 3641.

[201] J. VandeVondele and J. Hutter, Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases, *J. Chem. Phys*, **2007**, *127*, 114105.

[202] C. J. Mundy, I-F. W. Kuo, J. I. Siepmann, M. J. McGrath, J. Vondevondele, M. Sprik, J. Hutter, M. Parrinello, F. Mohamed, M. Krack, et al., Liquid water from first principles: Validation of different sampling approaches, *J. Phys. Chem. B*, **2004**, *108*, 12990–12998.

[203] J. S. Rowlinson and B. Widom, *Molecular theory of capillarity;* Courier Corporation, 2013.

[204] J. S. Rowlinson and F. Swinton, *Liquids and liquid mixtures: Butterworths monographs in chemistry;* Butterworth-Heinemann, 2013.

[205] M. Dinpajooh, P. Bai, D. A. Allan, and J. I. Siepmann, Accurate and precise determination of critical properties from Gibbs ensemble Monte Carlo simulations, *J. Chem. Phys*, **2015**, *143*, 114113.

[206] A. Kud, C. Coquelet, and S. Maixner, A new semi-empirical model for saturated vapor density of pure compounds, *J. Chem. Eng. Data*, **2020**, *65*, 577–590.

[207] B. L. Eggimann, Y. Sun, R. F. DeJaco, R. Singh, M. Ahsan, T. R. Josephson, and J. I. Siepmann, Assessing the quality of molecular simulations for vapor–liquid equilibria: An analysis of the TraPPE database, *J. Chem. Eng. Data*, **2019**, *65*, 1330–1344.

[208] Pubchem identifier: Cid 24404.

[209] A. A. Durrant, T. G. Pearson, and P. L. Robsinson, *J. Chem. Soc.*, **1934**, pages 730–735, Reaxys Database (retrieved December 1, 2020).

[210] Pubchem identifier: Cid 23969.

[211] A. L. G. Rees and K. Stewart, The density of liquid arsine, *Trans Faraday Soc*, **1949**, *45*, 1028–1032, Reaxys Database (retrieved December 1, 2020).

[212] A. Frank and K. Clusius, Vapor pressure of pure substances. organic and inorganic compounds, *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*, Eds. P.J. Linstrom and W.G. Mallard, National Institute of Standards and Technology, Gaithersburg MD, 20899 (retrieved December 1, 2020).

[213] A. A. Durrant, T. G. Pearson, and P. L. Robsinson, *J. Chem. Soc.*, **1934**, page 732, Reaxys Database (retrieved December 1, 2020).

[214] J. I. Siepmann, S. Karaborni, and B. Smit, Simulating the critical behaviour of complex fluids, *Nature*, **1993**, *365*, 330–332.

[215] P. G. Sennikov, Weak hydrogen-bonding by second-row (PH_3, H_2S) and third-row (AsH_3, H_2Se) hydrides, *J. Phys. Chem*, **1994**, *98*, 4973–4981.

[216] P. Helminger, E. L. Beeson Jr, and W. Gordy, Microwave spectra and molecular constants of arsine and stibine, *Phys. Rev. A*, **1971**, *3*, 122.

[217] M. Diraison, G. J. Martyna, and M. E. Tuckerman, Simulation studies of liquid ammonia by classical ab initio, classical, and path-integral molecular dynamics, *J. Chem. Phys*, **1999**, *111*, 1096–1103.

[218] M. A. Ricci, M. Nardone, F. P. Ricci, C. Andreani, and A. K. Soper, Microscopic structure of low temperature liquid ammonia: A neutron diffraction experiment, *J. Chem. Phys.*, **1995**, *102*, 7650–7655.

[219] A. D. Boese, A. Chandra, J. M. L. Martin, and D. Marx, From ab initio quantum chemistry to molecular dynamics: The delicate case of hydrogen bonding in ammonia, *J. Chem. Phys.*, **2003**, *119*, 5965–5980.

[220] A. P. Fröba and A. Leipertz, Accurate determination of liquid viscosity and surface tension using surface light scattering (sls): Toluene under saturation conditions between 260 and 380 k, *Int. J. Thermophys*, **2003**, *24*, 895–921.

[221] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, Jr. J. A. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, , and D. J. Fox; Gaussian 09, revision d.01; Gaussian, Inc., Wallingford CT, 2013.

[222] C. P. Kelly, C. J. Cramer, and D. G. Truhlar, SM6: a density functional theory continuum solvation model for calculating solvation free energies of neutral, ions, and solute-water clusters, *J. Chem. Theory Comput.*, **2005**, *1*, 1133–1152.

[223] R. M. Olson, A. V. Marenich, A. C. Chamberlin, C. P. Kelly, J. D. Thompson, J. D. Xidos, J. Li, G. D. Hawkins, P. D. Winget, T. Zhu, D. Rinaldi, D. A. Liotard, C. J. Cramer, D. G. Truhlar, and M. J. Frisch; MN-GSM version 2012; University of Minnesota, Minneapolis, MN, 55455-0431, **2012**.

[224] R. Span, E. W. Lemmon, R. T. Jacobsen, W. Wagner, and A. Yokozeki, A reference equation of state for the thermodynamic properties of nitrogen for temperatures from 63.151 to 1000 K and pressures to 2200 MPa, *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*, Eds. P.J. Linstrom and W.G. Mallard, National Institute of Standards and Technology, Gaithersburg MD, 20899 (retrieved December 1, 2020).

[225] A. Chapoy, A. H. Mohammadi, B. Tohidi, and D. Richon, Gas solubility measurement and modeling for the nitrogen+ water system from 274.18 K to 363.02 K, *J. Chem. Eng. Data*, **2004**, *49*, 1110–1115.

[226] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, et al., Gaussian 16, *Gaussian Inc. Wallingford CT*, **2016**.

[227] A. V. Marenich, S. V. Jerome, C. J. Cramer, and D. G. Truhlar, 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.

Appendix A. Supporting information for chapter 2

Table SA1. Surface tension of toluene obtained from simulations. Simulations for calculating the surface tension of toluene were performed in NVT ensemble with 260 toluene molecules and a box dimension of 32 Å$\times$ 32 Å$\times$ 260 Å. Simulations in NVT ensemble were first performed to obtain a stable liquid configuration of toluene. Then the box was elongated in the z dimension. For all state points, we performed 8 independent simulations, each consisted of at least 100,000 cycles. Error bars for simulation denote the 95% confidence interval.

T [K]	Simulated surface tension [dyn/cm]
293.15	31.79 ± 0.83
313.15	29.55 ± 0.58
333.15	26.08 ± 0.86
353.15	24.83 ± 0.54
373.15	22.52 ± 0.86

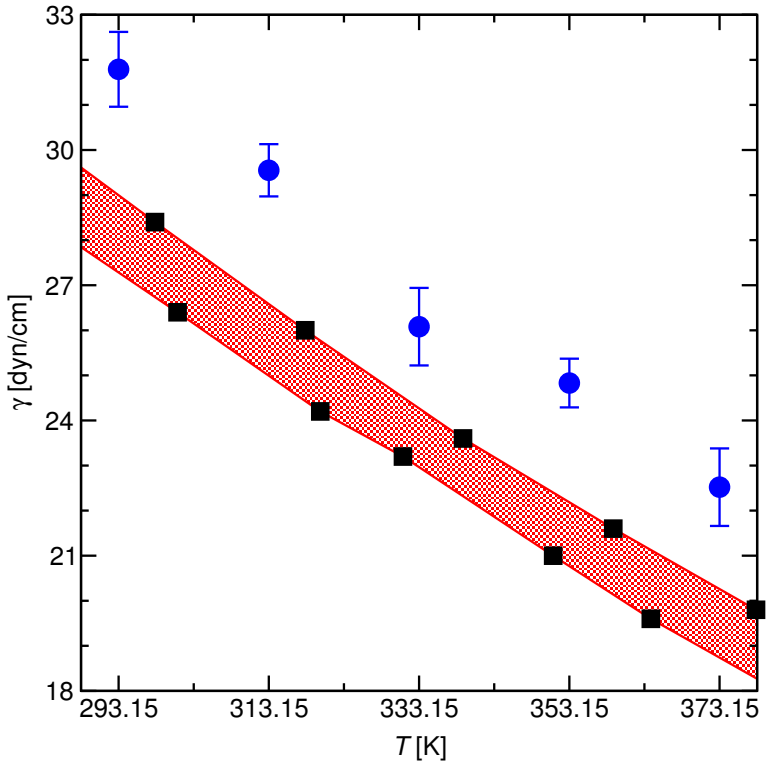
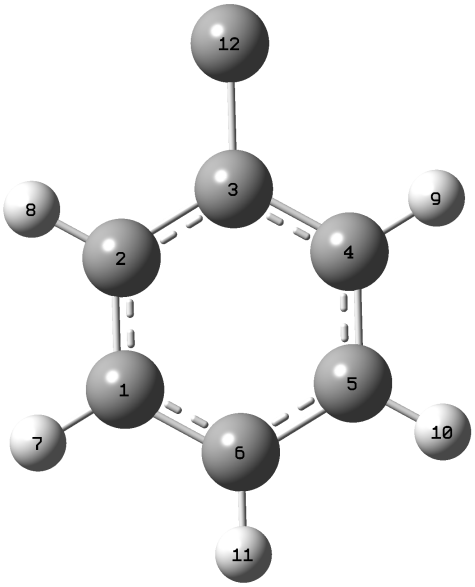


Figure SA1. Surface tension of toluene. The blue circles are simulation results. Red lines and shaded areas indicate specific values and range of experimental data [220].

Table SA2. Lennard-Jones parameters and partial charges for the molecular models. The partial charges for toluene are determined as follows: First, the toluene structure is optimized using density functional theory at the B3PW91/6-31++G(d,p) level of theory using the Gaussian 09 software [221]. The CM4 charges [222] for the optimized gas-phase structure are then determined at B3P86/6-31+G(d,p) level of theory with 1-octanol as the universal continuum solvent using the Minnesota Gaussian solvation model (MN-GSM) implemented in Gaussian 09 [223]. The united-atom methyl group is assigned the sum of the partial charges of its carbon and three hydrogen atoms. The site numbers are given in the molecular structure.



Molecule	Force Field	Site	ε/k_B [K]	σ [Å] [Å]	q $\lvert e\rvert$
Water	TIP4P/2005	O	93.2	3.1589	—
		H	—	—	+0.5564
		M	–	—	−1.1128
n-Alkane	TraPPE–UA	CH_3	98	3.75	—
		CH_2	46	3.95	—
Benzene	TraPPE–EH	C(aro)	30.7	3.600	−0.095
		H(aro)	25.5	2.360	+0.095
Toluene	TraPPE–EH/UA	C1, C5	30.7	3.600	−0.093
		C2, C4	30.7	3.600	−0.106
		C3	30.7	3.600	−0.042
		C6	30.7	3.600	−0.101
		H7, H10	25.5	2.360	+0.091
		H8, H9	25.5	2.360	+0.083
		H11	25.5	2.360	+0.090
		CH_3	98	3.75	+0.103

Table SA3. Bonded parameters for the molecular models

Molecule	Force Field	Bond Type	l_0 [Å]
Water	TIP4P/2005	O–H	0.9572
		O–M	0.1546
Alkane	TraPPE–UA	CH_x–CH_y	1.54
Arene	TraPPE–EH	C(aro)–H(aro)	1.08
		C(aro)–C(aro)	1.392
		C(aro)–CH_x	1.54

Molecule	Force Field	Angle Type	k/k_B [K/rad^2]	θ_0 [deg]
Water	TIP4P/2005	H–O–H	Rigid	104.52
Alkane	TraPPE–UA	CH_x–CH_2–CH_y	62500	114
		CH_x–CH–CH_y	62500	112
Arene	TraPPE–EH	C(aro)–C(aro)–C(aro)	Rigid	120
		C(aro)–C(aro)–H(aro)	Rigid	120
		C(aro)–C(aro)–CH_x	Rigid	120

Molecule	Model	Dihedral Type	c_0/k_B [K]	c_1/k_B [K]	c_2/k_B [K]	c_3/k_B [K]
Alkane	TraPPE–UA	CH_x–CH_2–CH_2–CH_y	0.00	355.03	−68.19	791.32

Table SA4. Simulations are carried out using the Monte Carlo for Complex Chemical Systems-Minnesota (MCCCS–MN, version 16.1) code developed in house [84]. Distribution of Monte Carlo moves for simulations of the binary mixtures in the $N_1 N_2 p_N AT$ ensemble. The CDDC-CBMC inner cutoff is set to 5.0 Å and the number of trial sites during the regrowth is set to 8.

Mixture	Move Type	Molecule	Probability
Water/Decane	Translations or Rotations	Decane	0.056
	Translations or Rotations	Water	0.344
	CBMC regrowth	Decane	0.199
	Volume		0.001
Water/Dodecane	Translations or Rotations	Dodecane	0.030
	Translations or Rotations	Water	0.370
	CBMC regrowth	Dodecane	0.199
	Volume		0.001
Water/Benzene	Translations or Rotations	Benzene	0.081
	Translations or Rotations	Water	0.4185
	Volume		0.001
Water/Toluene	Translations or Rotations	Toluene	0.074
	Translations or Rotations	Water	0.4255
	Volume		0.001

Table SA5. Simulations are carried out using the Monte Carlo for Complex Chemical Systems-Minnesota (MCCCS–MN, version 16.1) code developed in house [84]. Distribution of Monte Carlo moves for simulations of the terary mixtures in the $N_{\mathrm{W}}N_{\mathrm{D}}N_{\mathrm{T}}(p_{\mathrm{N}}A)V_{\mathrm{D}}V_{\mathrm{T}}T$ ensemble. The CDDC-CBMC inner cutoff is set to 5.0 Å, the numbers of trial sites during the insertion of the first bead and during regrowth of dodecane are set to 32 and 8, and the number of orientations explored during the transfer of toluene is set to 16.

Move Type	Molecule or Box	Probability
Translations or Rotations	Water	0.295
	Toluene	0.027
	Dodecane	0.021
	3-Methyl-pentane	0.0035
	Octane	0.0035
Regrowths	Dodecane	0.196
	3-Methyl-pentane	0.002
	Octane	0.002
Particle transfer	Toluene/IFT	0.015
	Dodecane/IFT	0.035
Identity switch	Toluene/3-Methyl-pentane	0.016
	3-Methyl-pentane/Octane	0.016
	Octane/Dodecane	0.014
Volume changes	IFT	0.001
	Toluene	0.001
	Dodecane	0.001

Table SA6. Liquid-phase mole fraction of toluene and ideal-gas number densities of toluene and *n*-dodecane from binary simulations at $p = 1.83$ MPa. Errors denote the 95% confidence intervals.

T [K]	x_T	$\rho_\mathrm{T}/10^{-2}$ [1/nm^3]	$\rho_\mathrm{D}/10^{-4}$ [1/nm^3]
383.15	0.6209 ± 0.0020	1.094 ± 0.047	3.81 ± 0.24
	0.6423 ± 0.0014	1.125 ± 0.026	3.53 ± 0.16
	0.6601 ± 0.0019	1.158 ± 0.048	3.23 ± 0.17
	0.6655 ± 0.0014	1.170 ± 0.050	3.19 ± 0.19
403.15	0.6157 ± 0.0010	1.711 ± 0.031	7.53 ± 0.27
	0.6392 ± 0.0007	1.812 ± 0.017	7.19 ± 0.15
	0.6482 ± 0.0009	1.808 ± 0.045	6.76 ± 0.28
	0.6652 ± 0.0010	1.863 ± 0.051	6.33 ± 0.25
423.15	0.6170 ± 0.0010	2.609 ± 0.050	14.05 ± 0.47
	0.6414 ± 0.0010	2.756 ± 0.057	12.98 ± 0.47
	0.6654 ± 0.0006	2.853 ± 0.043	11.94 ± 0.25
443.15	0.62395 ± 0.00041	3.871 ± 0.038	24.53 ± 0.31
	0.64464 ± 0.00045	3.999 ± 0.059	22.43 ± 0.36
	0.66816 ± 0.00051	4.151 ± 0.016	20.49 ± 0.29

Table SA7. Ideal gas pressures used for the reservoirs of the simulations in the $N_{\mathrm{W}}N_{\mathrm{D}}N_{\mathrm{T}}(p_{\mathrm{N}}A)V_{\mathrm{D}}V_{\mathrm{T}}T$ ensemble.

