MODELING THE ANOMALOUS BEHAVIORS OF SUPERCOOLED TETRAHEDRAL LIQUIDS

by

Jingxiang Guo

A dissertation submitted to the Department of Chemical and Biomolecular Engineering, Cullen College of Engineering in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in Chemical Engineering

Chair of Committee: Dr. Jeremy C. Palmer Committee Member: Dr. Jacinta C. Conrad Committee Member: Dr. Jeffrey D. Rimer Committee Member: Dr. Yashashree Kulkarni Committee Member: Dr. Pradeep Sharma

> University of Houston December 2020

Copyright 2020, Jingxiang Guo

Acknowledgements

This thesis is dedicated to my family, who supported me both mentally and financially throughout my undergraduate and Ph.D. studies in the U.S. Their patience for my long-term absence from home have made it possible for me to accomplish this fulfilling journey.

I also want to thank my advisor Professor Jeremy C. Palmer, who provided me with many challenging research projects from which I have learned so much both scientifically and technologically. His guidance and mentorship have helped me think critically about the research, and become a better problem solver.

Nowadays, scientific research has rarely been a personal endeavor. Throughout my Ph.D., I receive tremendous help from many current memebers (Ryan Roberts, Ankur Agarwal, Deepak Mangal, Lakshmanji Verma and Shivraj Kotkar) and former members (Renjie Chen and R. John Clark). They are always there for answering my technical and conceptual questions. I want to thank all of them.

Abstract

Tetrahedral liquids such as water and silica are among the most ubiquitous and important substances in our world, yet the origin of their well-known structural, thermophysical, and dynamic anomalies still remains elusive. It has been posited that anomalous behaviors of tetrahedral liquids under ambient and supercooled conditions are due to critical fluctuations associated with a low-temperature liquidliquid critical point (LLCP), below which two distinct metastable liquids undergo a first-order liquid-liquid phase transition (LLPT). Due to rapid homogeneous nucleation of the crystalline phase, experimental probing techniques have not yet been able to verify or falsify this hypothesis. Computational studies of molecular fluids demonstrate metastable LLPTs are possible, but they have not resolved the outstanding question of how such behavior can be characterized experimentally. We present results from large-scale molecular dynamics (MD) simulations of more than 100,000 molecules of two model tetrahedral liquids that exhibit LLPTs: the ST2 model of water and an ionic model of liquid silica. The simulations reveal that both models exhibit anomalous scattering, reminiscent of that observed in experiment, which is characterized by an increase in the static structure factor at low wavenumbers. This unusual behavior is linked with coupled fluctuations in density and local tetrahedral order in the liquid. The Ornstein-Zernike correlation length estimated from the anomalous scattering component exhibits power-law growth upon cooling, consistent with the existence of a liquid-liquid critical point in both models. Further, spontaneous liquid-liquid phase separation is observed in each model upon thermally quenching large systems into the two-phase regions. Lastly, we show that nucleation of the stable crystal phase from metastable tetrahedral liquids, a rare event, can be investigated by enhanced sampling methods based on the hybrid Monte Carlo (HMC) algorithm.

Table of Contents

Acknowledgements		iii	
A	bstra	ict	iv
Table of Contents			vi
Li	ist of	Tables	viii
Li	ist of	Figures	ix
1	Int	roduction	1
	1.1	Water is ubiquitous yet unusual	1
	1.2	The Liquid-Liquid Phase Transition (LLPT) hypothesis	3
	1.3	Water-like anomalies in silica	5
	1.4	Homogeneous ice nucleation	9
2	And	omalous scattering in supercooled ST2 water	14
	2.1	Introduction	14
	2.2	Methods	17
		2.2.1 Simulation protocol	17
		2.2.2 Structural analysis	18
	2.3	Results and Discussion	21
	2.4	Conclusions	36
3	Flu	ctuations near the liquid–liquid transition in a model of silica	38
	3.1	Introduction	38

4	Hył	brid Monte Carlo with LAMMPS	55
	4.1	Introduction	55
	4.2	Hybrid Monte Carlo	57
		4.2.1 Algorithm	57
		4.2.2 Implementation notes	58
		4.2.3 Model and sampling protocol	60
		4.2.4 Validation	61
	4.3	Free Energy Calculations	67
		4.3.1 Biased sampling	67
		4.3.2 Nucleation of mW water	69
	4.4	Conclusions	74
5	Cor	nclusions and Future Work	77
	5.1	Conclusions	77
	5.2	Future Work	79
		5.2.1 Future Work 1	79
		5.2.2 Future Work 2	80
		5.2.3 Future Work 3	82
Re	efere	ences	84

List of Tables

1	Results of ensemble consistency tests	 67

List of Figures

1.1	Schematic isobaric temperature dependence of thermodynamic re-	
	sponse functions of water in comparisons with those of typical liquid.	
	(a), (b) and (c) are temperature dependence of isothermal compressbil-	
	ity, κ_T , isobaric heat capacity, C_P and thermal expansion coefficient,	
	α_P , respectively	2
1.2	Water's complex noncrystalline phase diagram with equilibrium and	
	non-equilibrium states. 11 The red dot " C_2 " is the hypothetical location	
	of liquid-liquid critical point. 12	3
1.3	Schematic molecular representation of ST2 water model	5
1.4	Both liquid silica and D_2O exhibit the density maximum. 47	7
1.5	A schematic of rigid-ion model of silica. The number ratio of the silicon	
	atom to the oxygen atom is 1:2. Each Si atom carries +4 charges and	
	each oxygen atom carries -2 charges	8
1.6	Free energy $F(\rho)$ of mWAC model at a range of supercritical and sub-	
	critical temperature and pressure conditions. 55	9
1.7	A schematic diagram of CNT showing the free energy dependence of	
	the nucleus size assuming the shape of nucleus is spherical	10
91	Probability density distributions for the tetrahedral order parameter a	
2.1	a computed at different temperatures for a system containing $N = 22000$	
	computed at different temperatures for a system containing $N = 32000$	
	ST2 water molecules. The vertical dashed lines indicate the median	
	values of q at the highest and lowest temperatures. \ldots \ldots \ldots	22

2.2	(a) Total structure factors and (b) total pair correlation functions com-	
	puted at different temperatures for a system containing $N=32000~{ m ST2}$	
	water molecules. The inset in (b) shows an enlargement of the total	
	pair correlation function for $r \ge 2$ nm	23
2.3	Partial pair correlation functions for low- q and high- q species com-	
	puted at different temperatures for a system containing $N=32000~{ m ST2}$	
	water molecules.	25
2.4	Bhatia-Thornton structure factors $S_{\mathrm{CC}}(Q)$ (Eq. 2.3) and $S_{\mathrm{NC}}(Q)$ (Eq.	
	2.4) computed at different temperatures for a system containing $N =$	
	32000 ST2 water molecules. The structure factors $S_{\rm NN}(Q) = S(Q)$ (Eq.	
	2.2) are shown in Fig. 2.2(a)	27
2.5	Ornstein-Zernike plots for (a) $S^{A}(Q)$ and (b) $S_{CC}(Q)$. Points are data	
	from MD simulations of $N = 32000$ ST2 water molecules, whereas lines	
	in (a) and (b) are fits using Eqs. 2.13 and 2.14, respectively. \ldots	30
2.6	Estimates of the correlation length ξ calculated from $S^{\mathrm{A}}(Q)$ (circles)	
	and $S_{\rm CC}(Q)$ (squares) for a system containing $N=32000~{ m ST2}$ water	
	molecules. Lines are power-law fits using Eq. 2.9	32
2.7	(top) Density profile $\rho(z)$ at 237 K computed along the major axis	
	(z-axis) of a rectangular simulation cell containing $N = 256000$ ST2	
	molecules after the system was thermally quenched into the region of	
	liquid-liquid coexistence. (bottom) The HDL-LDL interfaces are ob-	
	served in renderings of the system in which the particles are shaded	
	according to the local density across the z -axis of the simulation cell	34
3.1	(a) Structure factor and (b) radial distribution function computed for a	
	system of $N = 36,000$ atoms. The inset in (a) shows $S(0)$ as a function	
	of <i>T</i> , and the inset in (b) shows an enlargement of $r^2[g(r)-1]$ for $r \ge 2.5$	

.

nm.