	Toluene	Dodecane
T [K]	p [kPa]	p [kPa]
383.15	53.7	1.61
403.15	101	3.77
423.15	162	7.39
443.15	246	13.6

Table SA8. Experimental and calculated water/n-decane interfacial tension data used for the k_{ij} parameterization of our IFPSC entry. Error bars for simulation denote the 95% confidence interval.

T	p	Experiment		Simulations	
[K]	[bar]	Ref.	γ [dyn/cm]	k_{ij}	γ [dyn/cm]
323	1	[6]	49.78 ± 0.04	1.00	48.3 ± 2.5
				1.05	43.6 ± 3.2
				1.10	42.4 ± 4.5
423	400	[4]	36.09 ± 0.33	1.00	40.1 ± 2.4
				1.05	36.3 ± 2.2
				1.10	34.4 ± 3.4
473	400	[4]	30.59 ± 0.37	1.00	32.7 ± 1.7
				1.05	31.3 ± 1.9
				1.10	28.6 ± 1.9

Table SA9. Experimental and calculated water/n-decane interfacial tension data used for the improved k_{ij} parameterization. Error bars for simulation denote the 95% confidence interval.

T	p	Experiment		Simulations	
[K]	[bar]	Ref.	γ [dyn/cm]	k_{ij}	γ [dyn/cm]
298	1	[6]	51.98 ± 0.04	0.90	55.4 ± 2.0
		[4]	51.1 ± 0.50	0.95	52.4 ± 1.5
		[7]	50.5	1.00	50.4 ± 1.5
		[9]	51.96 ± 0.28	1.05	45.4 ± 1.9
323	1	[6]	49.78 ± 0.04	0.90	53.9 ± 1.6
		[9]	47.78 ± 0.16	0.95	50.8 ± 1.4
		[8]	45.6	1.00	48.9 ± 1.7
				1.05	45.5 ± 2.1
423	400	[4]	36.09 ± 0.33	1.00	39.5 ± 1.3
				1.05	36.3 ± 1.9
				1.10	33.8 ± 1.9
473	400	[4]	30.59 ± 0.37	1.00	33.1 ± 1.0
				1.05	31.3 ± 1.5
				1.10	29.1 ± 1.4

Table SA10. Experimental and calculated water/benzene interfacial tension data used for the k_{ij} parameterization of our IFPSC entry. Error bars for simulation denote the 95% confidence interval.

T	p	Experiment		Simulations	
[K]	[bar]	Ref.	γ [dyn/cm]	k_{ij}	γ [dyn/cm]
373	200	[7]	25.0	1.00	40.0 ± 3.2
				1.20	28.4 ± 2.5
				1.25	23.6 ± 2.5
				1.30	20.4 ± 3.4
				1.35	14.8 ± 2.3
449	200	[7]	16.1	1.00	33.1 ± 1.8
				1.20	20.5 ± 1.7
				1.25	16.8 ± 1.5
				1.30	15.2 ± 1.7
				1.35	11.8 ± 1.4

Table SA11. Experimental and calculated water/benzene interfacial tension data used for the improved k_{ij} parameterization. Error bars for simulation denote the 95% confidence interval.

T	p	Experiment		Simulations	
[K]	[bar]	Ref.	γ [dyn/cm]	k_{ij}	γ [dyn/cm]
298	1	[7]	34.6	1.00	48.0 ± 1.5
		[8]	32.6	1.10	38.6 ± 2.0
				1.15	34.0 ± 1.9
				1.20	29.0 ± 2.1
323	1	[7]	31.1	1.00	45.0 ± 1.6
				1.15	32.3 ± 1.6
				1.20	28.4 ± 1.9
				1.25	24.9 ± 1.9
373	200	[7]	25.0	1.00	41.6 ± 1.3
		[8]	26.2	1.20	26.5 ± 1.9
				1.25	23.1 ± 1.7
				1.30	18.2 ± 1.6
449	200	[7]	16.1	1.00	32.7 ± 1.3
				1.20	21.0 ± 1.0
				1.25	16.2 ± 1.5
				1.30	13.8 ± 1.6

Investigation of localized k_{ij}. A localized k_{ij} parameter between oxygen atom in water and aromatic carbon atoms was constructed in a way that the normal constant k_{ij} was replaced with a distance-dependent value. A hyperbolic tangent function was introduced to represent the distance dependence of k_{ij}, where the value goes from $k_{ij,\max}$ that is larger than 1.0 to $k_{ij,\min}$ of 1.0 as two LJ interaction sites goes further away from each other. A sample plot of the hyperbolic tangent function was given in Figure SA2. Then we went through the same fitting procedure for $k_{ij,\max}$ as what we did for fitting the constant k_{ij}, except for that we only introduced the localized k_{ij} between the oxygen atom in water and the six aromatic carbon atoms to mimic localized induced polarization effect, while using normal LB combing rule for the interaction between oxygen and aromatic hydrogen atoms. The final $k_{ij,\max}$ for O–C (aromatic) was determined to be 1.67, which was used later in simulations for water/toluene and water/(50 wt% *n*-dodecane + 50 wt% toluene) mixture. For O–CH_x (aliphatic) and O–CH_3 in toluene, we used normal LB combing rule, or a k_{ij} value of 1.0. We believed that compared to the previous improved k_{ij} value of 1.01 for O–CH_x (aliphatic), a difference of 0.01 in k_{ij} would have negligible effects on the IFT for water/*n*-dodecane. Thus the IFT values for water/*n*-dodecane under the I2A protocol could be directly compared to without any additional simulations. The predicted IFT values for water/toluene and water/(50 wt% *n*-dodecane + 50 wt% toluene) mixture utilizing a localized k_{ij} are given in Table SA12 and Figure SA3. The simulation protocol corresponding to this localized k_{ij} is denoted as I3A.

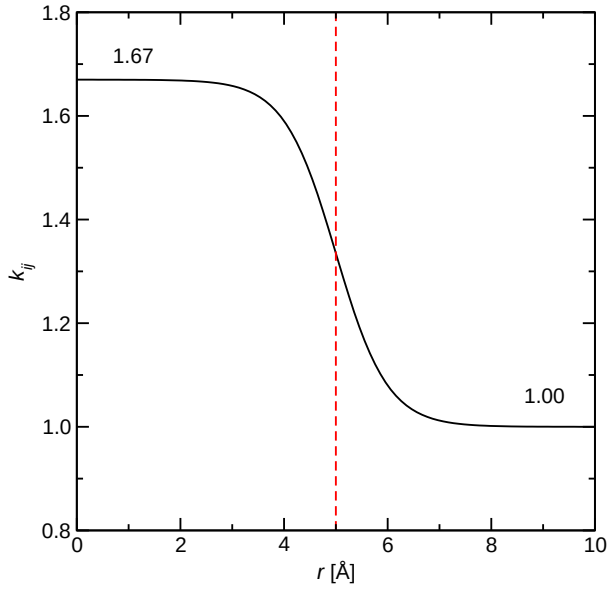


Figure SA2. A sample plot of how the k_{ij} varies with site-site distance. The expression of the function is $k_{ij} = -a * \tanh(\frac{r-b}{c}) + d$ where a, b, c, and d are adjustable parameters. a and d determine the starting and ending values of k_{ij}, respectively; while b and c determine the inflection point and the steepness of the curve, respectively. In this sample plot and also in our simulations for water/toluene and water/(50 wt% *n*-dodecane + 50 wt% toluene) mixture, the a, b, c, and d take values of 0.335, 5.0, 1.0, and 1.335, respectively. Among them, a and d are obtained from $k_{ij,\max}$ fitting procedure, while b is an estimation of the separation distance between water and benzene at the minimum point in a one dimensional potential energy surface obtained from quantum calculations. c is set to 1.0 for simplicity.

Table SA12. Calculated IFT for water/toluene and water/(50 wt% n-dodecane + 50 wt% toluene) mixture utilizing localized k_{ij} parameter. Error bars for simulation denote the 95% confidence interval.

Mixture	Protocol	T [K]			
		383.15	403.15	423.15	443.15
W/D	I2A or I3A	38.7 ± 1.8	35.8 ± 1.4	32.5 ± 1.1	30.7 ± 1.5
W/T	I3A	29.5 ± 1.5	25.5 ± 1.2	24.44 ± 0.95	20.7 ± 1.1
W/DT	I3A	32.3 ± 1.1	31.0 ± 1.1	28.8 ± 1.4	25.6 ± 1.1

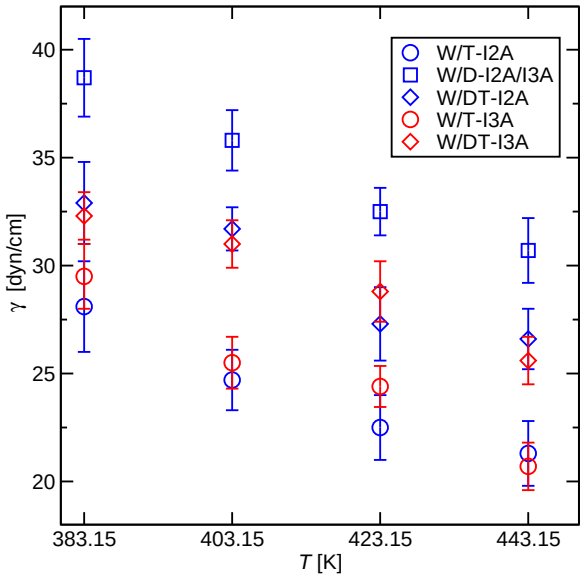


Figure SA3. Predicted IFT values utilizing I2A protocol (improved global k_{ij} values) and I3A protocol (localized k_{ij} values)

Appendix B. Supporting information for chapter 3

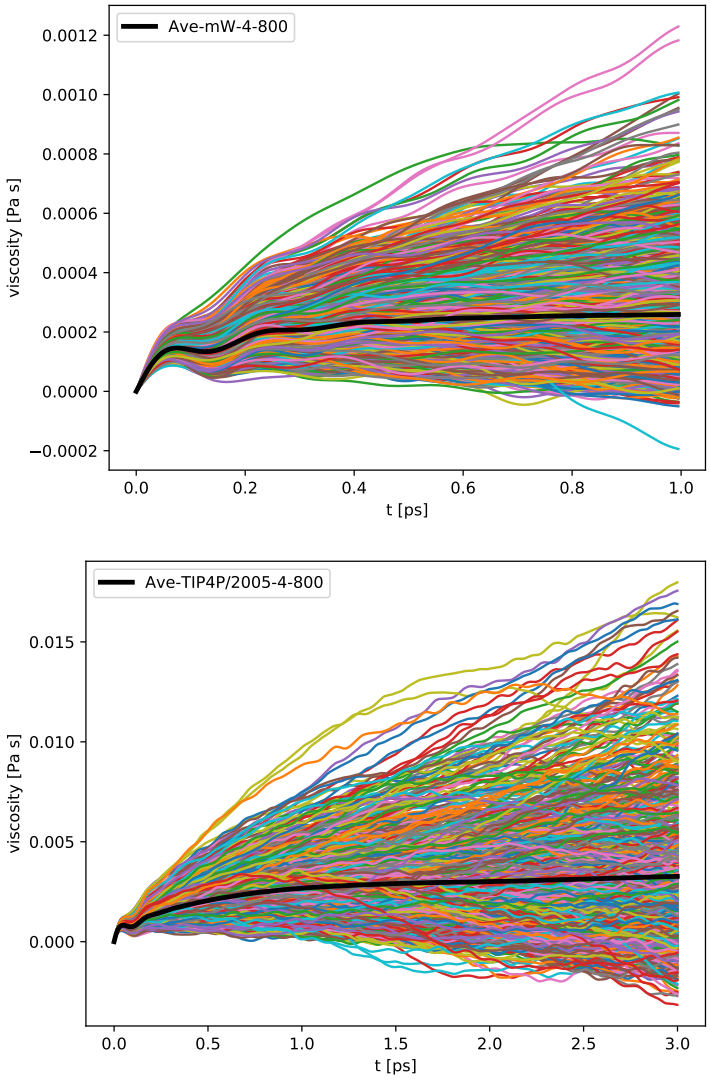


Figure SB1. Time evolution of the calculated viscosity from time integral of off-diagonal pressure tensor auto-correlation function (PCF) for the mW and TIP4P/2005 models with $L = 4$ nm and $\rho = 800$ kg/m^3. Due to the noise in molecular simulations, the time integral of PCF for individual simulation runs diverge, while their average still shows good convergence. The choice of the cutoff of the integral should not be too large (because the individual time integral of PCF will diverge wildly) nor too short (because the average will not converge). Based on these criteria, the time cutoffs were chosen to be 1.0 ps for the mW model and and 3.0 ps for the TIP4P/2005 model.

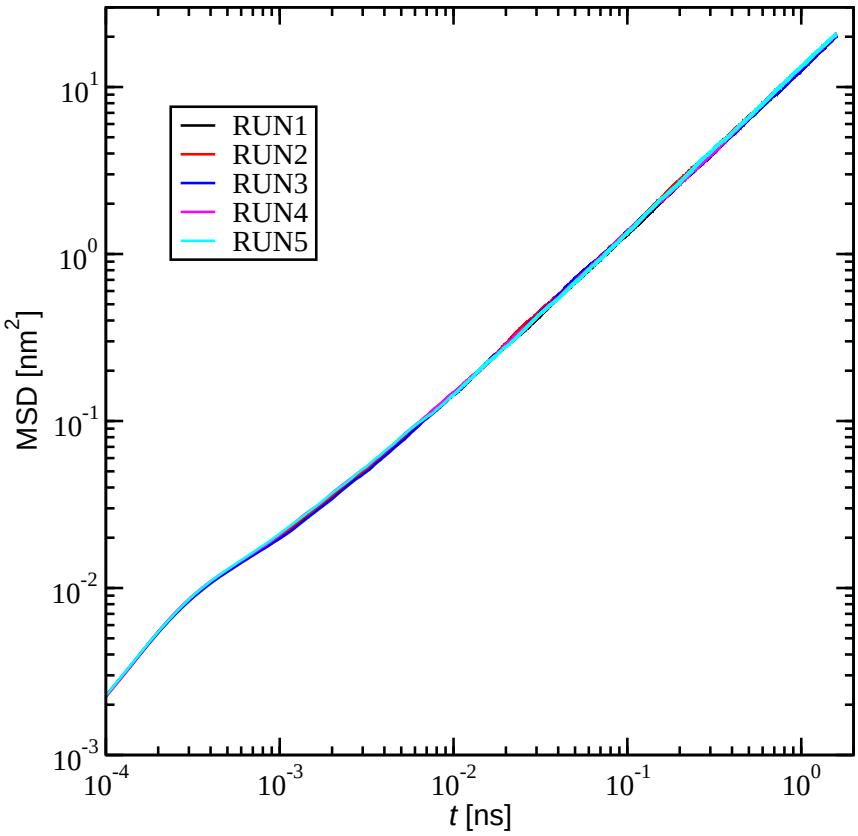


Figure SB2. Mean-square displacement (MSD) as a function of time for TIP4P/2005-8 system with density $\rho = 800$ kg/m^3 from 5 independent simulations.