3.2	(a) Probability density distribution for the tetrahedral order parameter	
	q. The median values of q at the highest (3800 K) and lowest (3375	
	K) temperatures are indicated by vertical dashed lines. (b) Bhatia-	
	Thornton structure factors $S_{ m CC}(Q)$ (Eq. 3.6) and $S_{ m NC}(Q)$ (Eq. 3.7). (c)	
	Simulation snapshots at 3800 K and 3375 K showing the H (red) and L $$	
	(blue) species of Si atoms. Data in (a) – (c) are for a system containing	
	N = 36,000 atoms	47
3.3	Ornstein-Zernike plots for (a) $S_{\rm CC}(Q)$ and (b) $S^{\rm A}(Q)$. Points are data	
	from MD simulations of $N = 36,000$ atoms. Solid lines in (a) and (b)	
	are fits to the Lorentzian functions given by Eqs. 3.12 and 3.13, respec-	
	tively.	49
3.4	Ornstein-Zernike correlation length ξ estimated from $S_{ ext{CC}}(Q)$ (squares)	
	and $S^{A}(Q)$ (circles) for a system of $N = 36,000$ atoms. Lines are power-	
	law fits using Eq. 3.14	50
3.5	Density profiles at $T_{\rm low}$ = 3000 K computed along the major axes	
	(z-axes) of rectangular simulation cells containing (a) $N = 4500$, (b)	
	36,000, and (c) 432,000 atoms. The abscissa in each panel is normal-	
	ized by the length of the simulation cell, L_z . (d) Liquid-liquid interfaces	
	are observed in snapshots of the largest system ($N = 432,000$) in which	
	Si atoms are colored according to the local density profile in (c)	53
4.1	Equation of state data for liquid argon calculated using HMC and MD	
	(filled and open symbols, respectively). Excellent agreement between	
	HMC and MD is observed at all conditions examined. Statistical un-	

Chapter 1: Introduction

1.1 Water is ubiquitous yet unusual

Water or H_2O is essential to the existence and prosperity of lives on earth. It has tremendous impact on agriculture, manufacturing, pharmaceuticals, and our lifestyle in general.¹ Despite its ubiquitous presence in our daily lives and its seemingly simple molecular structure, water poses many unsolved scientific puzzles that still baffle scientists today.^{2–4} Ice floats on water, for example, which conflicts with our life experience that solid is typically denser than liquid. Water becomes more fluid rather than viscous when compressed below the room temperature. The density of a simple liquid exhibits monotonic temperature dependence while that of water reaches density maximum at 4 °C at room pressure. These examples are just the tip of the iceberg of water's anomalous thermophysical and dynamic properties. In light of water's many odd behaviors, it is natural to ask: what is the origin of such a broad spectrum of anomalies ?

The effort to understand physical properties of supercooled water has been going on for many decades.^{5,6} In the 1970s, Angell and his coworkers performed experiment to probe supercooled water down to -26 °C.⁷ In their work, the temperature dependence of isothermal compressibility κ_T can be well characterized by a criticallike power-law relation with extrapolated divergence at a singular temperature T_s = 228 K. This temperature is inaccessible by experimental probing techniques due to rapid homogeneous ice nucleation. The isobaric heat capacity C_P of supercooled water also becomes increasingly enhanced and can be described by a power-law relation.^{8,9} In Fig.1.1(a) and (b), the isothermal compressbility and the isobaric heat capacity of water exhibit non-monotonic temperature dependencies. After passing



(c) α_T vs. T

Figure 1.1: Schematic isobaric temperature dependence of thermodynamic response functions of water in comparisons with those of typical liquid. (a), (b) and (c) are temperature dependence of isothermal compressibility, κ_T , isobaric heat capacity, C_P and thermal expansion coefficient, α_P , respectively

through a minimum, their magnitudes become anomalously enhanced as temperature decreases under ambient pressure. In contrast, the thermodynamic response functions of a simple typical liquid vary monoatonically with respect to temperature. These pioneering works not only provide early experimental evidence of water's thermodynamic anomalies but also inspire many physical interpretations to elucidate water's low temperature anomalous behaviors.¹⁰



1.2 The Liquid-Liquid Phase Transition (LLPT) hypothesis

Figure 1.2: Water's complex noncrystalline phase diagram with equilibrium and nonequilibrium states.¹¹ The red dot " C_2 " is the hypothetical location of liquid-liquid critical point.¹²

At this time, there does not yet exist a thermodynamically consistent scenario that could account for both thermodynamic and transport properties of water.^{10,12–14} One scenario that is supported by many experimental and simulation studies suggests that water exhibits a low-temperature liquid-liquid critical point (LLCP), in addition to its normal gas-liquid critical point at higher temperatures. This hypothesis was proposed based on a series of molecular dynamics studies of an atomistic water model at supercooled conditions.^{12,15} In this scenario, water's anomalous thermophysical properties at ambient and increasingly supercooled conditions can be attributed to critical fluctuations associated with the LLCP terminating a liquidliquid phase transition (LLPT) line. Fig.1.2 shows the hypothetical location of this critical point as " C_2 " in the middle of phase diagram.¹¹ Below this critical point, supercooled water undergoes a first-order phase transition between a low density liquid (LDL) and a high density liquid (HDL). The LDL is characterized by the local tetrahedral structure where a water molecule is four-coordinated with its nearest neighbors within the first primary coordination shell. The HDL, on the other the hand, is less tetrahedral due to the intrusion of a fifth neighbor into the first shell, resulting in higher density, energy and entropy than those of the LDL. Above the LLCP, in the supercritical region, the Widom line emanates as a continuation of first-order LLPT line. By definition, Widom line is the locus of maximum of correlation length.¹⁶⁻¹⁹ Crossing this line, the system continuously fluctuates between a LDL-like and HDLlike state.²⁰ In the vicinity of the LLCP, the loci of thermodynamic response functions κ_T and C_P maxima asymptotically converge onto the Widom line. According to this hypothesis, the experimentally observed high-density and low-density amorphous glasses (HDA and LDA) are considered as kinetically arrested forms of HDL and LDL, respectively.²¹ Although this intriguing hypothesis was advanced based on the evidence from computer simulations of a molecular model of water, only the experimental evidence can eventually prove or disprove this scenario. If the LLPT hypothesis could be verified experimentally, then, it would provide a unified interpretation of water's metastable and stable phase behaviors.

Until now, no conclusive evidence of LLPT in water has been reported yet in experimental studies of supercooled water.^{22,23} While current probing techniques and methods are still in development toward breaching the temperature limit imposed by spontaneous homogeneous ice nucleation, computer simulations have played a guiding role in understanding water's low temperature anomalies. After the seminal work by Poole et al.,¹² many computer models of water have been scrutinized for the possible existence of LLCP.^{24–31} For now, only the ST2 water model has been



Figure 1.3: Schematic molecular representation of ST2 water model. confirmed to exhibit an LLPT through rigorous free energy calculations.^{28,32–35}

The ST2 water model is a rigid and five-site atomistic model of water.³⁶ Fig.1.3 shows the molecular structure of ST2 water model. Two positive charges are positioned on hydrogen atom and two negative charge are positioned to represent the oxygen's lone pair of electrons. The four point charges are equal in magnitude and are located at corners of a perfect tetrahedron. The oxygen atom is modeled as a soft sphere with excluded volume and van der Waals interactions. The unique characteristic of ST2 is its built-in tetrahedrality, which enables this water model to exhibit its anomalies at more accessible temperature.¹² As a result, the LLPT of ST2 water model can be characterized at relatively higher temperature where other water models suffer from slow equilibration at deeply supercooled condition.^{24,33} Although this water model fails to quantitatively reproduce the properties of real water (e.g., the temperature of maximum density (TMD) of ST2 water at room pressure is above 50 °C as opposed to 4 °C in real water³⁷), such deficiency seems to be an advantage that can be exploited to study the LLPT phenomena in *in silico*.³⁸⁻⁴²

1.3 Water-like anomalies in silica

Although the LLPT hypothesis still remains putative in the experimental studies of supercooled bulk water, it was investigated and verified in many molecular liquids other than water, such as silica, silicon, phosphorus, $Al_2O_3 - Y_2O_3$ and triphenyl phosphate.^{43–47} A very recent study by Henry et al. showed that a first-order LLPT and its associated critical point can also be found in liquid sulfur through X-ray diffraction and Raman scattering techniques.⁴⁸ A preponderance of evidence pointing to the existence of an LLPT in many single-component substances suggests that certain common physical attributes must play a role.¹⁸