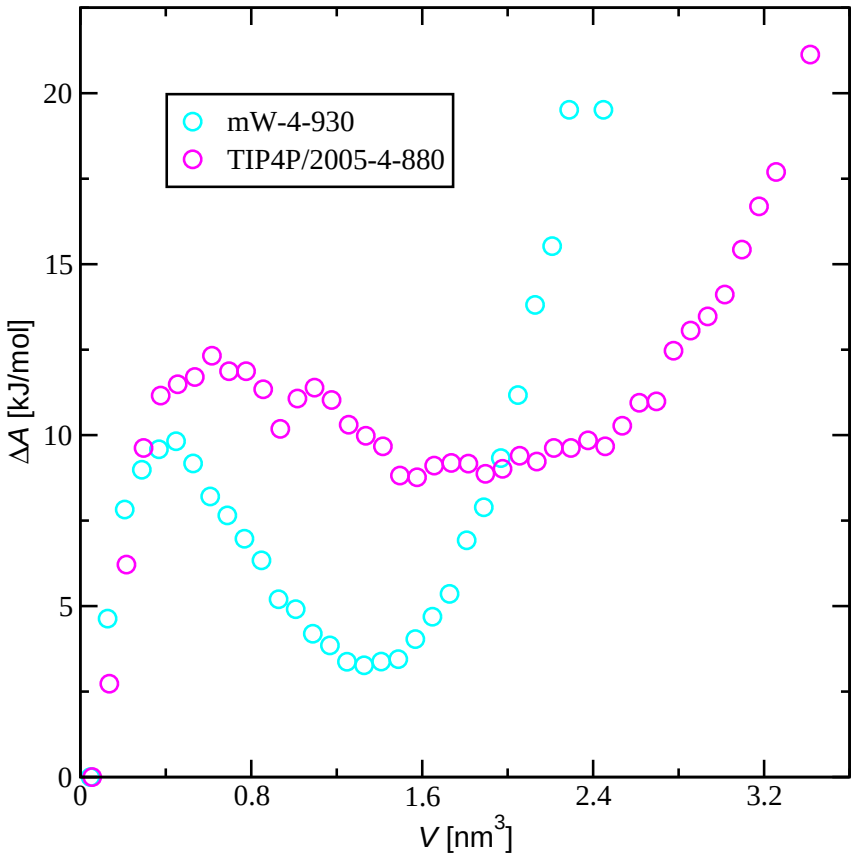


Figure SB3. Change in Helmholtz free energy for mW-4 and TIP4P/2005-4 systems at 930 kg/m^3 and 880 kg/m^3, respectively, as function of the volume of the largest void or bubble. The volume with the largest probability is assigned a Helmholtz free energy of zero.

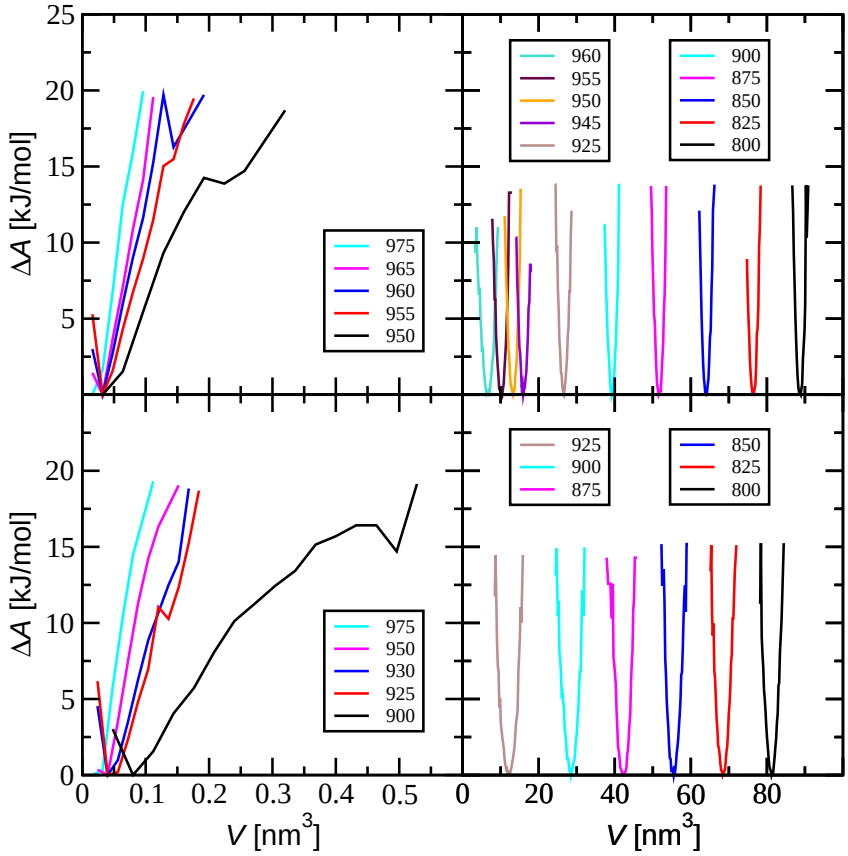


Figure SB4. Change in Helmholtz free energy for mW-8 (top) and TIP4P/2005-8 (bottom) systems as function of the volume of the largest void for homogeneously stretched phases (left) or of the bubble for the bubbly phases (right). The volume with the largest probability is assigned a Helmholtz free energy of zero. From the free energy profiles for mW-8 at 950 kg/m^3 and for TIP4P/2005-8 at 900 kg/m^3, we cautiously estimate that the nucleation free energy barriers are higher than 30 and 20 kJ/mol, respectively.

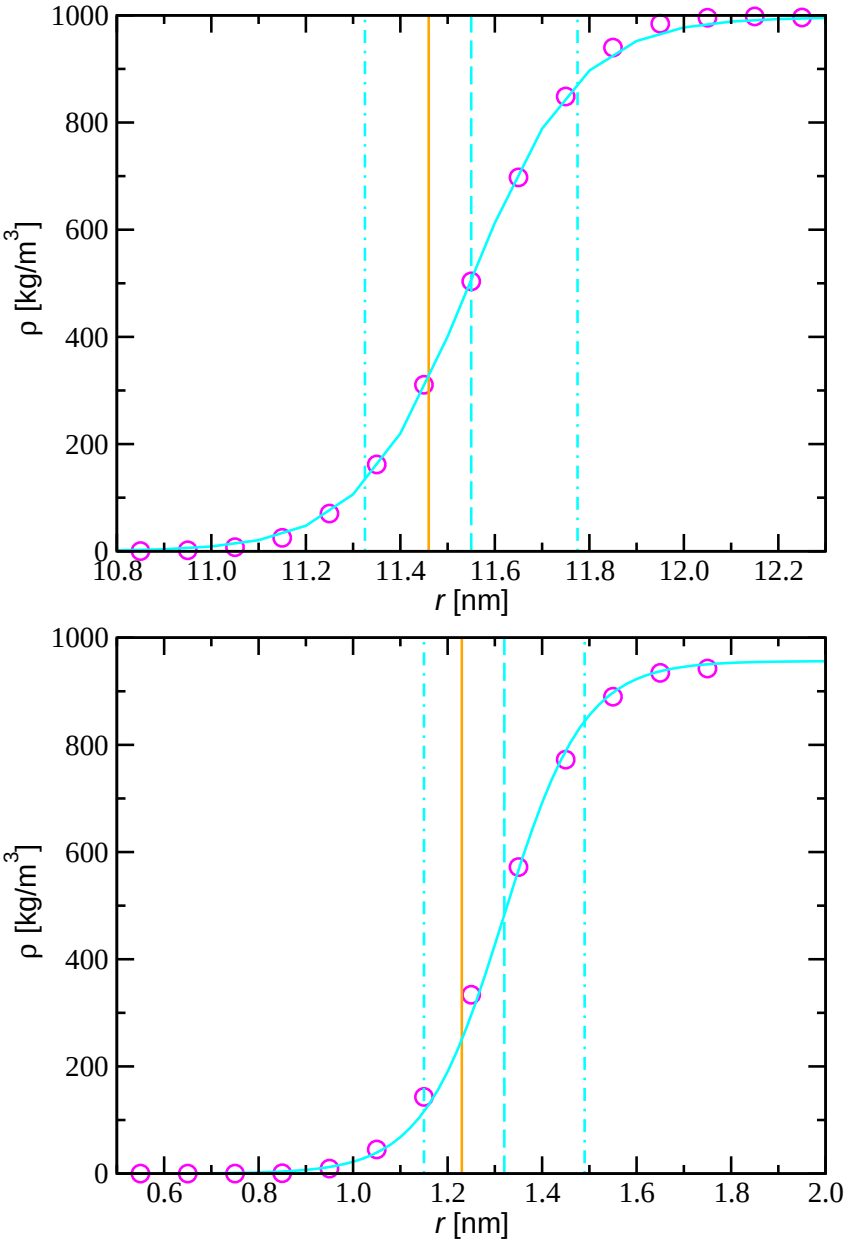


Figure SB5. Shell density profile for the mW-32 (top) and the TIP4P/2005-4 (bottom) systems with density $\rho = 800$ kg/m^3. The solid orange line denotes the bubble radius calculated from the bubble volume via the cell method (r^*), the dashed cyan line denotes the bubble radius obtained from fitting to hyperbolic tangent function (r_0), and the dash-dotted cyan lines denote the interfacial width ($r_0 \pm d/2$).

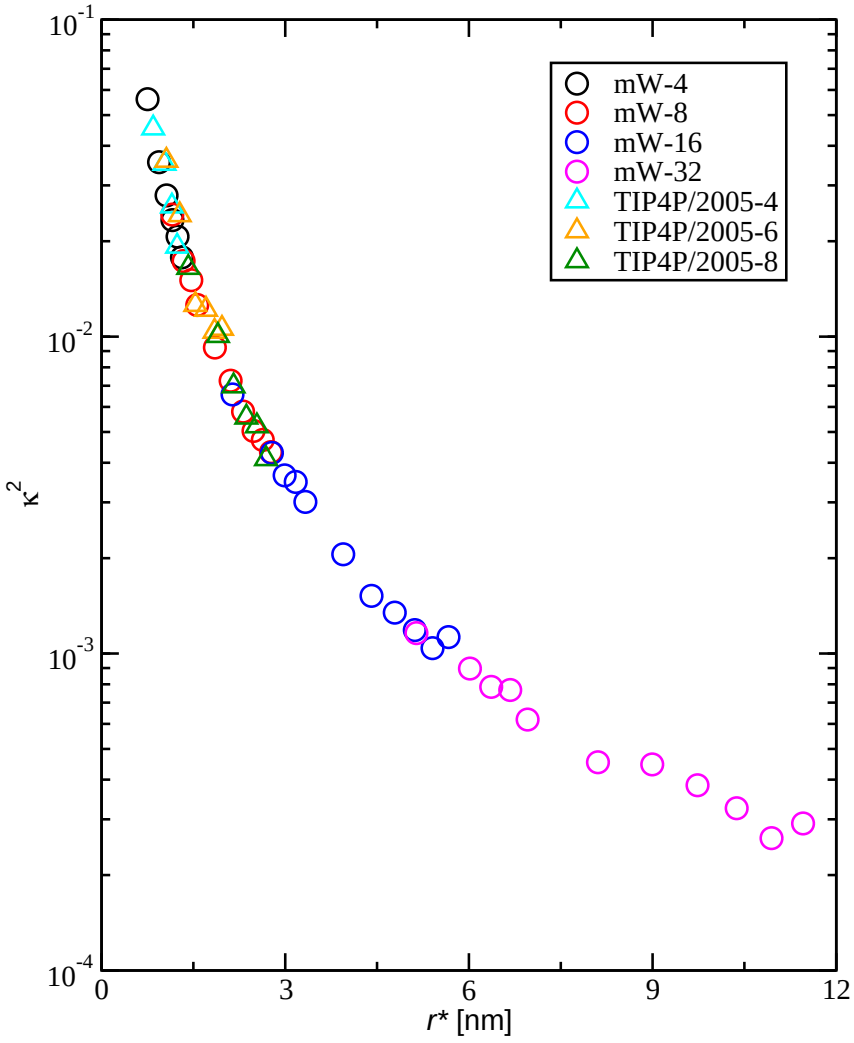


Figure SB6. Shape anisotropy factor as function of bubble radius.

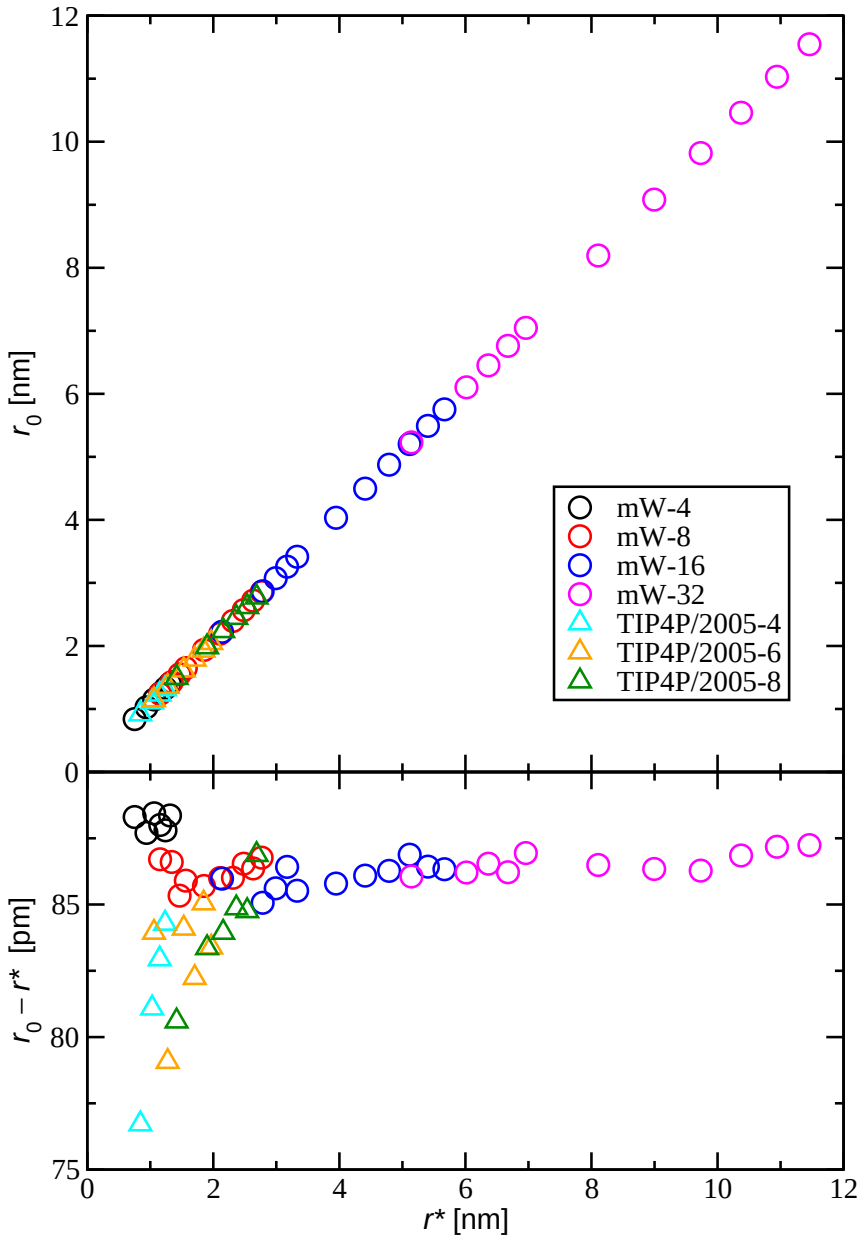


Figure SB7. Radius obtained from the shell density profile (r_0) versus radius obtained from bubble volume (r^*). The lower part shows the difference between r_0 and r^* as function of r^*.

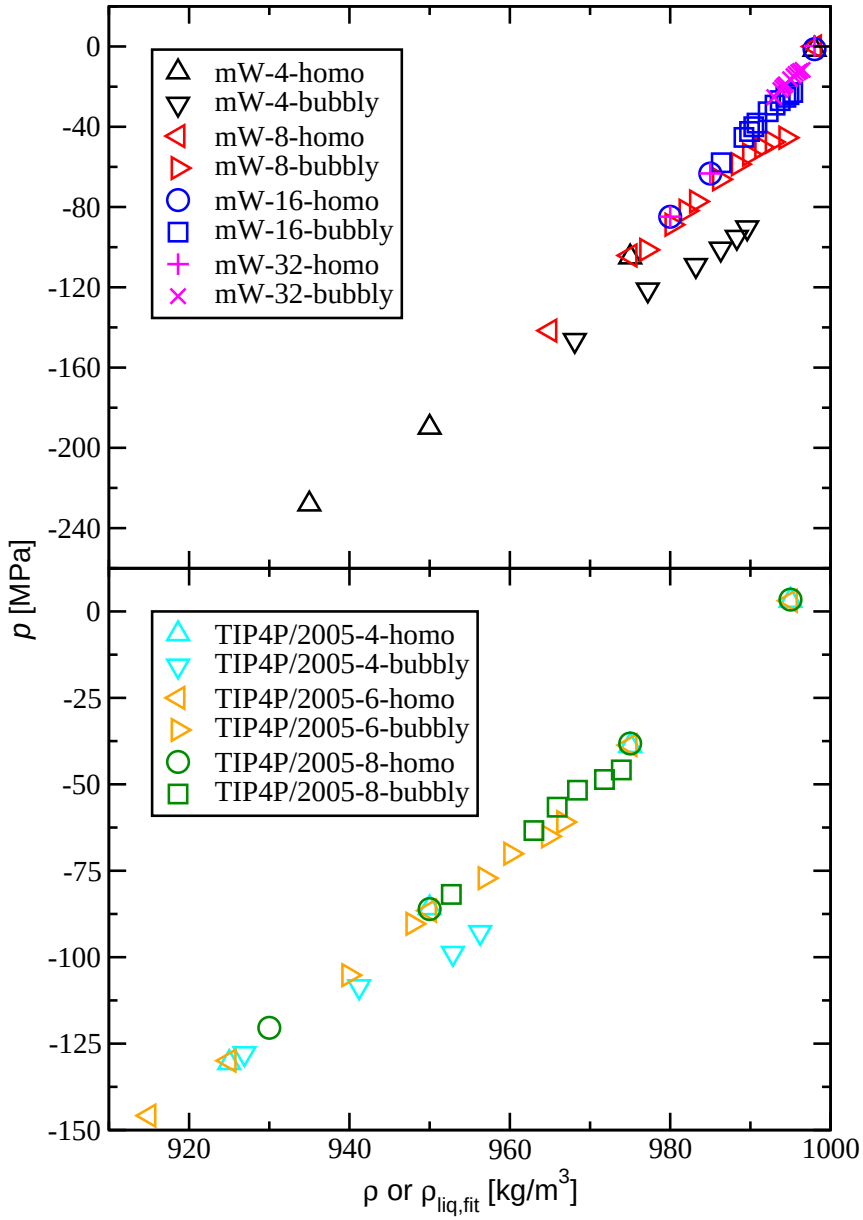


Figure SB8. Pressure for the mW (top) and TIP4P/2005 (bottom) models as function of ρ and $\rho_{\mathrm{liq,fit}}$ for homogeneously stretched phases and the bubbly systems, respectively.

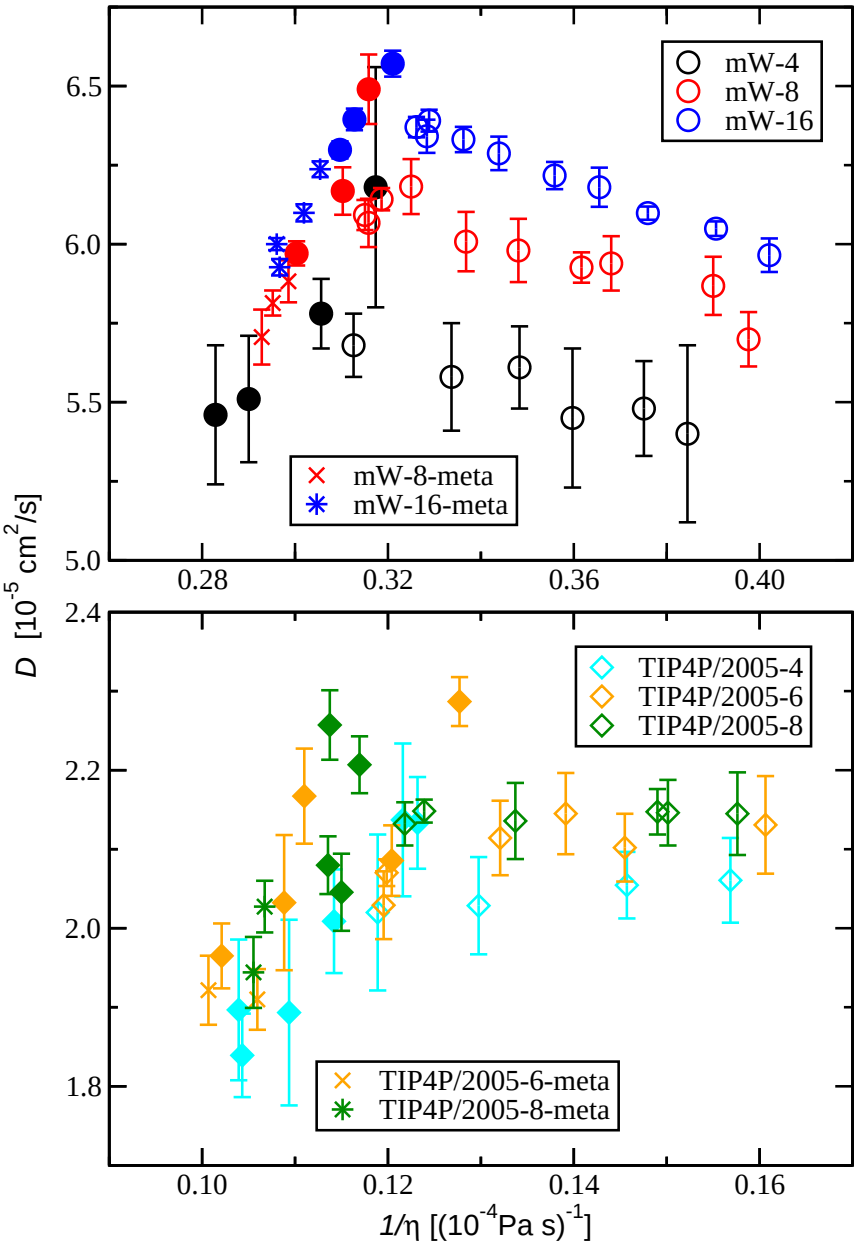


Figure SB9. Self-diffusion coefficient for the mW (top) and TIP4P/2005 (bottom) models as function of inverse shear viscosity. Open and filled symbols represent bubbly and homogeneously stretched water systems, respectively. The plusses, crosses, and stars show data for metastable homogeneous systems.

Table SB1. Number of molecules for all water systems investigated in this work at $T = 298$ K. The densities in this table are rounded to three significant figures but, due to the finite number of molecules, the actual densities differ slightly, and the precise density values are given in Tables SB2–SB8.

Name	ρ [kg/m^3]							
	800	825	850	875	885	890	895	900
mW-4	1711	1764	1818	1871				1925
mW-8	13688	14115	14543	14971				15398
mW-16	109469	112890	116311	119732				123153
mW-32	875753	903120	930488	957855				985222
TIP4P/2005-4	1711	1764	1818	1871	1892			1925
TIP4P/2005-6	5773	5953	6134	6314		6422	6458	6494
TIP4P/2005-8	13688	14115	14543	14971				15398

Name	ρ [kg/m^3]							
	910	915	925	930	935	945	950	955
mW-4			1978		2000		2032	
mW-8			15826			16168	16254	16340
mW-16			126574				129994	130678
mW-32			1012590				1039957	1045431
TIP4P/2005-4			1978				2032	
TIP4P/2005-6	6567	6603	6675				6855	
TIP4P/2005-8			15826	15911			16254	

Name	ρ [kg/m^3]						
	960	965	975	980	985	995	998
mW-4			2085				2134
mW-8	16426	16511	16682				17075
mW-16	131362	132047	133415	134099	134783		136563
mW-32	1050904	1056377	1067325	1072798	1078272		1092850
TIP4P/2005-4			2085			2128	
TIP4P/2005-6			7036			7180	
TIP4P/2005-8			16682			17023	

Table SB2. Numerical data for system properties obtained for the **mW-4 system** at $T = 298$ K: Pressure (P), potential energy (U), residual isochoric heat capacity ($C_{V,\mathrm{res}}$), viscosity (η), and diffusion coefficient (D). Density values followed by (m) denote metastable homogeneous systems. All uncertainties (u) represent the standard error of the mean at 95% confidence interval from block analysis.

ρ	P	u	U	u	$C_{V,\mathrm{res}}$	u	η	u	D	u
kg/m^3	MPa		kJ/mol		J/(mol·K)		10^{-4} Pa·s		10^{-5} cm^2/s	
800.25	−90.30	0.22	−41.31107	0.00091	22.60	0.34	2.601	0.062	5.40	0.28
825.04	−95.02	0.20	−41.39805	0.00093	22.66	0.31	2.666	0.033	5.48	0.15
850.30	−100.86	0.21	−41.48226	0.00078	22.79	0.22	2.780	0.035	5.45	0.22
875.09	−109.06	0.34	−41.56422	0.00079	22.65	0.50	2.871	0.034	5.61	0.13
900.34	−121.24	0.52	−41.64581	0.00061	22.93	0.25	2.997	0.047	5.58	0.17
925.13	−146.40	0.75	−41.7179	0.0010	23.62	0.33	3.199	0.068	5.68	0.10
935.42	−228.08	0.34	−41.7096	0.0015	24.84	0.13	3.535	0.066	5.46	0.22
950.39	−189.83	0.34	−41.83544	0.00062	23.58	0.38	3.448	0.051	5.51	0.20
975.18	−104.97	0.28	−42.01662	0.00077	21.48	0.31	3.272	0.063	5.78	0.11
998.10	−1.47	0.31	−42.1554	0.0018	19.56	0.19	3.151	0.055	6.18	0.38

Table SB3. Numerical data for bubble properties obtained from the cell analysis (top) and from radial density profiles (bottom) for the **mW-4 system** at $T = 298$ K: Bubble volume fraction (α), bubble radius (r^*) estimated from the bubble volume, bubble shape anisotropy factor (κ^2), bubble radius (r_0) estimated from shell density profile, fitted bulk liquid density ($\rho_{\mathrm{liq,fit}}$) from shell density profile, and fraction of molecules in the interfacial region (α_{int}). Uncertainties (u) for α are calculated from the standard deviation. All other uncertainties (u) represent the standard error of the mean at 95% confidence interval from block analysis.

ρ	α	u	r^*	u	κ^2	u
kg/m^3			nm		10^{-2}	
800.25	0.1481	0.0032	1.31270	0.00061	1.778	0.073
825.04	0.1251	0.0032	1.2409	0.0010	2.068	0.092
850.30	0.1017	0.0033	1.1580	0.0015	2.331	0.093
875.09	0.0784	0.0034	1.06166	0.00099	2.79	0.12
900.34	0.0542	0.0036	0.9388	0.0023	3.55	0.15
925.13	0.0280	0.0039	0.7513	0.0046	5.61	0.25

ρ	r_0	u	$\rho_{\mathrm{liq,fit}}$	u	α_{int}	u
kg/m^3	nm		kg/m^3			
800.25	1.40106	0.00076	989.6	1.1	0.0787	0.0015
825.04	1.3287	0.0013	988.3	1.9	0.0684	0.0015
850.30	1.2460	0.0022	986.3	2.0	0.0570	0.0014
875.09	1.1501	0.0024	983.2	2.6	0.04677	0.00064
900.34	1.0265	0.0020	977.2	1.3	0.03516	0.00079
925.13	0.8396	0.0043	968.1	1.3	0.02208	0.00042

Table SB4. Numerical data for system properties obtained for the **mW-8 system** at $T = 298$ K: Pressure (P), potential energy (U), residual isochoric heat capacity ($C_{V,\mathrm{res}}$), viscosity (η), and diffusion coefficient (D). Density values followed by (m) denote metastable homogeneous systems. All uncertainties (u) represent the standard error of the mean at 95% confidence interval from block analysis.

ρ	P	u	U	u	$C_{V,\mathrm{res}}$	u	η	u	D	u
kg/m^3	MPa		kJ/mol		J/(mol·K)		10^{-4} Pa·s		10^{-5} cm^2/s	
800.25	−45.445	0.10	−41.7150	0.0014	21.72	0.49	2.515	0.042	5.699	0.086
825.22	−47.579	0.10	−41.7581	0.0012	21.59	0.25	2.564	0.060	5.868	0.092
850.24	−50.25	0.10	−41.80020	0.00082	21.80	0.38	2.717	0.038	5.939	0.086
875.26	−53.867	0.086	−41.84168	0.00078	21.60	0.28	2.765	0.056	5.926	0.048
900.23	−58.68	0.13	−41.8822	0.0012	21.61	0.26	2.873	0.045	5.98	0.10
925.25	−66.19	0.21	−41.92037	0.00091	21.79	0.23	2.969	0.048	6.008	0.094
945.25	−77.23	0.26	−41.94971	0.00046	21.90	0.42	3.077	0.046	6.182	0.087
950.27	−81.85	0.24	−41.95573	0.00031	22.18	0.39	3.139	0.049	6.142	0.035
955.30	−88.82	0.52	−41.96016	0.00069	22.19	0.39	3.174	0.085	6.092	0.048
960.32	−101.37	0.61	−41.95962	0.00046	22.88	0.38	3.167	0.055	6.067	0.076
965.30	−141.57	0.12	−41.94590	0.00044	22.30	0.21	3.330	0.057	5.971	0.038
975.30	−104.13	0.10	−42.01622	0.00099	21.75	0.41	3.223	0.066	6.168	0.075
998.27	0.01	0.12	−42.1567	0.0011	19.77	0.34	3.166	0.033	6.49	0.11
950.27(m)	−189.70	0.11	−41.83200	0.00098	23.50	0.38	3.436	0.037	5.706	0.087
955.30(m)	−174.68	0.11	−41.87214	0.00060	23.23	0.29	3.401	0.065	5.814	0.040
960.32(m)	−158.732	0.063	−41.90971	0.00022	22.84	0.49	3.353	0.040	5.882	0.066

Table SB5. Numerical data for bubble properties obtained from the cell analysis (top) and from radial density profiles (bottom) for the **mW-8 system** at $T = 298$ K: Bubble volume fraction (α), bubble radius (r^*) estimated from the bubble volume, bubble shape anisotropy factor (κ^2), bubble radius (r_0) estimated from shell density profile, fitted bulk liquid density ($\rho_{\mathrm{liq,fit}}$) from shell density profile, and fraction of molecules in the interfacial region (α_{int}). Uncertainties (u) for α are calculated from the standard deviation. All other uncertainties (u) represent the standard error of the mean at 95% confidence interval from block analysis.

ρ	α	u	r^*	u	κ^2	u
kg/m^3			nm		10^{-3}	
800.25	0.1733	0.0011	2.76675	0.00045	4.32	0.18
825.22	0.1492	0.0011	2.63242	0.00044	4.72	0.19
850.24	0.1251	0.0011	2.48196	0.00066	5.04	0.20
875.26	0.1009	0.0012	2.31037	0.00043	5.80	0.22
900.23	0.0767	0.0012	2.1080	0.0011	7.26	0.31
925.25	0.0519	0.0013	1.8512	0.0020	9.24	0.37
945.25	0.0313	0.0013	1.5632	0.0025	12.58	0.48
950.27	0.02575	0.00010	1.46501	0.00187	15.0	2.0
955.30	0.01971	0.00027	1.3397	0.0061	17.3	1.8
960.32	0.01268	0.00045	1.154	0.014	24.3	3.3

ρ	r_0	u	$\rho_{\mathrm{liq,fit}}$	u	α_{int}	u
kg/m^3	nm		kg/m^3			
800.25	2.85352	0.00070	994.41	0.26	0.0464	0.0015
825.22	2.71878	0.00076	992.71	0.61	0.04054	0.00040
850.24	2.56850	0.00031	991.40	0.23	0.03417	0.00027
875.26	2.3964	0.0011	989.88	0.32	0.02847	0.00067
900.23	2.1940	0.0015	988.47	0.27	0.02294	0.00021
925.25	1.9369	0.0022	986.13	0.16	0.01701	0.00025
945.25	1.6491	0.0027	983.26	0.13	0.01168	0.00015
950.27	1.5503	0.0022	981.957	0.054	0.01032	0.00035
955.30	1.4263	0.0069	980.20	0.44	0.00855	0.00013
960.32	1.241	0.014	977.05	0.56	0.00675	0.00046

Table SB6. Numerical data for system properties obtained for the **mW-16 system** at $T = 298$ K: Pressure (P), potential energy (U), residual isochoric heat capacity ($C_{V,\mathrm{res}}$), viscosity (η), and diffusion coefficient (D). Density values followed by (m) denote metastable homogeneous systems. All uncertainties (u) represent the standard error of the mean at 95% confidence interval from block analysis.