Silica or silicon dioxide (SiO_2) is a technologically and scientifically important substance. It is ubiquitous in the form of rock and sand in the nature, and occupies a significant percentage of compositions of the earth's crust.⁴⁹ It is also known for its remarkable glass-forming ability, and is often used as an important ingredients for glass production.⁵⁰ Understanding and then controlling the phase behaviors of molten silica is of great scientific and industrial interest. Molton silica exhibits many peculiar properties reminiscent of those found in water.^{47,51} Just like ice, silica can exist in a diverse forms of crystalline phases under various temperature and pressure conditions. ^{52,53} Both β -cristobalite and Ice I_c share the same diamond unit cell structure, and similarly, β -tridymite is the structural analog of Ice I_h .⁵⁴ It's not surprising that parallel studies on water and silica, two chemically distinct substances with the same local tetrahedral bonding arrangements, have uncovered that they exhibit similar anomalous behaviors. Angell et al. measured the temperature dependence of densities for liquid silicon dioxide, locating it's temperature of maximum density (TMD) at 1773 K.47 Fig.1.4 shows the temperature dependence of density of both silica and D₂O plotted together on a reduced scale for comparisons.⁴⁷ Their experimental results revealed the importance of averaged bridge-bond angles (O-H-O angle in D_2O and Si-O-Si angle in SiO_2) in determining the sharpness of density variation around the TMD, and the extent of anomalies in tetrahedral liquids.⁴⁷ Liquid silica also exhibits water-like dynamical anomalies in which the self-diffusion coefficient reaches a maximum at a given pressure.⁵¹

Computational studies of several silica models have been devoted to investigat-



Figure 1.4: Both liquid silica and D₂O exhibit the density maximum.⁴⁷

ing the origin of its water-like anomalies. In computer simulation, silica interactions are often treated by a rigid-ion model in which silicon and oxygen carry formal charges and interact through electrostatic interactions, but do not exhibit explicit bonding. The models developed by Woodcock, Angell, Cheeseman (WAC) and van Beest, Kramer, and van Santen (BKS) are examples of such ionic model of silica commonly used in simulation studies.^{56–58} Because of the simplified treatment of interactions such as the absence of covalent bonding and the lack of polarizability, none of these models are able to reproduce quantitatively the physical properties of real silica, but they do qualitatively reproduce many water-like anomalies such as temperature of maximum density.^{56,57,59} The search for signatures of LLCP in BKS and WAC models of silica have been performed by running extensive equation of state calculations for a range of temperature and pressure conditions.^{59,60} By using the crossing of isochores as an indicator of phase transition, the sign of critical point can be readily identified.⁶⁰ In an accessible range of conditions examined, BKS model does not show the tendency of isochores crossing, and therefore, no evidence of an LLPT, while the WAC model seems to be close to criticality in the glassy region.⁶⁰ Further studies by Lascaris proposed a simple way of tipping the WAC model toward a true criticality. By scaling the charge magnitude on both silicon and oxygen by a



Figure 1.5: A schematic of rigid-ion model of silica. The number ratio of the silicon atom to the oxygen atom is 1:2. Each Si atom carries +4 charges and each oxygen atom carries -2 charges

factor of 0.84 in the original mWAC model, an LLCP can be unambiguously located in the phase diagram where the extrema of thermodynamic response functions coincide.⁶¹ Fig.1.5 shows the schematic of WAC model of silica. A non-bonded pairwise interaction between Si^{+4} ion and O^{-2} ion in WAC model can be described by

$$U_{WAC}(r_{ij}) = f^2 \frac{1}{4\pi\varepsilon_0} \frac{z_i z_j e^2}{r_{ij}} + A_{ij} \exp(-B_{ij} r_{ij}), \qquad (1.1)$$

where r_{ij} is separation distance between ions, z_i is the charge number of either $z_{Si} = +4$ or $z_O = -2$, ε_0 is the permittivity of free space, and e is the elementary charge. The exponential term in Eq.1.1 describes the short-range and excluded volume interactions with $A_{SiSi} = 1.917991469 \times 10^5$ kJ/mol, $A_{SiO} = 1.751644217 \times 10^5$ kJ/mol, $A_{OO} = 1.023823519 \times 10^5$ kJ/mol, and $B_{ij} = 34.48$ nm⁻¹ for all ion pairs. When the scaling charge f = 1, the original WAC model is recovered. For f = 0.84, the modified version of WAC model (mWAC) is obtained.⁶¹

The unambiguous evidence of an LLPT in mWAC was further confirmed through free energy calculations using density as an order parameter.⁵⁵ Fig.1.6 shows free energy profiles of mWAC at conditions above and below the estimated critical point.⁵⁵



Figure 1.6: Free energy $F(\rho)$ of mWAC model at a range of supercritical and subcritical temperature and pressure conditions.⁵⁵

Above the critical temperature, there exists only one metastable liquid phase shown as single free energy minimum. Below the critical temperature, the model exhibits two free energy minima of equal depths, suggesting two metastable liquid phases differing in their densities coexist (see Fig.1.6).

1.4 Homogeneous ice nucleation

Thanks to the metastable nature of pure supercooled tetrahedral liquids, homogeneous nucleation of more stable crystalline phases will prevail on long time scales. As discussed in the previous section, the origin of puzzling properties of tetrahedral liquids still remain unresolved because of temporal limitations of current experimental measurement techniques. To further explore the "no man's land" in supercooled water, understanding the nucleation mechanism and computing the nucleation rate will be helpful in designing an experiment that can alleviate such temporal issues.¹⁶ However, for current experimental methods, probing crystal nucleation events at the molecular scale in real time is still challenging, and therefore, intrinsic mechanisms and kinetics of homogeneous nucleation processes are still elusive.⁶²

Nucleation is the first step of crystallization, and it plays a central role in de-



Figure 1.7: A schematic diagram of CNT showing the free energy dependence of the nucleus size assuming the shape of nucleus is spherical.

termining the outcome of a crystallization process.⁶³ There exists two major classes of nucleation mechanisms. In homogeneous nucleation, the formation of a stable embryo of the crystalline phase is triggered by thermal fluctuations, whereas in heterogeneous nucleation it is triggered by contact with an external surface. Although the dominant mechanism for generating the initial crystalline phase in many industrial and natural phenomena is mostly likely to be heterogeneous, homogeneous nucleation still plays a pivotal role in many disciplines of science and engineering.² For example, supercooled water can often be found in high-altitude cloud of atmosphere,⁶⁴ and cloud development is influenced by the homogeneous ice nucleation rate.⁶⁵

Classical nucleation theory (CNT), though initially proposed to account for the process of vapor condensation⁶⁶ or liquid-vapor transition, provides a qualitative description of thermodynamics and kinetics of crystal nucleation processes (see Fig.1.7). The theory is based a assumption that a vanishingly thin layer exists between the

supersaturated and the condensed phases, and both phases have the macroscopic properties of bulk phases respectively.⁶² If the nucleus is assumed to have a spherical shape, the thermodynamics of nucleation can be described by a free energy change due to the formation of crystalline phase from the supersatured phase

$$\Delta G(r) = \frac{4}{3}\pi r^3 \Delta \mu + 4\pi r^2 \gamma, \qquad (1.2)$$

where r is the raidus of spherical nucleus and γ is the interfacial free energy per unit area, and $\Delta \mu$ is the chemical potential difference per unit volume between bulk supersaturated and condensed phase. By taking the first derivative of ΔG with respect to the radius r and setting it equal to 0, we can obtain the maximum value of ΔG and the radius r^* at which the maximum occurs. Using this procedure, we find that the number of molecules in the critical nucleus is given by

$$n^* = \frac{32\pi\rho_c\gamma^3}{3\Delta\mu^3},\tag{1.3}$$

where the ρ_c is the density of crystalline phase. The Gibbs free energy barrier at this critical nucleus size is given by

$$\Delta G^* = \frac{16\pi\gamma^3}{3\Delta\mu^2}.\tag{1.4}$$

As the above equations indicates, interfacial free energy and chemical potential difference are factored into the equation cubically and quadratically, and therefore, the free energy barrier is very sensitive to any uncertainties in those values. An Arrhenius rate law can be applied to describe the kinetics of nucleation by the following equation

$$J = J_0 \exp\left(-\frac{\Delta G^*}{k_B T}\right). \tag{1.5}$$

According to this equation, the nucleation rate depdends expoentially on the Gibbs free energy barrier, which determines the probability of forming a critical nucleus at the top of barrier. At moderate supercooling, the free energy barrier ΔG^* is dominating term controlling the nucleation rate, while at high supercooling condition, the free energy barrier ΔG^* become diminished, and the expotential term approaches one. Under this condition, the kinetic prefactor J_0 , which is associated with the mobility of the system comes into play and becomes a dominating factor. Thus, the interplay of the kinetic factor J_0 and thermodynamic factor ΔG^* predicts a maximum in temperature dependence of the nucleation rate.⁶²