ρ	P	u	U	u	$C_{V,\mathrm{res}}$	u	η	u	D	u
kg/m^3	MPa		kJ/mol		J/(mol·K)		10^{-4} Pa·s		10^{-5} cm^2/s	
800.00	−22.927	0.044	−41.93076	0.00038	21.26	0.42	2.487	0.046	5.965	0.053
825.00	−23.999	0.043	−41.95283	0.00022	21.20	0.22	2.560	0.040	6.049	0.023
850.00	−25.318	0.050	−41.97388	0.00045	20.97	0.15	2.660	0.042	6.098	0.021
875.00	−26.982	0.045	−41.99452	0.00022	20.94	0.22	2.736	0.052	6.180	0.062
900.00	−29.210	0.037	−42.01461	0.00017	21.18	0.33	2.810	0.023	6.217	0.043
925.00	−32.501	0.051	−42.03425	0.00026	21.32	0.25	2.908	0.034	6.287	0.053
950.00	−38.179	0.075	−42.05251	0.00029	21.10	0.37	2.974	0.048	6.331	0.040
954.99	−39.976	0.15	−42.05581	0.00023	21.38	0.35	3.045	0.059	6.341	0.052
960.00	−42.357	0.097	−42.05877	0.00014	21.51	0.25	3.066	0.028	6.370	0.032
965.00	−45.31	0.14	−42.06145	0.00021	21.44	0.15	3.041	0.059	6.391	0.035
975.00	−57.92	0.21	−42.06108	0.00024	22.09	0.28	3.155	0.049	6.361	0.011
980.00	−84.855	0.056	−42.046120	0.000068	21.02	0.26	3.229	0.095	6.298	0.027
985.00	−63.293	0.040	−42.07769	0.00011	20.59	0.22	3.197	0.043	6.395	0.034
998.00	−1.342	0.051	−42.15516	0.00038	19.65	0.38	3.115	0.061	6.571	0.041
954.99(m)	−175.576	0.026	−41.86978	0.00034	23.30	0.31	3.380	0.076	5.927	0.026
960.00(m)	−159.664	0.035	−41.90755	0.00064	22.75	0.34	3.389	0.053	6.000	0.019
965.00(m)	−142.614	0.041	−41.94410	0.00036	22.41	0.28	3.307	0.059	6.099	0.027
975.00(m)	−105.285	0.033	−42.01319	0.00033	21.67	0.33	3.260	0.064	6.237	0.025

Table SB7. Numerical data for bubble properties obtained from the cell analysis (top) and from radial density profiles (bottom) for the **mW-16 system** at $T = 298$ K: Bubble volume fraction (α), bubble radius (r^*) estimated from the bubble volume, bubble shape anisotropy factor (κ^2), bubble radius (r_0) estimated from shell density profile, fitted bulk liquid density ($\rho_{\mathrm{liq,fit}}$) from shell density profile, and fraction of molecules in the interfacial region (α_{int}). Uncertainties (u) for α are calculated from the standard deviation. All other uncertainties (u) represent the standard error of the mean at 95% confidence interval from block analysis.

ρ	α	u	r^*	u	κ^2	u
kg/m^3			nm		10^{-3}	
800.00	0.18606	0.00040	5.66643	0.00023	1.127	0.045
825.00	0.16151	0.00041	5.40531	0.00028	1.040	0.039
850.00	0.13691	0.00042	5.11568	0.00019	1.186	0.045
875.00	0.11233	0.00042	4.78917	0.00035	1.346	0.052
900.00	0.08768	0.00043	4.40951	0.00020	1.521	0.065
925.00	0.06289	0.00042	3.94717	0.00068	2.057	0.077
950.00	0.03774	0.00045	3.32928	0.00086	3.01	0.13
954.99	0.032598	0.000031	3.1706	0.0010	3.48	0.62
960.00	0.027331	0.000061	2.9897	0.0022	3.65	0.64
965.00	0.022031	0.000079	2.7824	0.0034	4.30	0.36
975.00	0.01001	0.00011	2.1386	0.0078	6.57	0.88

ρ	r_0	u	$\rho_{\mathrm{liq,fit}}$	u	α_{int}	u
kg/m^3	nm		kg/m^3			
800.00	5.75276	0.00059	995.28	0.17	0.02641	0.00064
825.00	5.49174	0.00063	994.81	0.30	0.02258	0.00079
850.00	5.20256	0.00045	994.420	0.095	0.01952	0.00046
875.00	4.87544	0.00057	993.71	0.13	0.01651	0.00049
900.00	4.49560	0.00053	993.073	0.093	0.01341	0.00017
925.00	4.03296	0.00071	992.21	0.10	0.01040	0.00022
950.00	3.4148	0.0010	990.819	0.045	0.00709	0.00014
954.99	3.25706	0.00078	990.419	0.046	0.00644	0.00018
960.00	3.0753	0.0026	989.89	0.13	0.00562	0.00017
965.00	2.8674	0.0032	989.19	0.11	0.004841	0.000086
975.00	2.2246	0.0078	986.362	0.065	0.002776	0.000043

Table SB8. Numerical data for system properties obtained for the **mW-32 system** at $T = 298$ K: Pressure (P), potential energy (U), residual isochoric heat capacity ($C_{V,\mathrm{res}}$), and viscosity (η). Density values followed by (m) denote metastable homogeneous systems. All uncertainties (u) represent the standard error of the mean at 95% confidence interval from block analysis.

ρ	P	u	U	u	$C_{V,\mathrm{res}}$	u	η	u
kg/m^3	MPa		kJ/mol		J/(mol·K)		10^{-4} Pa·s	
800.00	−11.672	0.011	−42.04228	0.00011	20.93	0.26	2.463	0.063
825.00	−12.201	0.014	−42.05345	0.00014	20.94	0.28	2.543	0.058
850.00	−12.8579	0.0055	−42.064031	0.000093	21.04	0.27	2.697	0.051
875.00	−13.684	0.012	−42.07431	0.00012	21.22	0.31	2.746	0.038
900.00	−14.761	0.020	−42.08435	0.00039	21.07	0.32	2.808	0.050
925.00	−16.315	0.018	−42.09411	0.00032	21.17	0.40	2.886	0.040
950.00	−18.877	0.057	−42.10349	0.00016	20.86	0.32	2.938	0.061
955.00	−19.662	0.030	−42.105139	0.000037	20.89	0.25	2.976	0.058
960.00	−20.591	0.038	−42.106811	0.000083	20.95	0.25	2.999	0.084
965.00	−21.721	0.022	−42.108277	0.000062	21.26	0.34	3.059	0.066
975.00	−25.260	0.035	−42.110456	0.000057	20.97	0.21	3.113	0.035
980.00	−84.825	0.013	−42.046036	0.000050	21.21	0.19	3.214	0.060
985.00	−63.250	0.013	−42.077718	0.000046	20.63	0.22	3.202	0.048
998.32	0.281	0.013	−42.157149	0.000065	18.68	0.92	3.086	0.060
960.00(m)	−159.633	0.010	−41.90762	0.00014	22.76	0.54	3.334	0.071
965.00(m)	−142.6070	0.0075	−41.944156	0.000094	22.52	0.38	3.350	0.073
975.00(m)	−105.267	0.012	−42.00850	0.00031	21.75	0.20	3.218	0.056

Table SB9. Numerical data for bubble properties obtained from the cell analysis (top) and from radial density profiles (bottom) for the **mW-32 system** at $T = 298$ K: Bubble volume fraction (α), bubble radius (r^*) estimated from the bubble volume, bubble shape anisotropy factor (κ^2), bubble radius (r_0) estimated from shell density profile, fitted bulk liquid density ($\rho_{\mathrm{liq,fit}}$) from shell density profile, and fraction of molecules in the interfacial region (α_{int}). Uncertainties (u) for α are calculated from the standard deviation. All other uncertainties (u) represent the standard error of the mean at 95% confidence interval from block analysis.

ρ	α	u	r^*	u	κ^2	u
kg/m^3			nm		10^{-4}	
800.00	0.19231	0.00013	11.45840	0.00022	2.91	0.11
825.00	0.16751	0.00015	10.94286	0.00029	2.616	0.096
850.00	0.14270	0.00015	10.37355	0.00019	3.25	0.15
875.00	0.11788	0.00015	9.73342	0.00015	3.84	0.15
900.00	0.09303	0.00016	8.99490	0.00026	4.47	0.21
925.00	0.06813	0.00014	8.10783	0.00011	4.54	0.16
950.00	0.04308	0.00016	6.95914	0.00076	6.19	0.28
955.00	0.038022	0.000016	6.67521	0.00091	7.7	1.6
960.00	0.0329523	0.0000077	6.36426	0.00050	7.9	1.2
965.00	0.027845	0.000017	6.0169	0.0012	9.0	1.7
975.00	0.017395	0.000033	5.1435	0.0033	11.6	1.7

ρ	r_0	u	$\rho_{\mathrm{liq,fit}}$	u	α_{int}	u
kg/m^3	nm		kg/m^3			
800.00	11.54564	0.00036	996.508	0.080	0.01451	0.00028
825.00	11.03004	0.00051	996.202	0.037	0.01243	0.00029
850.00	10.46040	0.00028	995.956	0.052	0.01097	0.00022
875.00	9.81970	0.00074	995.636	0.055	0.00939	0.00025
900.00	9.08124	0.00088	995.377	0.063	0.00768	0.00032
925.00	8.19432	0.00053	994.962	0.048	0.00589	0.00010
950.00	7.04608	0.00098	994.401	0.048	0.004131	0.000067
955.00	6.7614	0.0013	994.185	0.036	0.00386	0.00011
960.00	6.45080	0.00059	993.982	0.030	0.003401	0.000076
965.00	6.1031	0.0014	993.740	0.050	0.00304	0.00013
975.00	5.2296	0.0036	992.972	0.055	0.001164	0.000026

Table SB10. Numerical data for system properties obtained for the **TIP4P/2005-4 system** at $T = 298$ K: Pressure (P), potential energy (U), residual isochoric heat capacity ($C_{V,\mathrm{res}}$), viscosity (η), and diffusion coefficient (D). Density values followed by (m) denote metastable homogeneous systems. All uncertainties (u) represent the standard error of the mean at 95% confidence interval from block analysis.

ρ	P	u	U	u	$C_{V,\mathrm{res}}$	u	η	u	D	u
kg/m^3	MPa		kJ/mol		J/(mol·K)		10^{-4} Pa·s		10^{-5} cm^2/s	
800.25	−92.8	1.0	−46.752	0.010	70.2	1.5	6.375	0.468	2.061	0.054
825.04	−98.9	1.3	−46.8623	0.0075	71.6	2.8	6.863	0.601	2.054	0.042
850.30	−108.5	1.4	−46.9702	0.0096	72.1	1.8	7.706	0.886	2.029	0.062
875.09	−127.9	3.2	−47.0776	0.0091	71.9	2.9	8.411	0.860	2.020	0.099
884.91	−188.54	0.73	−47.097	0.011	74.3	2.7	9.586	0.482	1.839	0.053
900.35	−168.34	0.40	−47.2220	0.0090	76.6	1.9	9.622	0.773	1.897	0.089
925.13	−130.54	0.70	−47.4167	0.0079	74.1	1.8	9.145	0.691	1.89	0.12
950.39	−85.76	0.67	−47.554	0.011	69.0	3.0	8.757	0.739	2.009	0.066
975.18	−38.81	0.71	−47.6459	0.0076	66.1	2.2	8.224	0.481	2.137	0.097
995.29	3.16	0.56	-47.7193	0.0078	63.1	1.3	8.118	0.844	2.133	0.058

Table SB11. Numerical data for bubble properties obtained from the cell analysis (top) and from radial density profiles (bottom) for the **TIP4P/2005-4 system** at $T = 298$ K: Bubble volume fraction (α), bubble radius (r^*) estimated from the bubble volume, bubble shape anisotropy factor (κ^2), bubble radius (r_0) estimated from shell density profile, fitted bulk liquid density ($\rho_{\mathrm{liq,fit}}$) from shell density profile, and fraction of molecules in the interfacial region (α_{int}). Uncertainties (u) for α are calculated from the standard deviation. All other uncertainties (u) represent the standard error of the mean at 95% confidence interval from block analysis.

ρ	α	u	r^*	u	κ^2	u
kg/m^3			nm		10^{-2}	
800.25	0.1231	0.0013	1.2340	0.0043	1.93	0.39
825.04	0.0996	0.0015	1.1498	0.0057	2.58	0.29
850.30	0.0712	0.0010	1.0280	0.0045	3.52	0.90
875.09	0.0396	0.0019	0.843	0.023	4.5	1.6

ρ	r_0	u	$\rho_{\mathrm{liq,fit}}$	u	α_{int}	u
kg/m^3	nm		kg/m^3			
800.25	1.3183	0.0076	956.3	2.7	0.0747	0.0060
825.04	1.2328	0.0062	952.9	2.5	0.0638	0.0027
850.30	1.1090	0.0035	941.2	2.4	0.0451	0.0019
875.09	0.919	0.027	926.9	1.5	0.0312	0.0034

Table SB12. Numerical data for system properties obtained for the **TIP4P/2005-6 system** at $T = 298$ K: Pressure (P), potential energy (U), residual isochoric heat capacity ($C_{V,\mathrm{res}}$), viscosity (η), and diffusion coefficient (D). Density values followed by (m) denote metastable homogeneous systems. All uncertainties (u) represent the standard error of the mean at 95% confidence interval from block analysis.

ρ	P	u	U	u	$C_{V,\mathrm{res}}$	u	η	u	D	u
kg/m^3	MPa		kJ/mol		J/(mol·K)		10^{-4} Pa·s		10^{-5} cm^2/s	
800.02	−60.92	0.69	−47.0361	0.0039	69.0	2.3	6.23	0.61	2.131	0.062
824.97	−65.05	0.61	−47.1045	0.0028	69.4	2.4	6.87	0.39	2.102	0.043
850.05	−70.08	0.65	−47.1766	0.0054	68.8	3.3	7.19	0.89	2.145	0.051
874.99	−77.14	0.70	−47.2476	0.0018	69.5	3.3	7.57	0.62	2.114	0.047
899.94	−90.4	1.2	−47.3200	0.0044	70.7	2.5	8.3	1.0	2.070	0.017
910.05	−105.2	4.0	−47.3406	0.0041	71.1	2.1	8.37	0.51	2.029	0.043
915.04	−145.82	0.30	−47.3449	0.0046	72.5	4.5	9.79	0.57	1.965	0.041
925.02	−129.93	0.32	−47.4091	0.0057	73.0	4.3	9.19	0.51	2.032	0.086
949.96	−86.55	0.30	−47.5474	0.0067	70.7	2.2	8.12	0.49	2.077	0.030
975.05	−38.68	0.21	−47.6512	0.0034	65.5	2.2	9.01	0.75	2.167	0.060
995.00	3.11	0.34	−47.7203	0.0023	64.2	2.3	7.83	0.67	2.287	0.031
889.96(m)	−181.61	0.42	−47.1385	0.0058	76.4	4.1	9.93	0.53	1.922	0.044
894.95(m)	−175.11	0.30	−47.1877	0.0033	77.3	2.4	9.44	0.57	1.910	0.038

Table SB13. Numerical data for bubble properties obtained from the cell analysis (top) and from radial density profiles (bottom) for the **TIP4P/2005-6 system** at $T = 298$ K: Bubble volume fraction (α), bubble radius (r^*) estimated from the bubble volume, bubble shape anisotropy factor (κ^2), bubble radius (r_0) estimated from shell density profile, fitted bulk liquid density ($\rho_{\mathrm{liq,fit}}$) from shell density profile, and fraction of molecules in the interfacial region (α_{int}). Uncertainties (u) for α are calculated from the standard deviation. All other uncertainties (u) represent the standard error of the mean at 95% confidence interval from block analysis.

ρ kg/m^3	α	u	r^*	u	κ^2	u
			nm		10^{-2}	
800.00	0.14749	0.00033	1.9665	0.0015	1.06	0.41
825.00	0.12190	0.00064	1.8455	0.0032	1.038	0.090
850.00	0.09617	0.00042	1.7052	0.0025	1.22	0.63
875.00	0.0697	0.0013	1.5311	0.0098	1.26	0.38
900.00	0.0403	0.0019	1.276	0.020	2.42	0.27
910.00	0.0240	0.0017	1.069	0.024	3.60	0.93

ρ kg/m^3	r_0	u	$\rho_{\mathrm{liq,fit}}$	u	α_{int}	u
	nm		kg/m^3			
800.00	2.0499	0.0027	966.7	2.0	0.0578	0.0042
825.00	1.9305	0.0033	964.8	1.3	0.0494	0.0013
850.00	1.7874	0.0071	960.1	2.2	0.0401	0.0026
875.00	1.615	0.011	956.9	1.4	0.03083	0.00097
900.00	1.355	0.020	947.9	2.0	0.0218	0.0017
910.00	1.143	0.058	939.9	4.5	0.0151	0.0022

Table SB14. Numerical data for system properties obtained for the **TIP4P/2005-8 system** at $T = 298$ K: Pressure (P), potential energy (U), residual isochoric heat capacity ($C_{V,\mathrm{res}}$), viscosity (η), and diffusion coefficient (D). Density values followed by (m) denote metastable homogeneous systems. All uncertainties (u) represent the standard error of the mean at 95% confidence interval from block analysis.

ρ	P	u	U	u	$C_{V,\mathrm{res}}$	u	η	u	D	u
kg/m^3	MPa		kJ/mol		J/(mol·K)		10^{-4} Pa·s		10^{-5} cm^2/s	
800.25	−45.84	0.15	−47.1885	0.0026	67.8	1.2	6.35	0.38	2.145	0.052
825.22	−48.61	0.25	−47.2415	0.0031	68.7	1.6	6.71	0.56	2.147	0.029
850.24	−51.70	0.24	−47.2972	0.0014	67.8	1.1	6.66	0.37	2.146	0.042
875.26	−56.61	0.35	−47.3479	0.0021	68.4	1.8	7.48	0.46	2.136	0.048
900.23	−63.38	0.40	−47.4002	0.0017	69.9	2.0	8.07	0.55	2.148	0.015
925.25	−81.8	1.5	−47.4470	0.0026	69.7	1.5	8.21	0.37	2.132	0.027
930.21	−120.4	1.0	−47.4430	0.0024	72.4	2.9	8.70	0.29	2.046	0.049
950.27	−86.09	0.17	−47.5471	0.0016	69.4	1.4	8.81	0.53	2.080	0.037
975.30	−38.20	0.12	−47.6495	0.0013	67.5	1.3	8.55	0.52	2.207	0.036
995.23	3.396	0.086	−47.7184	0.0016	63.5	1.5	8.82	0.38	2.257	0.044
900.23(m)	−167.909	0.058	−47.2289	0.0011	76.6	2.9	9.48	0.54	1.944	0.045
925.25(m)	−129.48	0.24	−47.4114	0.0024	74.1	2.0	9.06	0.38	2.027	0.033

Table SB15. Numerical data for bubble properties obtained from the cell analysis (top) and from radial density profiles (bottom) for the **TIP4P/2005-8 system** at $T = 298$ K: Bubble volume fraction (α), bubble radius (r^*) estimated from the bubble volume, bubble shape anisotropy factor (κ^2), bubble radius (r_0) estimated from shell density profile, fitted bulk liquid density ($\rho_{\mathrm{liq,fit}}$) from shell density profile, and fraction of molecules in the interfacial region (α_{int}). Uncertainties (u) for α are calculated from the standard deviation. All other uncertainties (u) represent the standard error of the mean at 95% confidence interval from block analysis.

ρ	α	u	r^*	u	κ^2	u
kg/m^3			nm		10^{-3}	
800.25	0.15858	0.00063	2.6862	0.0036	4.11	0.80
825.22	0.13359	0.00037	2.5369	0.0024	5.22	0.88
850.24	0.10811	0.00050	2.3641	0.0037	5.57	0.88
875.26	0.08204	0.00067	2.1563	0.0058	6.98	0.72
900.23	0.0562	0.0010	1.900	0.011	10.1	1.0
925.25	0.02331	0.00158	1.416	0.032	16.5	2.5

ρ	r_0	u	$\rho_{\mathrm{liq,fit}}$	u	α_{int}	u
kg/m^3	nm		kg/m^3			
800.25	2.7731	0.0043	973.9	1.6	0.0458	0.0010
825.22	2.6217	0.0029	971.8	1.4	0.0420	0.0022
850.24	2.4490	0.0043	968.43	0.80	0.0320	0.0015
875.26	2.2403	0.0047	965.9	1.2	0.0268	0.0014
900.23	1.984	0.012	963.0	1.7	0.01988	0.00048
925.25	1.497	0.032	952.7	1.5	0.0109	0.0013

Appendix C. Supporting information for chapter 4

Fitting of $\epsilon_{N_2-N_2}$ and $\sigma_{N_2-N_2}$. $N_2 - N_2$ interaction takes the form of shifted-force LJ potential. Isochoric–Isothermal (NVT) Gibbs ensemble Monte Carlo (GEMC) simulations were performed using the Monte Carlo for Complex Chemical Systems–Minnesota (MCCCS–MN) software developed in house, at T = 100, 105, 108, 111, 114, 117, 120, 123 K, with total number of molecules N = 2000. For each temperature, 8 independent simulations were run and statistics was collected from 100000 Monte Carlo cycles (MCC; one MCC consists of N randomly selected moves) of production run. Uncertainties were estimated as the standard error of the mean at 95% confidence interval. Critical points were obtained by fitting the densities of the four highest temperatures to the law of rectilinear diameters with critical exponent of 0.326. The simulated vapor–liquid coexistence curve (VLCC) and Clausius-Clapeyron plot with $\epsilon_{N_2-N_2/k_B}$ = 135 K and $\sigma_{N_2-N_2}$ = 3.62 Å are given in Figure SC1, along with the experimental data taken from NIST [224].

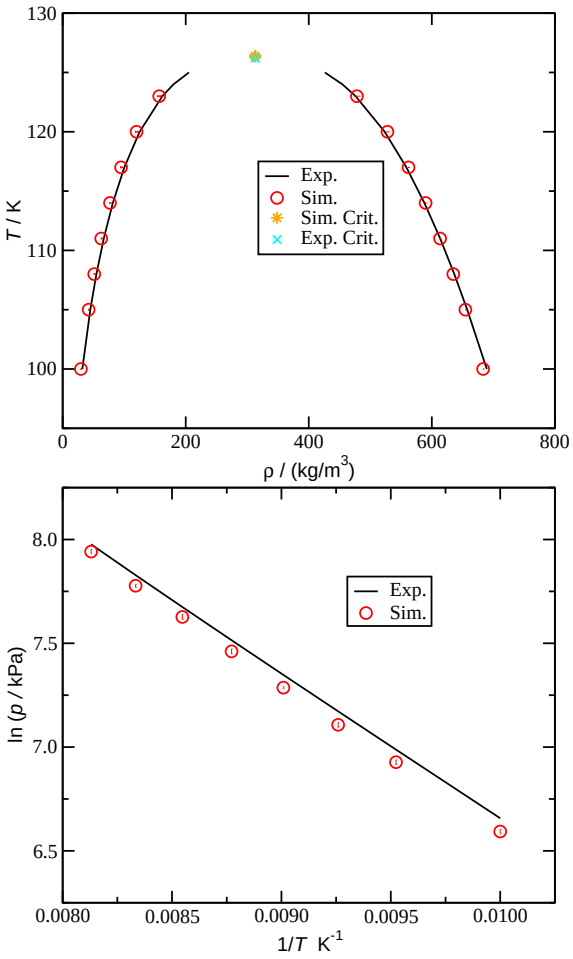


Figure SC1. VLCC (top) and Clausius-Clapeyron plot of N_2.

Fitting of $\epsilon_{\text{water}-\text{N}_2}$ and $\sigma_{\text{water}-\text{N}_2}$. water $-$ N_2 interactions takes the form of shifted-force LJ potential. To calculate solubility of N_2 in water, Isobaric–Isothermal (NPT) GEMC simulations were performed with $N_{\text{N}_2} = 50$ and $N_{\text{water}} = 2000$ at two state points: $T = 298.07$ K and $p = 1.004$ MPa; $T = 342.98$ K and $p = 6.941$ MPa. The fitting procedure started with coarse-grained grid points in both $\epsilon_{\text{water}-\text{N}_2}$ and $\sigma_{\text{water}-\text{N}_2}$, and then went through several iterations with higher grid resolutions to locate the best parameters. Results from the final iteration are plotted in Figure SC2. Statistics was collected from 100000 MCCs of production run from 8 independent simulations. Uncertainties were estimated as standard error of the mean at 95% confidence interval. The final values for $\epsilon_{\text{water}-\text{N}_2}$ and $\sigma_{\text{water}-\text{N}_2}$ are chosen as those giving the smallest relative mean squared error for all 32 data points.

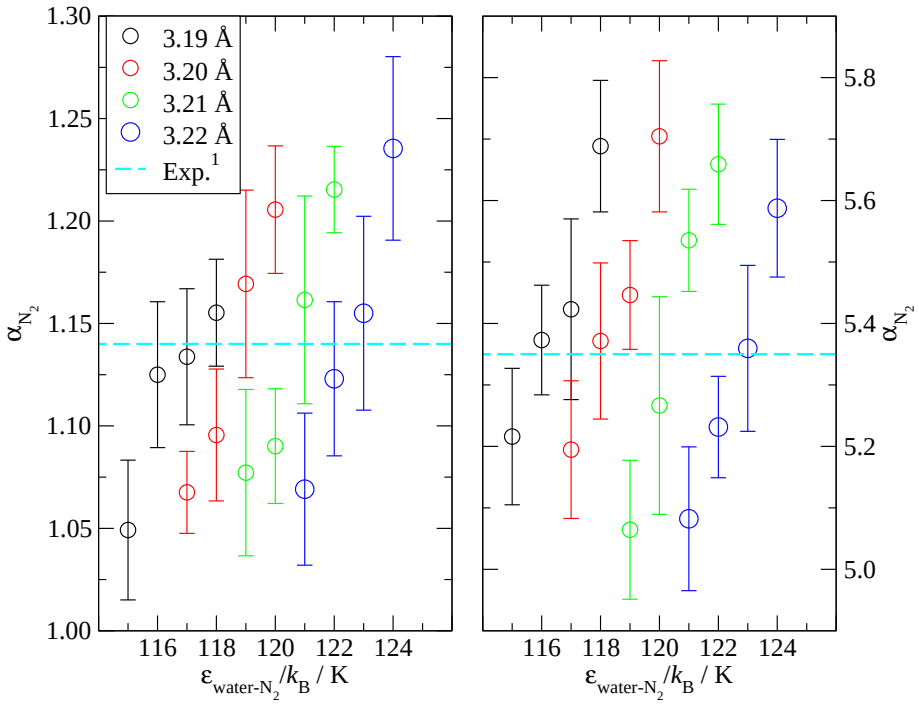


Figure SC2. Simulated solubility of N_2 in liquid water as a function of $\epsilon_{\text{water}-\text{N}_2}$ under different values of $\sigma_{\text{water}-\text{N}_2}$ at $T = 298.07$ K and $p = 1.004$ MPa (left) and $T = 343.98$ K and $p = 6.941$ MPa (right). Circles denote simulation data and cyan lines denote experimental solubility values [225].

Table SC1. Comparison between simulated solubility of N_2 in water χ_{N_2} using TIP4P/2005 water and TraPPE nitrogen model and experimental solubility $\chi_{N_2,\,exp}$ at different state points [225]. Solubility is represented as mole fraction of N_2 in the mixture. Uncertainties (u) are estimated as the standard error of the mean at 95% confidence interval obtained from 8 independent production runs of 50000 MCCs.

T	p	χ_{N_2}	u	$\chi_{N_2,\,exp}$
K	MPa	10^{-4}	10^{-4}	10^{-4}
298.07	1.004	0.8	0.10	1.14
343.98	6.941	4.2	0.27	5.35

Table SC2. Linear box dimension L, ratio between LJ wall radius and box length R_w/L, number of water molecules N_{water}, and number of nitrogen molecules N_{N_2} for all simulated systems.

System	L nm	R_w/L	N_{water}	N_{N_2}
water (mW)	32	0.15	1093177	0
	64	0.15, 0.075	8745414	0
	128	0.15, 0.075	69963308	0
	256	0.15, 0.075	559706465	0
	512	0.075	4477651718	0
water (mW)/nitrogen (single-site)	128	0.15	69963308	832
water (TIP4P/2005)	32	0.15	1089891	0
	48	0.15	3678381	0
	64	0.15	8719125	0
water (TIP4P/2005)/nitrogen (TraPPE)	64	0.15	8719125	104

Table SC3. Initial bubble radius R_0, initial wall velocity $\frac{dR}{dt}_0$, bubble collapse time t_c, and most negative wall velocity captured preceding initial collapse point $\frac{dR}{dt}_m$ for all simulated systems. Value of $\frac{dR}{dt}_m$ for TIP4P/2005-32-0.15 system is not shown due to large noise.

system	L nm	R_w/L	R_0 nm	$\frac{dR}{dt}_0$ m/s	t_c ps	$\frac{dR}{dt}_m$ m/s
water (mW)	32	0.15	4.75	-31	18.75	-987
	64	0.15	9.56	-33	33.125	-3122
		0.075	4.81	-7	34.875	-788
	128	0.15	19.16	-25	63.25	-4055
		0.075	9.64	-11	69.625	-2607
	256	0.15	38.37	-28	123.625	-8012
		0.075	19.26	-7	149	-2844
	512	0.075	38.48	-6	309	-4662
water (mW)/nitrogen (single-site)	128	0.15	19.16	-24.6	63.625	-1515
water (TIP4P/2005)	32	0.15	4.84	-25.3	69.7	
	48	0.15	7.21	-35.4	99	-319
	64	0.15	9.74	-2.8	122	-413
water (TIP4P/2005)/nitrogen (TraPPE)	64	0.15	9.71	-8.9	131	-41

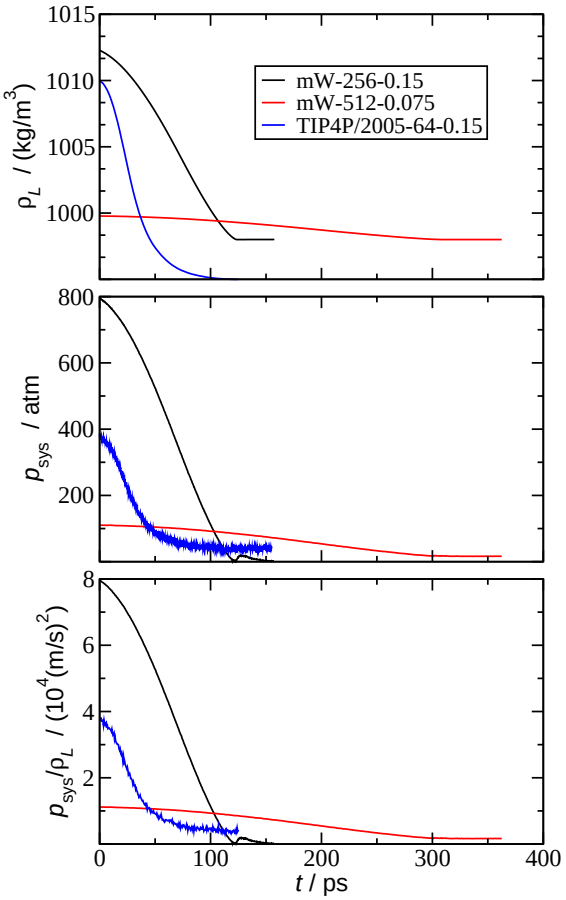


Figure SC3. Time evolution of outside liquid density (top), system pressure (middle), and the ratio between them (bottom) for mW-256-0.15, mW-512-0.075, and TIP4P/2005-64-0.15 systems.

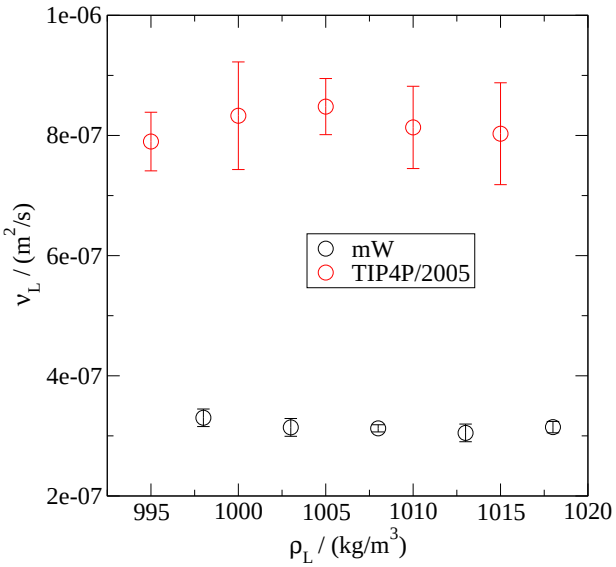


Figure SC4. Kinematic viscosity of liquid water as a function of density at T = 298 K. Calculation method is described in reference [155]. Uncertainties are estimated as the standard error of the mean at 95% confidence interval from 8 block averages, with block size of 2.5 ns and 0.5 ns for mW and TIP4P/2005 systems, respectively.