The limitations of CNT are also obvious. First, at deeply supercooled condition, the size of critical nucleus become very small, blurring the boundary between the supersaturated and condensed phase. The assumption of capillary approximation breaks down, and the macroscopic properties used for nucleus in the CNT may not be physically relevant.⁶⁷ Additionlly, CNT predicts a one-step transformation of supersaturated phase into a more stable condense phase. The metastable phase as an intermediate state during the nucleation is not predicted by CNT. However, in many experimental and simulations studies, such metastable phase plays a crucial role as suggested by Ostwald's step rule, which states that the solid phase that crystallizes initially from the liquid phase may not be the one that is most thermodynamically stable but the one that is closest in free energy to the liquid phase.⁶⁸

In comparison with the spatial and temporal limitations of experimental techniques, computer simulations provide appropriate length and time scale to investigate the physical phenomena such as nucleation. The first direct observations of homogeneous ice nucleation in computer simulations were reported by Matsumoto et al.⁶⁹ They performed long molecular dynamics simulations using TIP4P water model, demonstrating the possibility of observing ice nucleation in computer simulations. Because the ice nucleation is a rare event, and the time scale of its emergence in supercooled water can be easily beyond the reach of current computer computer simulations, 22,70 brute-force approach or direct simulations using Monte Carlo (MC) and molecular dynamics (MD) with atomistic water models is usually not computationally affordable to perform. To facilitate the ice nucleation process in computer simulations, either a coarse-grained water model can be used $^{26,71-73}$ or enhanced sampling methods can be adopted to overcome the time scale challenges associated with modeling ice nucleation. $^{74-77}$

The coarse-grained mW model of water has often been used in the study of homogeneous ice nucleation.^{26,68,75,76} This water model can quantitatively reproduce some physical properties of water such as the melting point at ambient condition with less computational cost compared with atomsitic model of water.⁷¹ The mW water model treats a water molecule as a coarse-grained bead interacting through two-body and three-body interactions.

The force-field of mW was derived from Stillinger-Weber potential which was initially proposed for silicon. All force-field parameters were kept the same as those for silicon except that those parameters governing the energy and length scales are fitted to the experimental properties of water.⁷¹ The resulting water model favors a local tetrahedral symmetry and depends on short-ranged interactions to mimic the hydrogen bonding effect. Without explicit hydrogen atoms, mW suffers from several shortcomings. It tends to overestimates the self-diffusion coefficient of real water at given pressure since the hydrogen reorientation is not taken into account.⁷¹ It also underestimates the isobaric heat capacity at low temperature due to the lack of rotational degree of freedom.⁷¹

Chapter 2: Anomalous scattering in supercooled ST2 water

This chapter was previously published by Taylor & Francis Group: Guo, J.; Singh,
R. S.; Palmer, J. C. Anomalous scattering in supercooled ST2 water. *Molecular Physics* 2018, 116, 1953–1964

2.1 Introduction

Water exhibits a number of anomalous thermophysical behaviours that are not observed in "simple" non-network-forming liquids.^{3,16,19} Whereas most liquids densify as temperature decreases, for example, liquid water expands as it is cooled below 4 °C at ambient pressure.³ The magnitudes of water's isothermal compressibility κ_T , thermal expansivity α_P , and isobaric heat capacity C_P also begin to increase rapidly as the liquid is cooled near its freezing point; the rate of increase continues to accelerate with decreasing temperature as far as can be measured in the supercooled regime.^{3,7-9,78,79}

The unusual properties of water are thought to result from fluctuations that drive the formation of transient low-density and high-density domains in the liquid.^{3,16,17} The low-density domains are associated with open tetrahedral coordination structures that are stabilised by water's hydrogen bonding network. This local tetrahedral order is partially disrupted in the high-density domains, however, by the entropically favoured intrusion of additional molecules into water's primary coordination shell. The competition between these two locally favoured coordination structures thus gives rise to mixture-like behaviour that is uncommon in pure liquids. Indeed, thermodynamic models that treat water as a binary mixture of two partially miscible liquids have been remarkably successful in describing its unusual properties.^{80,81} It has been hypothesised that water's mixture-like behaviour may even culminate in a first-order phase transition between a high-density liquid and a low-density liquid (HDL and LDL, respectively) under deeply supercooled conditions.^{3,12,16,17} The apparent divergence of κ_T , α_P , and C_P upon cooling is thus explained in this scenario by the presence of a low-temperature critical point that terminates the metastable HDL–LDL coexistence line. This theory also posits that water's experimentally observed high-density and low-density amorphous glasses (HDA and LDA)^{11,82–84} are kinetically arrested manifestations of the two liquids. Despite the intriguing nature of the liquid-liquid critical point (LLCP) hypothesis, however, it has proved to be extremely challenging to falsify experimentally. Extrapolation of water's thermodynamic response functions indicates that the hypothesised LLCP lies below the homogeneous nucleation temperature for ice ($T_H \approx 232$ K at ambient pressure^{6,7}). Hence, the LLCP is predicted to occur in a region of water's phase diagram where direct measurement of liquid state properties is frustrated by rapid ice nucleation.^{3,16,17,74}

Numerous computational studies have also searched for a liquid-liquid phase transition (LLPT) in molecular models of water.^{24,27,29,38,40,68,85–87} Unfortunately, the low-temperature phase behaviour of many models remains unclear because sampling under supercooled conditions is often hampered by slow relaxation processes in the liquid. The few exceptions include the coarse-grained mW water model⁷¹ and the rigid, point-charge ST2 model of Stillinger and Rahman.³⁶ Although mW does not exhibit an LLPT,^{26,88} clear evidence of such a transition has been reported for ST2.^{12,28,30,32–35,39,42} Free energy calculations performed with enhanced sampling methods show that two liquids (HDL and LDL) can be brought into phase coexistence in ST2 under deeply supercooled conditions.^{28,30,32–35,42} At coexistence, both liquids are metastable with respect to cubic ice, which can nucleate from the LDL phase.^{28,42} Recent free energy calculations show, however, that the thermodynamic stability of ST2's LLPT with respect to crystallization can be systematically tuned

by incorporating bond-angle flexibility into the model.⁴² The metastable LLPT in ST2 can therefore be traced in a continuous fashion, by adjusting a single model parameter, to a thermodynamically stable transition that can be observed without the possibility of interference from ice nucleation.⁴²

Whereas free energy methods can be used to directly detect LLPTs in molecular models, obtaining unambiguous experimental evidence to falsify the existence of such a transition in water is significantly more challenging. Recent measurements on micron-sized droplets using femtosecond x-ray laser pulses have been successful, however, in probing the structure of liquid water down to 227 K at ambient pressure.^{22,23} These investigations mark the first successful attempts to characterise the properties of liquid water below its homogeneous nucleation temperature. The x-ray scattering measurements suggest that water's LLCP, if it exists, may lie at colder conditions^{23,89} that are even more difficult to access experimentally. Nevertheless, they reveal that κ_T exhibits a maximum at 229 K,²² which is consistent with crossing the Widom line that is predicted to extend from the LLCP into the supercritical region.^{16,17}

Motivated by these recent experiments, here we perform large-scale molecular dynamics (MD) simulations of systems containing up to 256000 molecules to investigate the small-angle scattering behaviour of supercooled ST2 water. We show that ST2 exhibits anomalous scattering, reminiscent of that observed in experiment, ^{22,90} which is characterised by an increase in the static structure factor at low wavenumbers. The anomalous scattering becomes significantly more pronounced at low temperature, indicating that fluctuations in the liquid increase rapidly upon cooling.