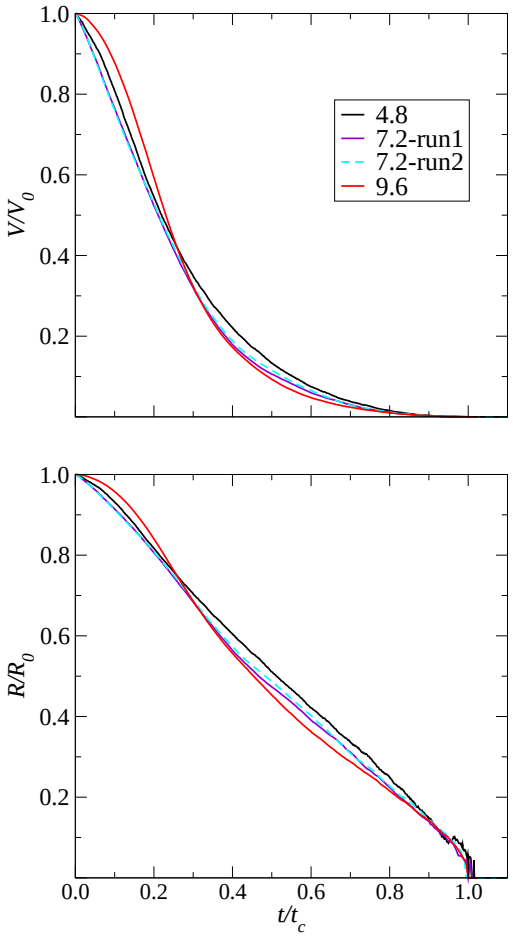


Figure SC5. Reduced bubble volume (top) and bubble radius (bottom) as a function of reduced time for all TIP4P/2005 water systems. Different colors represent systems with different initial bubble radius or different independent runs.

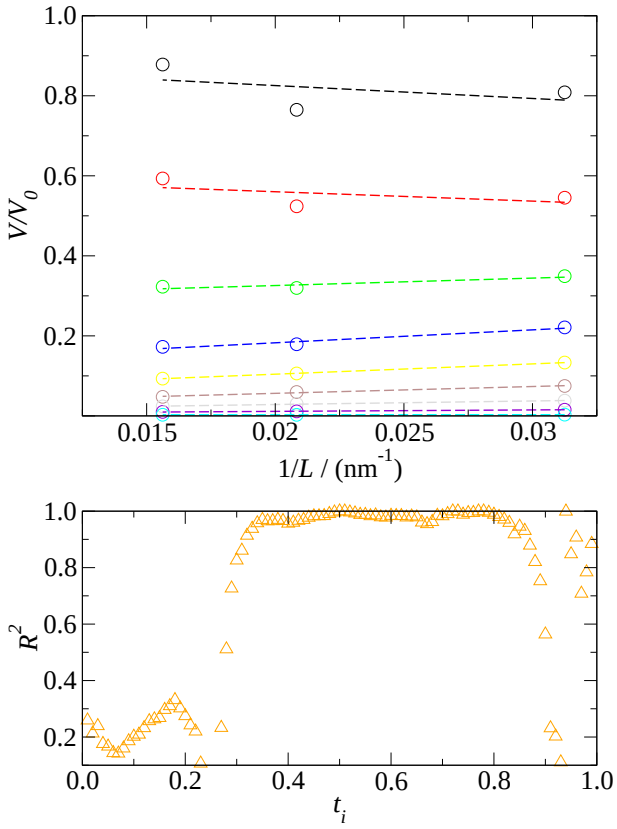


Figure SC6. V/V_0 as a function of $1/L$ obtained for TIP4P/2005 systems at $t/t_c = 0.1$ (black), 0.2, 0.3...0.9 (cyan). Circular dots represent simulation data and and dashed lines represent linearly fitted lines. Bottom graph shows R^2 as a function of t/t_c.

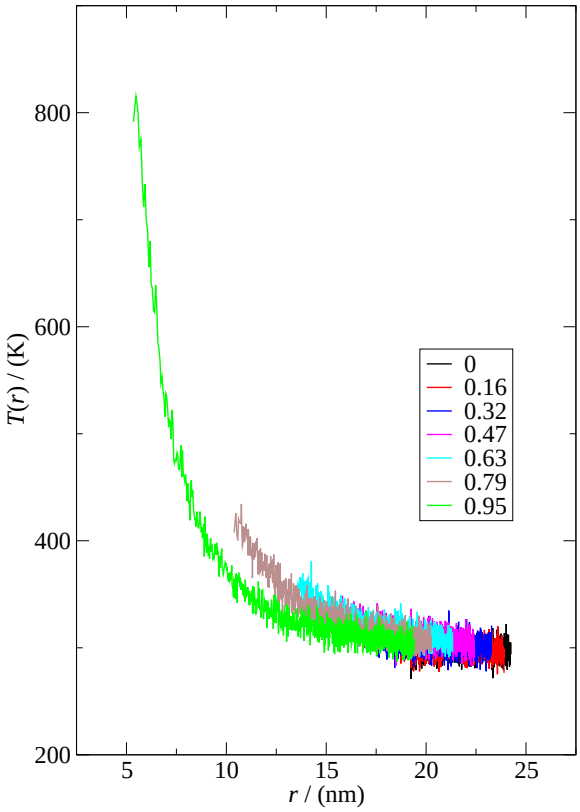


Figure SC7. Temperature shell profile at different values of t/t_c for the mW-128-0.15 system. Each shell contains approximately 1000 molecules.

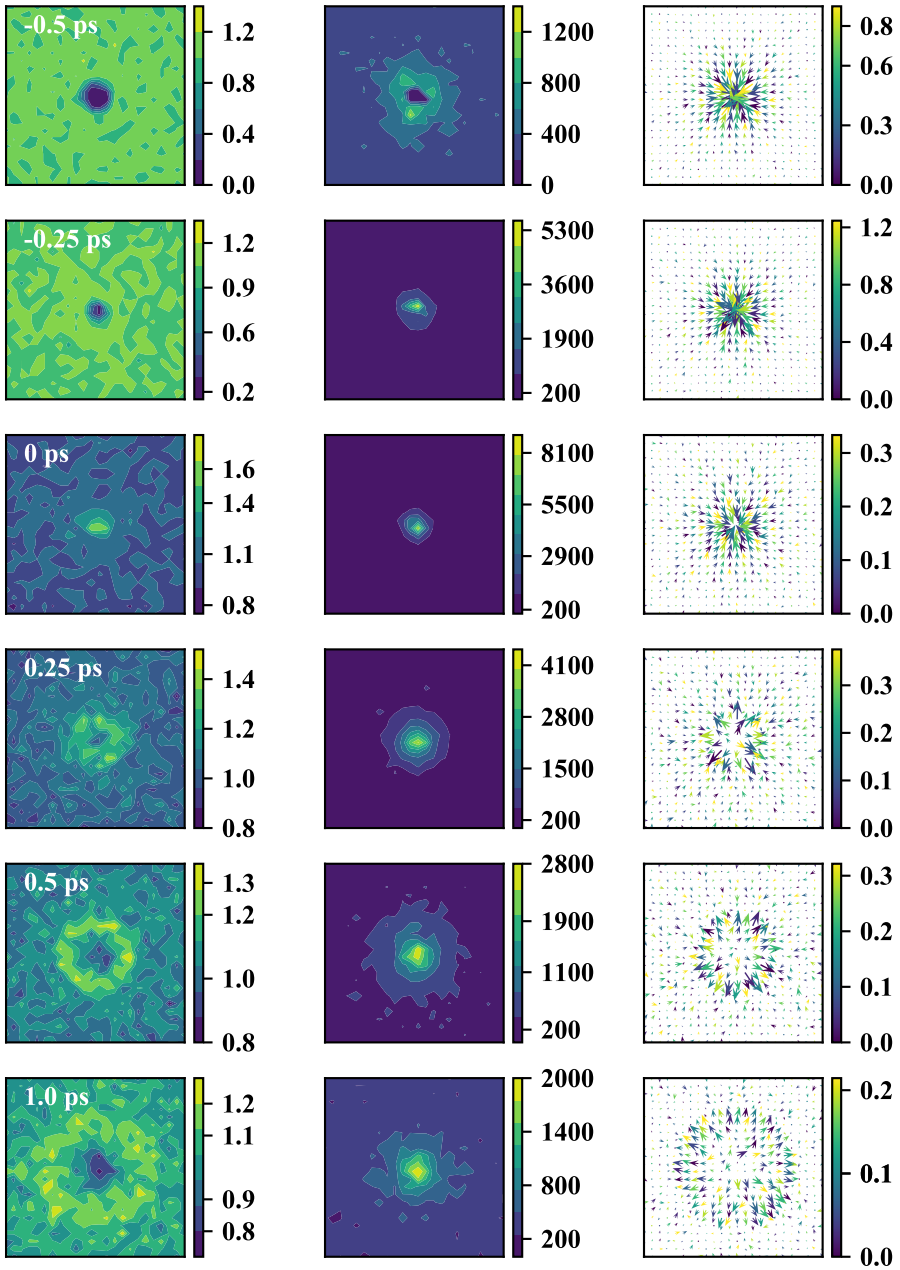


Figure SC8. Cross-sectional heat maps at $t - t_c = -0.5$, -0.25, 0, 0.25, 0.5, and 1 ps for mW-128-0.15 system. The central 24×24 nm^2 region of the simulation box is shown, with properties computed on meshes of size $1 \times 1 \times 1$ nm^3. Left: water number density normalized by the liquid density at $T = 298$ K and $p = 1$ bar; middle: kinetic temperature in units of Kelvin; right: voxel velocity normalized by speed of sound in bulk liquid water at $T = 298$ K.

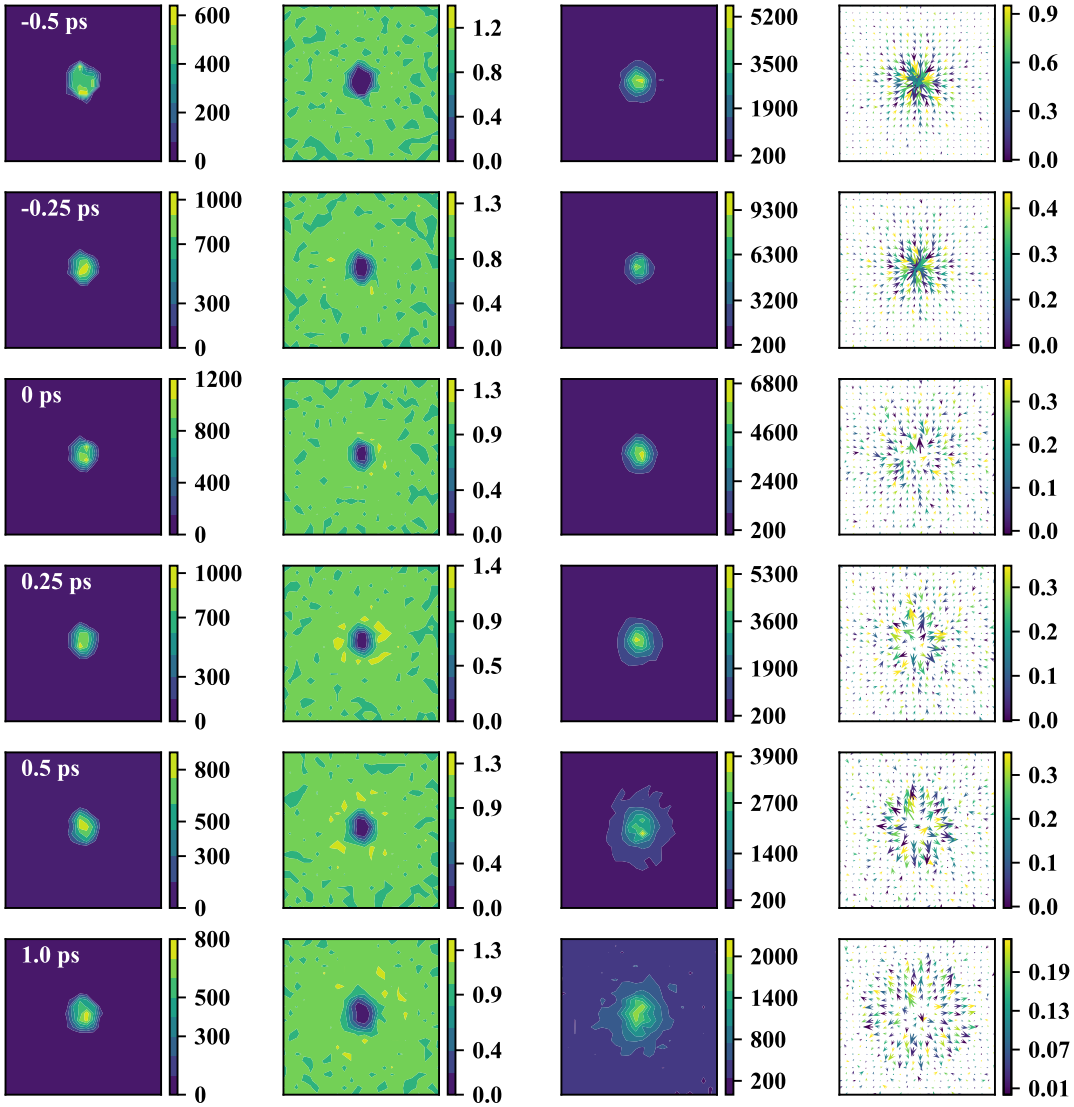


Figure SC9. Cross-sectional heat maps at $t - t_c = -0.5$, -0.25, 0, 0.25, 0.5, and 1 ps for mW-128-0.15-N_2 system. The central 24×24 nm^2 region of the simulation box is shown, with properties computed on meshes of size $1 \times 1 \times 1$ nm^3. Left: nitrogen number density normalized by the vapor density at $T = 298$ K and $p = 1$ bar; left middle: water number density normalized by the liquid density at $T = 298$ K and $p = 1$ bar; right middle: kinetic temperature in units of Kelvin; right: voxel velocity normalized by speed of sound in bulk liquid water at $T = 298$ K.