We scrutinise this unusual behaviour by analysing fluctuations in density and local tetrahedral order in ST2. Specifically, we apply a variant of the conventional Bhatia-Thornton analysis⁹¹ in which ST2 water is treated as a binary mixture of molecular species with low and high tetrahedral order in their primary coordina-

tion shell, respectively.^{92,93} This analysis reveals that "concentration" fluctuations in these species, and hence spatial variations in local tetrahedral order, contribute significantly to ST2's anomalous behaviour. We also estimate the static correlation lengths associated with concentration fluctuations and the anomalous scattering component. As expected, both correlation lengths exhibit power-law divergence as temperature is decreased, reflecting "critical opalescence" from long-wavelength fluctuations in the liquid.⁹⁴ The critical temperature $T_c \approx 245$ K estimated from the correlation lengths is in excellent agreement with values for ST2's LLCP reported in the literature.^{38,39} Finally, we show that spontaneous liquid-liquid phase separation is observed when a large system containing 256000 molecules is quenched below $T_{\rm c}$ into the two-phase region. Our large-scale MD simulations therefore confirm the existence of an LLPT in ST2 and support findings from previous computational studies performed using significantly smaller systems containing only a few hundred molecules.^{28,30,32–35,42} We anticipate that the analysis presented here may prove useful in interpreting scattering experiments on supercooled water and trends observed in computational studies of other models.

2.2 Methods

2.2.1 Simulation protocol

Canonical (*NVT*) ensemble molecular dynamics simulations of a system containing N = 32000 water molecules were performed using a widely-studied variant of the ST2 model in which the reaction field method is used to account for long-range electrostatic interactions. Potential parameters and other details of our implementation of this ST2 variant are identical to those reported in Refs. 38, 39 and 30. The simulations were performed using a cubic simulation cell with an edge length of $L \approx 10$ nm and a fixed bulk density of 0.94 g cm⁻³, which is close to the estimated critical isochore ($\approx 0.94-0.96$ g cm⁻³) for the ST2 model.^{38,86} The equations of motion were integrated using the velocity-Verlet algorithm with a 2 fs time step, and a Nosé-Hoover thermostat^{95,96} with a 0.2 ps time constant was applied to maintain the system's temperature. The trajectories were propagated for durations ranging from 150 to 300 ns, depending on the temperature. At the highest temperature studied (277 K), the system was equilibrated for 25 ns, followed by a production phase of 125 ns. At the lowest temperature examined (247 K), the equilibration and production phases were extended to 50 and 250 ns, respectively. The duration of the initial equilibration period was chosen to be at least an order of magnitude longer than characteristic relaxation times associated with the system's configurational energy and global bond-orientational order;⁸⁸ equilibration during this period was also carefully assessed by monitoring for drift in these quantities. Configurations were saved to file every 5 ps during the production phase of each simulation for further analysis.

2.2.2 Structural analysis

The simulation trajectories were used to perform Bhatia-Thornton mixture analysis^{91,92,94} and characterise fluctuations in density and local tetrahedral order in the liquid as ST2's LLCP was approached from above by cooling the system. The implementation and description of the Bhatia-Thornton approach presented here closely follow Ref. 92, where it was applied to study the behaviour of the TIP4P/2005 water model. In contrast with ST2, however, which exhibits a well-characterised LLPT, ^{28,42} clear numerical evidence of such a transition has not been reported in TIP4P/2005.⁸⁷ Examination of ST2's behaviour is therefore expected to provide insight into the coupling between fluctuations in density and local tetrahedral order in the vicinity of an LLPT.

We treat ST2 water as a binary mixture of species with low and high tetrahedral symmetry in their first coordination shell, respectively. The extent of local tetrahedral order was quantified for each molecule i in the system using 92,97

$$q_{i} = 1 - \frac{3}{8} \sum_{j=1}^{3} \sum_{l=j+1}^{4} \left(\cos \phi_{jl} + \frac{1}{3} \right)^{2}, \qquad (2.1)$$

where ϕ_{jl} is the angle formed by the lines joining molecule *i*'s oxygen atom and those on its *j*th and *l*th nearest neighbours (≤ 4). Molecules with q_i larger or smaller than a threshold value of \tilde{q} were classified as low-*q* (L) and high-*q* (H) species, respectively. We have found that the analysis is not particularly sensitive to the choice of \tilde{q} and thus used the median value of *q* for each configuration.⁹² This procedure resulted in an equimolar mixture of L and H species and allowed the classification of each molecule to change as its local environment evolved in time.⁹²

We analysed density (NN), concentration (CC), and coupled density-concentration (NC) fluctuations in ST2 upon cooling using the Bhatia-Thornton structure factors for binary mixtures.^{91,92} For the present case in which water is treated as an isotropic binary mixture of L and H species, these structure factors are given by⁹²

$$S_{\rm NN}(Q) = S(Q)$$
$$= x_{\rm H}S_{\rm HH}(Q) + x_{\rm L}S_{\rm LL}(Q) + 2\sqrt{x_{\rm H}x_{\rm L}}S_{\rm HL}(Q), \qquad (2.2)$$

$$S_{\rm CC}(Q) = x_{\rm H} x_{\rm L} [x_{\rm L} S_{\rm HH}(Q) + x_{\rm H} S_{\rm LL}(Q) - 2\sqrt{x_{\rm H} x_{\rm L}} S_{\rm HL}(Q)], \qquad (2.3)$$

and

$$S_{\rm NC}(Q) = x_{\rm H} x_{\rm L} [S_{\rm HH}(Q) - S_{\rm LL}(Q) + \frac{x_{\rm L} - x_{\rm H}}{\sqrt{x_{\rm H} x_{\rm L}}} S_{\rm HL}(Q)], \qquad (2.4)$$

where $Q = |\mathbf{Q}|$ is the norm of the scattering vector and $x_{\{\mathrm{H,L}\}}$ are the species mole fractions ($x_{\{\mathrm{H,L}\}} = 0.5$ in this case). Here, S(Q) is the total structure factor

$$S(Q) = N^{-1} \sum_{j,l=1}^{N} \langle \exp\left[i\mathbf{Q} \cdot (\mathbf{r}_l - \mathbf{r}_j)\right] \rangle, \qquad (2.5)$$

where $\{\mathbf{r}_i\}_{i=1}^N$ are the positions of the oxygen atoms, and $\langle ... \rangle$ denotes an ensemble average. For an isotropic system, Eq. 2.5 can also be evaluated by taking the Fourier transform of the radial distribution function, g(r):

$$S(Q) = 1 + 4\pi\rho \int_0^\infty r^2 \frac{\sin(Qr)}{Qr} h(r) dr,$$
 (2.6)

where h(r) = g(r) - 1 is the total pair correlation function, $\rho = N/V$, and V is the volume of the system. The Bhatia-Thornton expressions also depend on the Ashcroft-Langreth partial structure factor⁹⁸ for species α and β

$$S_{\alpha\beta}(Q) = \delta_{\alpha\beta} + 4\pi\rho \sqrt{x_{\alpha}x_{\beta}} \int_{0}^{\infty} r^{2} \frac{\sin(Qr)}{Qr} h_{\alpha\beta}(r) dr, \qquad (2.7)$$

where $\delta_{\alpha\beta}$ is the Kronecker delta function, $h_{\alpha\beta}(r) = g_{\alpha\beta}(r) - 1$, and $g_{\alpha\beta}(r)$ is the partial radial distribution function.

Although $S_{\rm NN}(Q)$ is generally not equal to S(Q) for binary mixtures, they are equivalent in this case because the scattering lengths of both species are identical. The functions S(Q) and $S_{\alpha\beta}(Q)$ can be evaluated directly using Eq. 2.5 and an analogous expression for the partial structure factor or by computing the Fourier transform of the radial distribution functions numerically, as suggested by Eqs. 2.6 and 2.7. The former approach is significantly more computationally demanding, whereas the latter approach can introduce artifacts associated with terminating the integration at a finite value $r = r_{\rm max}$, which must be less than or equal to L/2 to adhere to the minimum image convention. Discrepancies between the two approaches can be minimised by applying finite-size corrections,⁹⁹ which allow the total and partial structure factors to be accurately estimated at low wavenumbers. The assumption underlying both approaches and the finite-size correction scheme, however, is that the simulated system is large enough to capture the longest correlations in the liquid.^{93,99} As a result, significant care must be taken when applying these methods to analyse fluctuations in the vicinity of a critical point.

For the system sizes and thermodynamic conditions examined in this study, we found that the two approaches for evaluating the structure factors produced very similar results when finite-size corrections were applied. We therefore report results obtained using the Fourier transform method (Eqs. 2.6 and 2.7) with finite-size corrections. As discussed in the following section, we have explicitly checked for finite-size effects at the lowest temperature examined (247 K), where the correlation length is expected to be longest, by simulating a significantly larger system containing N = 256000 molecules. The linear dimension of this system is $L \approx 20$ nm and thus it can accommodate fluctuations twice as large. The correlation lengths estimated from the structure factors computed for both systems are in statistical agreement, indicating that results presented for N = 32000 molecules are not significantly affected by finite-size effects.

2.3 Results and Discussion

We use the parameter q (Eq. 2.1) to analyse fluctuations in local tetrahedral order in supercooled ST2 water. The probability density distributions of q calculated from the simulation trajectories exhibit a strong high-q peak around 0.9, which is close to the limiting value q = 1 expected for a regular tetrahedron (Fig. 2.1). A second much weaker peak at lower q is also observed in the distributions near 0.5 (Fig. 2.1). The high-q peak becomes more pronounced upon cooling, whereas the magnitude of the low-q peak decreases as T is lowered. This behaviour has been reported in numerous computational studies, ^{92,93,100} and it signifies an increase in tetrahedral order, which



Figure 2.1: Probability density distributions for the tetrahedral order parameter q computed at different temperatures for a system containing N = 32000 ST2 water molecules. The vertical dashed lines indicate the median values of q at the highest and lowest temperatures.

is consistent with the interpretation of x-ray scattering experiments on supercooled water.^{89,101}

Bimodal q-distributions have also been previously reported for the TIP5P, TIP4P/2005 and SPC/E water models.^{92,93,100} The weak bimodality reflects the existence of two subpopulations of molecules with distinct local environments. Although these subpopulations are not completely distinguished by the order parameter q, the trajectory-averaged value of the median separates the low-q and high-q peaks at each temperature investigated (Fig. 2.1). Similar behaviour is also found when examining the q-distributions computed for single configurations (not shown). This observation suggests that the median is a reasonable choice for the threshold value \tilde{q} . As noted in Sec. 2.2.2, we have found that Bhatia-Thornton analysis is not particularly sensitive to the choice of \tilde{q} , provided that it separates the low- and high-qpeaks in the q-distribution. The median was also found to be a suitable choice for \tilde{q} in a similar study of the TIP4P/2005 water model.⁹²

The total structure factor S(Q) exhibits distinct peaks S_1 and S_2 at $Q \approx 18$ and 28 nm⁻¹, respectively (Fig. 2.2(a)). These peaks are also observed in x-ray scattering ex-



Figure 2.2: (a) Total structure factors and (b) total pair correlation functions computed at different temperatures for a system containing N = 32000 ST2 water molecules. The inset in (b) shows an enlargement of the total pair correlation function for $r \ge 2$ nm.

periments, 23,101,102 and they are associated with the peaks at approximately 0.7 and 0.45 nm in water's total pair correlation function.³⁹ The latter value is the second-nearest-neighbor separation distance expected for a tetrahedral coordination structure $\sqrt{8/3}a \approx 0.45$ nm, where $a \approx 0.28$ nm is the distance between the oxygen sites on nearest neighbours.^{23,101} The growing separation between S_1 and S_2 observed upon cooling water in both simulation and experiment thus reflects an increase in tetrahedral order as T decreases,²³ which is in agreement with the behaviour reported above for the parameter q.

The magnitude of S(Q) in the low-Q region increases dramatically as T decreases

(Fig. 2.2(a)). This behaviour has been reported in experiment and previous simulation studies, 22,90,92,100,103 and it is associated with the enhancement of fluctuations in water as it is cooled below its normal freezing point. The limiting value of S(Q)as $Q \rightarrow 0$ is related to particle number (density) fluctuations and the isothermal compressibility $\kappa_T^{90,104}$

$$\lim_{Q \to 0} S(Q) = \frac{\langle N^2 \rangle - \langle N \rangle^2}{\langle N \rangle} = \rho k_{\rm B} T \kappa_T, \qquad (2.8)$$

where $k_{\rm B}$ is Boltzmann's constant. Hence, the increase in S(0) as T decreases indicates that the metastable liquid becomes more compressible. "Simple" liquids (e.g., the Lennard-Jones fluid), by contrast, become less compressible and exhibit a concomitant decrease in S(0) as T is lowered.^{22,90} The increase in compressibility upon cooling is thus unusual behaviour for a pure liquid and is one of supercooled water's characteristic thermophysical anomalies.^{3,16}

The total pair correlation function shows that ST2 exhibits increasing short- and intermediate-range order as it is cooled, in accord with an enhancement of local tetrahedral symmetry in the liquid (Fig. 2.2(b)). We analyse the function $r^2h(r)$ rather than h(r) directly to facilitate detection of long-length-scale correlations. At the highest temperature (277 K), no correlations are observed beyond ~ 1.8 nm (Fig. 2.2(b)) inset). As T is lowered, however, longer-range correlations emerge. At the lowest temperature examined (247 K), very weak correlations can be detected up to ca. 3.5–4.0 nm. This behaviour suggests a growing correlation length in ST2 water. As we demonstrate below, the increase in the correlation length in ST2 water upon cooling is consistent with the approach towards a low-temperature LLCP at ~ 245 K.

Insight into the nature of the growing correlations in ST2 can be gained by examining the partial pair correlation functions computed by treating ST2 water as an equimolar binary mixture of low-q (L) and high-q (H) species (Fig. 2.3). The functions $h_{LL}(r)$ and $h_{HH}(r)$ reveal the presence of very weak, long-range (r > 2 nm) density correlation between like species in the system (Fig. 2.3(a),(b)). The strength and range of


Figure 2.3: Partial pair correlation functions for low-q and high-q species computed at different temperatures for a system containing N = 32000 ST2 water molecules.

these positive correlations grow as T decreases. By contrast, $h_{\rm HL}(r)$ shows that longrange density correlations between unlike species become increasingly negative as T is lowered. Long-range density correlations between like and unlike species thus make positive and negative contributions to the total pair correlation function h(r), respectively (Fig. 2.3(c)). These results demonstrate that density and concentration fluctuations are coupled in the system, and they suggest that decreasing miscibility of the two species contributes to the growth of static correlations in the liquid as T is lowered, as expected in the vicinity of a demixing transition.^{92,93}

The mixture-like behaviour of supercooled ST2 water motivates analysis of the Bhatia-Thornton structure factors to characterise fluctuations in the liquid (Fig. 2.4). As noted in Ref. 92, $S_{CC}(0)$ will diverge upon approaching a demixing transition in a binary system; $S_{NN}(0)$ and $S_{NC}(0)$ will also exhibit similar, albeit significantly less pronounced, divergent behaviour if the two components have different partial molar volumes. We find that $S_{CC}(0)$ increases rapidly as T decreases. The magnitude of the change with T is substantially larger than that observed for $S_{NN}(0) = S(0)$. The significant rise in $S_{CC}(Q)$ as $Q \rightarrow 0$ demonstrates that strong concentration fluctuations are present at each temperature examined. Similarly, the increase in $S_{NC}(Q)$ at low wavenumber reveals that density and concentration fluctuations are coupled, as suggested by the pair correlation functions. An appreciable increase in $S_{NC}(0)$ is also observed as T decreases. The behaviour of the Bhatia-Thornton structure factors is therefore consistent with the approach to a demixing transition.