Appendix D. Supporting information for chapter 5

Table SD1. Simulated temperature T, the corresponding total volume of the two phases V, saturated liquid density ρ_{liq} and vapor density ρ_{vap}, vapor pressure p, heat of vaporization ΔH_{vap}, and their corresponding uncertainties u for all systems. Uncertainties are estimated as the standard error of the mean at 95% confidence interval obtained from block averages from the production period. For ΔH_{vap}, only frames with at least one molecule in the vapor phase are counted.

system	T	V	ρ_{liq}	u	ρ_{vap}	u	p	u	ΔH	u
	K	nm^3	$\mathrm{kg/m}^3$		$\mathrm{kg/m}^3$		kPa		kJ/mol	
	210	8.2	638	9	4.9	0.9	253	48	13.45	0.05
	230	8.2	608	17	9	7	498	399	12.7	0.6
	250	8.2	586	7	18	7	1081	418	12.4	0.4
PH_3-PBE-48	270	8.2	558	14	30	4	1951	256	11.8	0.4
	285	8.2	529	11	49	25	3412	1723	10.7	0.8
	290	8.2	524	13	59	11	4177	748	10.6	0.4
	295	8.2	522	22	54	12	3905	868	10.7	0.9
	210	16.0	645	10	4	2	230	90	13.9	0.2
	230	16.0	608	13	9	3	527	180	13.0	0.5
PH_3-PBE-96	250	16.0	596	19	22	7	1328	444	12.6	0.4
	270	16.0	564	13	34	5	2266	297	11.8	0.3
	285	16.0	544	17	75	26	5239	1779	10.4	0.4
	300	16.0	513	10	74	14	5461	1032	10.0	0.4
	190	8.2	705	33	14	12	655	564	13.1	0.5
	210	8.2	670	13	16	7	805	372	12.6	1.0
PH_3-BLYP-48	230	8.2	642	6	46	22	2615	1233	11.2	0.4
	250	8.2	590	17	84	36	5137	2173	9.5	0.5
	260	8.2	543	35	72	9	4595	584	9.2	0.5
	330	9.3	1310	45	28	14	998	488	16.9	0.2
	350	9.3	1280	39	53	29	1990	1079	16.0	1.2
AsH_3-PBE-48	370	9.3	1148	53	54	14	2149	551	14.4	1.3
	390	9.3	1143	36	113	30	4690	1235	13.9	1.4
	410	9.3	1092	75	190	61	8318	2678	11.6	1.0
	270	8.2	1845	67	8	8	147	143	17.1	0.8
	300	8.2	1704	68	20	3	407	62	16.1	0.9
	330	8.2	1685	94	45	23	992	504	16.4	0.8
SbH_3-PBE-48	360	11.7	1555	67	90	54	2160	1300	14.8	1.4
	370	11.7	1444	102	95	29	2332	707	13.9	0.8
	380	11.7	1543	31	185	81	4677	2059	13.8	2.0

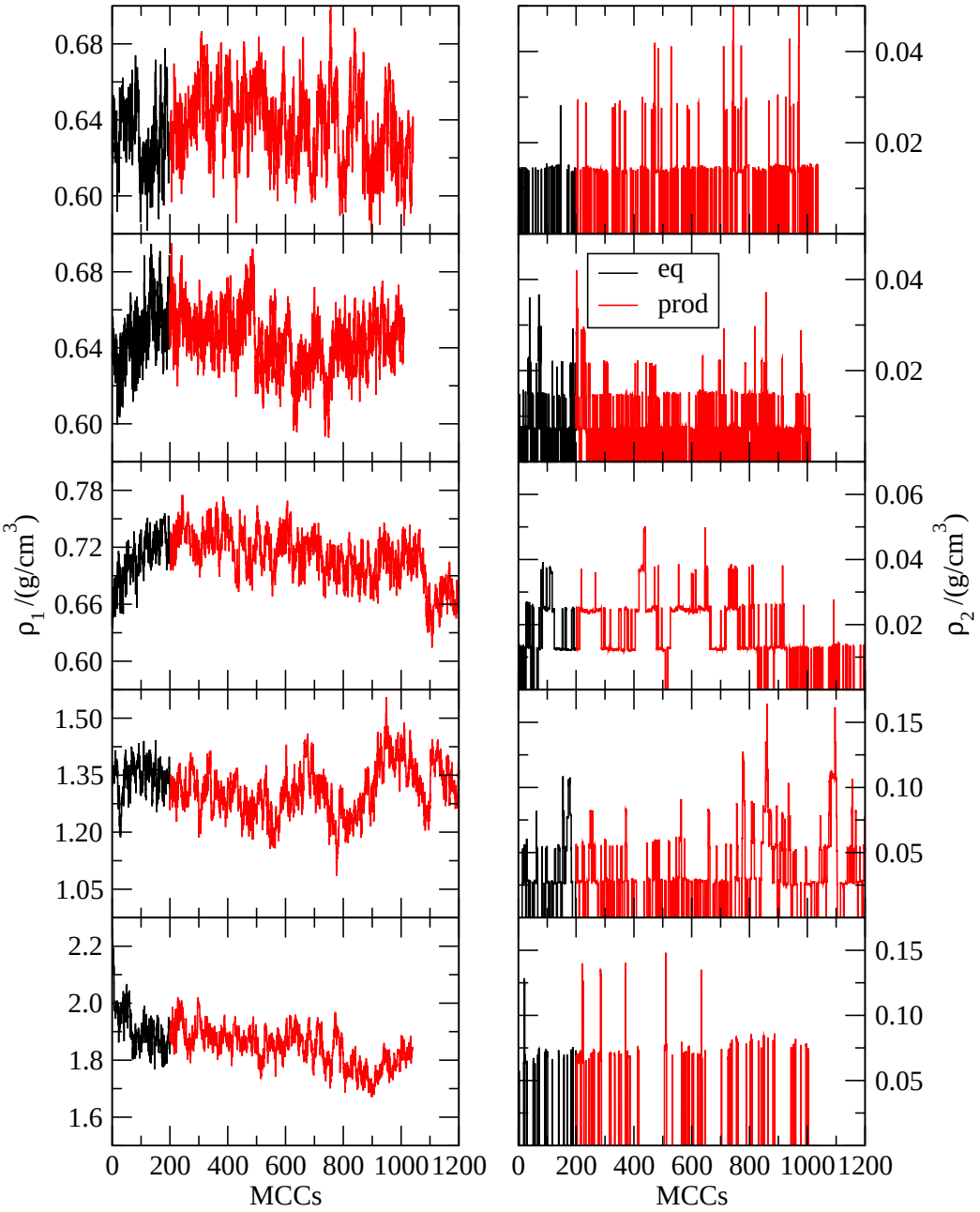


Figure SD1. Specific densities of box 1 (left) and box 2 (right) as a function of Monte Carlo cycles at the lowest temperature for all five systems (from top to bottom: PH_3-PBE-48, PH_3-PBE-96, PH_3-BLYP-48, AsH_3-PBE-48, SbH_3-PBE-48). Trajectories parts colored in black and red represent the equilibration and production periods, respectively.

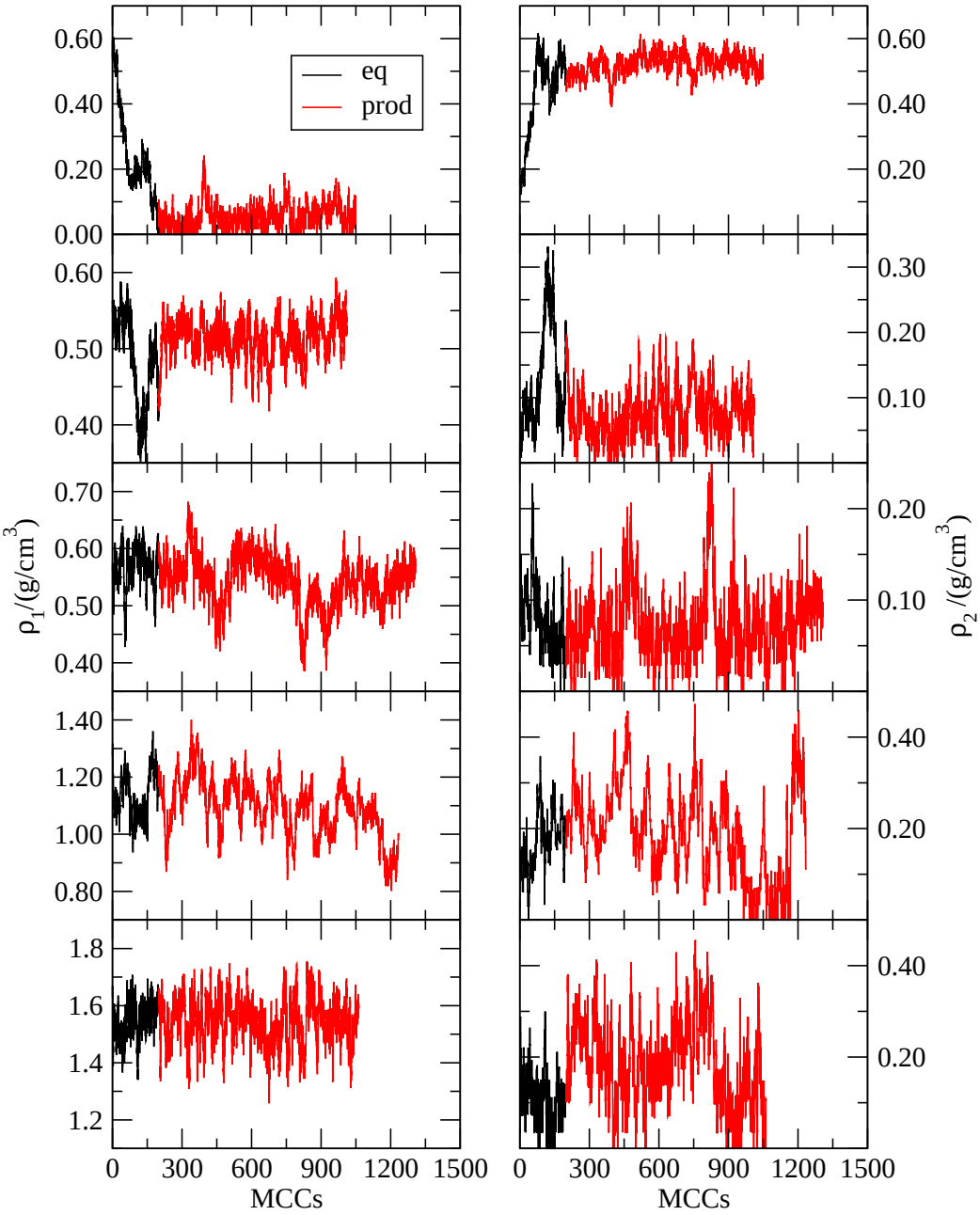


Figure SD2. Specific densities of box 1 (left) and box 2 (right) as a function of Monte Carlo cycles at the highest temperature for all five systems (from top to bottom: PH_3-PBE-48, PH_3-PBE-96, PH_3-BLYP-48, AsH_3-PBE-48, SbH_3-PBE-48). Trajectories parts colored in black and red represent the equilibration and production periods, respectively.

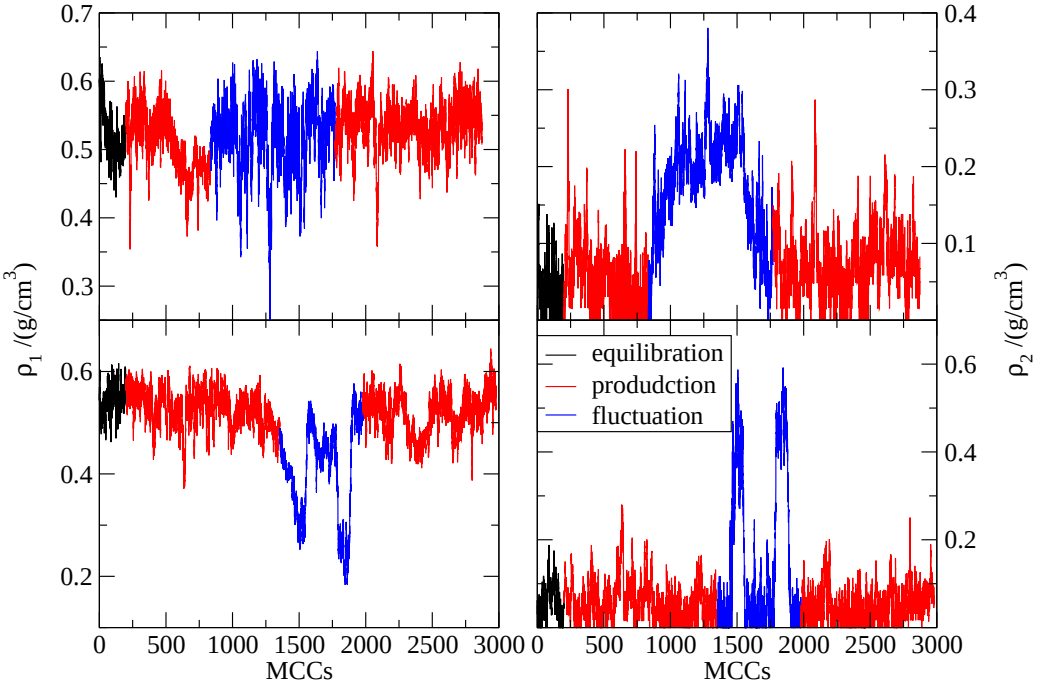


Figure SD3. Specific densities of box 1 (left) and box 2 (right) as a function of MC cycles at $T = 290$ K for PH_3-PBE-48 system (top: run 1; bottom: run 2). Trajectory parts colored in black, blue, and red represent the equilibration period, periods of large density fluctuations not included in the averages, and production periods, respectively.

Description of force fields. A simple harmonic potential was used for both X-H bond stretching and H-X-H angle bending

$$U_{\text{bond}} = \frac{1}{2}K_R(R - R_0)^2 \tag{5.1}$$

$$U_{\text{angle}} = \frac{1}{2}K_\theta(\theta - \theta_0)^2 \tag{5.2}$$

where the prefactors K_R and K_θ, equilibrium bond length R_0, and equilibrium angle θ_0 were obtained from potential energy scan in Gaussian [226] with the PBE functional, and are listed in Table SD2.

The intermolecular interactions are described by a combination of Lennard-Jones and Coulomb potentials

$$U_{\text{inter}} = 4\epsilon_{ij}[(\frac{\sigma_{ij}}{r_{ij}})^{12} - (\frac{\sigma_{ij}}{r_{ij}})^{6}] + \frac{q_i q_j}{4\epsilon_0 r_{ij}} \tag{5.3}$$

where q_i, q_j, r_{ij}, and ϵ_0 are the charges on interaction site i and j, distance between the two interaction sites, and permittivity of vacuum, respectively. Charges for atoms were obtained using the Charge Model 5 developed by Marenich *et al* [227]. Lorentz-Berthelot combining rules were used to obtain cross interaction parameters ϵ_{ij} and σ_{ij} between unlike interaction sites. ϵ_{ii} and σ_{jj} were obtained through minimizing the energy difference squared between first-principles and force-field based calculations. More specifically, a trajectory of frames were generated using first-principles molecular dynamics simulations, and used as input for running CP2K in REFTRAJ mode to obtain the corresponding energies described by empirical force fields. Then the mean squared errors for energy between these two methods were minimized by varying the force field parameters. The resulting parameters for intermolecular interactions are listed in Table SD3. The error is found to be most sensitive to σ_{HH} and the resulting mean squared errors of around around 1.5×10^{-4} hartree^2 remained relatively high. Thus, in this case the force fields mostly function to avoid the generation of Monte Carlo trial configurations that would yield to very high energies and may be problematic for convergence of KS-DFT calculations.

Table SD2. Intramolecular force field parameters.

molecule	bond		angle	
	K	R_0	K	θ_0
	hartree/bohr^2	bohr	hartree/rad^2	rad
PH_3	0.1878	2.7446	0.1534	1.6249
AsH_3	0.1658	2.9217	0.1518	1.6005
SbH_3	0.1312	3.2609	0.1459	1.5880

Table SD3. Intermolecular force field parameters.

molecule	atom	charge	ϵ/k_B	σ
		e	K	Å
PH_3	P	−0.2376	125	3.6
	H	0.0792	50	2.5
AsH_3	As	−0.15	120	3
	H	0.05	5	1.8
SbH_3	Sb	−0.0498	200	4
	H	0.0166	7	1.8

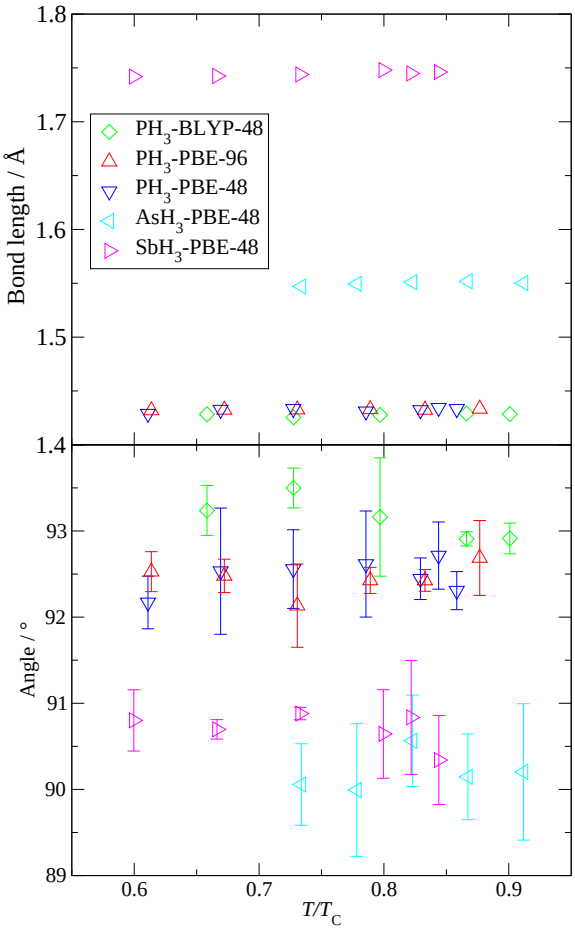


Figure SD4. Ensemble averages for X-H bond lengths (top) and H-X-H bending angles (bottom) calculated for the saturated liquid phase for PH_3, AsH_3-PBE-48, and SbH_3-PBE-48 systems as a function of reduced temperature.