The Ornstein-Zernike correlation length ξ can also be estimated by analysing the low-wavenumber behaviour of the structure factor.^{22,90,104} This quantity provides a measure of the characteristic size of correlations in the liquid, and its behaviour can be used to detect the presence of critical singularities. Specifically, along a critical



Figure 2.4: Bhatia-Thornton structure factors $S_{CC}(Q)$ (Eq. 2.3) and $S_{NC}(Q)$ (Eq. 2.4) computed at different temperatures for a system containing N = 32000 ST2 water molecules. The structure factors $S_{NN}(Q) = S(Q)$ (Eq. 2.2) are shown in Fig. 2.2(a).

isochore, ξ should exhibit power-law divergence as *T* decreases⁹³

$$\xi = \xi_0 \left(\frac{T}{T_c} - 1 \right)^{-\nu}, \tag{2.9}$$

where ξ_0 is a constant, T_c is the critical temperature, and v is the critical exponent associated with ξ . As noted in Ref. 93, ξ can be obtained from S_{CC} or the anomalous contribution to the total structure factor. The latter quantity is obtained by splitting S(Q) into normal $S^N(Q)$ and anomalous $S^A(Q)$ components:^{22,90,104}

$$S(Q) = S^{N}(Q) + S^{A}(Q).$$
 (2.10)

This division of S(Q) is motivated by the unusual behaviour of water's isothermal compressibility, which passes through a minimum around 319 K at ambient pressure.⁷ Above the minimum, κ_T decreases with *T*, as expected for a "normal" liquid. Below the minimum, however, κ_T begins to increases with decreasing *T*. The anomalous increase in κ_T upon cooling continues as far as can be measured experimentally in the supercooled regime.⁷ Thus, it has been proposed that the non-monotonic behaviour of water's isothermal compressibility can be modelled by considering κ_T as the superposition of two functions $\kappa_T = \kappa_T^N + \kappa_T^A$.^{7,78} The component κ_T^N follows the behaviour expected for a "normal" liquid and decreases monotonically with T.^{7,90,104} The anomalous component κ_T^A , on the other hand, generally increases as T decreases and accounts for enhanced density fluctuations that arise from the competition between local coordination structures in the liquid. The relationship between S(0) and κ_T given in Eq. 2.8 therefore suggests that S(Q) can be split in a similar fashion. The division of S(Q) can also be rationalised by considering water as a mixture of two "normal" liquids.¹⁰¹

Defining $S^{N}(Q)$ and $S^{A}(Q)$ is to some extent arbitrary in practice.^{101,104} In experimental studies, it is usually assumed that the normal component is independent of Q in the low wavenumber region, i.e., $S^{N}(Q) = S^{N}(0)$ for small Q.¹⁰⁴ Alternatively, the behaviour of $S^{N}(Q)$ at small Q may be approximated by extrapolation from larger wavenumbers, where the anomalous contribution to S(Q) is assumed to be negligible.^{90,104} The extrapolation is performed by imposing the constraint that $S^{N}(0) = \rho k_{\rm B} T \kappa_{T}^{\rm N}$, where $\kappa_{T}^{\rm N}$ is obtained from values tabulated in the literature.^{78,105} It has also been proposed that $S^{N}(Q)$ may be computed in simulation using the Bhatia-Thornton expression for the structure factor⁹²

$$S(Q) = S_{\rm NC}^2(Q) / S_{\rm CC}(Q) + \theta(Q),$$
 (2.11)

where

$$\theta(Q) = x_{\rm H} x_{\rm L} [S_{\rm HH}(Q) S_{\rm LL}(Q) - S_{\rm HL}^2(Q)] / S_{\rm CC}(Q).$$
(2.12)

If density and concentration fluctuations are uncoupled in the system, $S_{\text{NC}}(Q) = 0$ and Eq. 2.11 reduces to $S(Q) = \theta(Q)$. The function $\theta(Q)$ thus provides a measure of S(Q) in the absence of coupled density-concentration fluctuations,^{91,92} motivating the choice of $S^{N}(Q) = \theta(Q)$ as an operational definition for the normal contribution.

Estimates of the correlation length are obtained by fitting the low-Q regions of $S^{\rm A}(Q)$ and $S_{\rm CC}(Q)$ to Lorentzian functions^{90,104}

$$S^{\rm A}(Q) = \frac{a(T)}{\xi^{-2} + Q^2} \tag{2.13}$$

and

$$S_{\rm CC}(Q) = \frac{b(T)}{\xi^{-2} + Q^2},\tag{2.14}$$

where a(T) and b(T) are temperature-specific constants. We have used this procedure to estimate ξ from $S_{CC}(Q)$ and $S^A(Q) = S(Q) - \theta(Q)$. According to Eqs. 2.13 and 2.14, the functions $1/S^A(Q)$ and $1/S_{CC}(Q)$ should vary linearly with Q^2 over the fitted wavenumber range at each temperature. Additionally, although the finite-size corrections allow S(Q) to be estimated even at very small wavenumbers, we restrict our fitting to $Q \ge 2\pi (L/2)^{-1} \approx 1.26 \text{ nm}^{-1}$.¹⁰⁰ A suitable range of Q satisfying both of these criteria was identified at each temperature examined (Fig. 2.5).

Figure 2.6 shows estimates of the correlation length ξ calculated from $S^{A}(Q)$ (circles) and $S_{CC}(Q)$ (squares) for a system containing N = 32000 ST2 water molecules. Lines are power-law fits to the data for N = 32000 molecules using Eq. 2.9 with ξ_0 and T_c as adjustable parameters and v = 0.63, which is the critical exponent for the three-dimensional Ising universality class. The fits to ξ from $S^{A}(Q)$ and $S_{CC}(Q)$ yield $T_c = 244.6$ K and $T_c = 245.3$ K, respectively. The large black symbols denote estimates for ξ from simulations conducted using N = 256000 molecules at 247 K. Uncertainties were estimated using block analysis. The correlation lengths extracted from $S_{CC}(Q)$ and $S^{A}(Q)$ grow rapidly as T decreases, in accord with the power-law divergence predicted by Eq. 2.9 (Fig. 2.6). Histogram reweighting studies conducted



Figure 2.5: Ornstein-Zernike plots for (a) $S^{A}(Q)$ and (b) $S_{CC}(Q)$. Points are data from MD simulations of N = 32000 ST2 water molecules, whereas lines in (a) and (b) are fits using Eqs. 2.13 and 2.14, respectively.

using small systems (N < 800) suggest that ST2's LLPT belongs to the 3D Ising universality class.^{33,39} Very close to $T_{\rm c}$, it is therefore expected that ξ will grow via Eq. 2.9 with the 3D Ising scaling exponent $v \approx 0.63$.^{94,106–108} Further away from T_c , however, the scaling of ξ should cross over to mean-field behaviour with v = 0.5.^{106–108} Uncertainty in the location of $T_{\rm c}$ and statistical error in ξ prevent us from distinguishing between the crossover region and the asymptotic regime where Ising-like scaling is expected. Consequently, in fitting the data to Eq. 2.9, we set v = 0.63 and used ξ_0 and T_c as adjustable parameters. We observe that the data are well-described by power-law fits using the 3D Ising exponent $v \approx 0.63$ (Fig. 2.6). Ising-like scaling behavior for ξ has reported over the same range of $T/T_c - 1$ for the demixing transition in symmetric binary Lennard-Jones mixtures.⁹⁴ We have also fit the data by fixing $T_{\rm c}$ and using ν and ξ_0 as adjustable parameters (not shown). Our analysis suggests 245 ± 1 K and 0.6 ± 0.1 as best estimates for T_c and v. The value $T_c \approx 245$ K is in excellent agreement with estimates of 246 ± 1 and 247 ± 3 K for ST2's liquidliquid critical temperature obtained from histogram reweighting³⁹ and equation of state analysis,³⁸ respectively. Finally, to gauge the magnitude of finite-size effects at these conditions, we repeated the analysis at the lowest temperature (247 K) for a significantly larger system containing N = 256000 water molecules, as discussed in Sec. 2.2.2. We find that the values of ξ computed for both systems are in good agreement, indicating that the system with N = 32000 molecules is large enough to obtain accurate estimates of the correlation length over the range of conditions examined in this study.

Bhatia-Thornton analysis has also been used to characterise fluctuations in TIP5P and TIP4P/2005.⁹³ The behaviour of ξ for TIP5P was found to be consistent with the existence of a low-temperature LLCP near 212 K. The correlation length for TIP4P/2005, by contrast, was reported to exhibit much weaker scaling with T than expected in the vicinity of an LLCP. Equation of state analysis suggests that



Figure 2.6: Estimates of the correlation length ξ calculated from $S^{A}(Q)$ (circles) and $S_{CC}(Q)$ (squares) for a system containing N = 32000 ST2 water molecules. Lines are power-law fits using Eq. 2.9

both models may exhibit an LLPT at deeply supercooled conditions.^{24,29,87} To our knowledge, however, the existence of such a transition has not been confirmed in either model using rigorous free energy methods, such as those that have been applied to locate and characterise ST2's LLPT. Free energy calculations are expected to be extremely challenging to perform for these models, particularly for TIP4P/2005 where slow relaxation processes in the liquid frustrate sampling near the hypothesised LLPT.⁸⁷ Such calculations have only been possible for ST2 because its LLPT occurs at higher temperatures where relaxations are relatively fast.^{28,30,88} Thus, analysis of TIP5P and TIP4P/2005 is complicated by the fact that the location of their LLPTs, should they occur, is not precisely known.

Recent studies also demonstrated that the thermodynamic stability of LLPTs in tetrahedral fluids can be remarkably sensitive to model parameters.^{41,42} Adjustment of bond-angle flexibility in colloidal models of tetravalent DNA nanostars can cause the LLPTs in these systems to vanish completely.⁴¹ Similarly, ST2's metastable LLPT becomes thermodynamically stable with respect to crystallisation when bondangle flexibility is introduced into the model.⁴² An LLPT can also be induced in the WAC model of silica by tuning the Coulombic interactions.^{61,109} These results suggest that not all water models necessarily exhibit a low-temperature LLPT. Indeed, there seems to be no such transition in the mW water model.^{26,88}

The existence of a metastable LLPT in ST2 suggests that it should be possible to observe spontaneous phase separation and the formation of a stable liquid-liquid interface at $T < T_c$. Yagasaki et al.¹¹⁰ have reported evidence of a spontaneous phase separation in ST2. They used canonical ensemble molecular dynamic simulations to instantaneously quench systems containing N = 4000 molecules from 300 to 235 K (0.96 T_c) along the $\rho = 0.98$ g cm⁻³ isochore, which crosses into the two-phase region. The simulations were performed in a rectangular simulation cell with an aspect ratio of $L_x:L_y:L_z = 1:1:4$ to promote formation of a planar interface. As expected, they observed the development of low- and high-density domains (corresponding to ST2's LDL and HDL phases, respectively), which were separated by a diffuse interface oriented with its surface normal parallel to the major axis of the simulation cell. They also reported similar behaviour in complementary simulations performed for TIP5P and TIP4P/2005.

Overduin and Patey⁹³ subsequently applied the approach of Yagasaki et al. to study the behaviour of TIP5P and TIP4P/2005 in larger systems containing up to N = 32000 molecules. Interestingly, they found that density of the quenched systems became increasingly homogeneous as N was increased from 4000 to 32000. In the largest system with N = 32000 molecules, no evidence of liquid-liquid phase separation was observed. Thus they concluded that the behaviour reported by Yagasaki et al. for TIP5P and TIP4P/2005 is a finite-size artifact associated with using small systems and not evidence of an underlying LLPT. Although Overduin and Patey did not perform simulations for ST2, they hypothesised that the simulations conducted for this model by Yagasaki et al. may also suffer from finite-size artifacts.⁹³

To test this hypothesis, we have performed similar molecular dynamics simula-



Figure 2.7: (top) Density profile $\rho(z)$ at 237 K computed along the major axis (z-axis) of a rectangular simulation cell containing N = 256000 ST2 molecules after the system was thermally quenched into the region of liquid-liquid coexistence. (bottom) The HDL-LDL interfaces are observed in renderings of the system in which the particles are shaded according to the local density across the z-axis of the simulation cell.

tions for ST2 using a significantly larger system containing N = 256000 molecules. The density of the system was fixed at $\rho = 0.98$ g cm⁻³, which is the same value used by Yagasaki et al. We also used an elongated rectangular simulation cell, but chose a different aspect ratio ($L_x:L_y:L_z = 1:1:\sqrt{2}$). The minimum and maximum dimensions of the simulation cell were therefore set to $L_x = L_y = 17.67$ and $L_z = 24.99$ nm, respectively.

The system was equilibrated at 300 K for 1 ns and then quenched to 237 K by abruptly changing the set temperature of the thermostat. The quench was followed by HDL–LDL phase separation, as evidenced by the rapid formation of large lowand high-density domains that spanned the x-y plane of the simulation cell. Although their positions fluctuate, the domains persist for the duration of the simulation trajectory (ca. 200 ns). Analysis of the density profile $\rho(z)$ reveals that the liquid domains are separated by an interface normal to the *z*-axis and have densities of 0.89 and 1.07 g cm⁻³, respectively (Fig. 2.7). These values are consistent with the densities reported for LDL and HDL by Yagasaki et al.¹¹⁰ for smaller systems with N = 4000 molecules. They are also in good agreement with location of the LDL and HDL basins observed in free energy surfaces computed for ST2 using N = 216 molecules.^{30,42} These studies collectively show that an LLPT is observed in ST2, independent of system size, across a range of N that spans at least three orders of magnitude. Hence, they demonstrate that ST2 exhibits a genuine metastable LLPT that is not connected to, or strongly influenced by, artifacts stemming from finite-size effects.

Finally, we note that as $N \to \infty$, the characteristic time scale for ice nucleation will eventually become smaller than that required to relax the liquid and thus direct observation of ST2's LLPT will be prevented by rapid crystallisation. For the N = 256000 molecule system studied here, however, we did not observe evidence of crystallisation. Molecules with ice-like local order identified using the CHILL+ algorithm¹¹¹ were found to constitute less than 7 % of the total population over the duration of the ca. 200 ns MD trajectory. By contrast, the liquid-liquid separation process completed in a few nanoseconds, which is comparable to the liquid relaxation time under these conditions.¹¹⁰ The crystallisation and liquid relaxation time scales thus appear to be separated to some extent even in the largest systems and at the lowest temperatures examined in our study. As reported by Yagasaki et al.,¹¹⁰ we anticipate that crystallisation in ST2 may be observed by extending the duration of the simulations into the microsecond regime. Although this might be possible in future investigations, the computational cost associated with simulating these very large systems has prevented us from accessing such time scales in the present study.

2.4 Conclusions

We have performed large-scale MD simulations of systems containing up to N = 256000 molecules to investigate the low-temperature behaviour of the ST2 water model. Free energy studies of small systems with $N \approx 200$ molecules have demonstrated that ST2 exhibits a phase transition between two metastable liquids (HDL and LDL) below a critical temperature of $T_c \approx 245$ K. The liquids have different densities and are also distinguished by the extent of tetrahedral order in their local coordination environment. Here, we have exploited these differences to analyse fluctuations in supercooled ST2 water by treating it as an equimolar binary mixture of molecular species with low and high tetrahedral order in their primary coordination shell, respectively.

We found that ST2 exhibits anomalous scattering at low wavenumbers, similar to that observed experimentally in supercooled water. This behaviour is directly related to supercooled water's well-known increase in compressibility upon cooling, and it demonstrates that fluctuations in the liquid grow in magnitude as temperature is lowered. Analysis of the Bhatia-Thornton structure factors revealed that density, concentration, and coupled density-concentration fluctuations contribute to ST2's anomalous scattering. Estimates of the correlation lengths associated with the anomalous scattering component and concentration fluctuations were also extracted using Bhatia-Thornton analysis. Both correlation lengths increase rapidly upon cooling and exhibit power-law growth consistent with the presence of a critical singularity at $T_c \approx 245$ K. Thus the rapid growth of the correlation lengths points to the existence of a low-temperature LLCP in ST2, in accord with previous computational studies of significantly smaller systems.^{38,39}

Finally, we observed that spontaneous liquid-liquid phase separation occurs when large systems with N = 256000 molecules are thermally quenched below $T_{\rm c}$ into the two-phase region. Similar behaviour was also reported by Yagasaki et al.¹¹⁰ for sig-

nificantly smaller systems containing N = 4000 molecules. The consistency between the two studies therefore demonstrates that ST2's behaviour is not sensitive to system size. By contrast, Overduin and Patey⁹³ found that low- and high-density liquid domains in TIP5P and TIP4P/2005 gradually vanish as N is increased from 4000 to 32000. As discussed in Ref. 87, there are many possible explanations for this behaviour. Distinguishing between the different scenarios, however, will be challenging and require performing numerous large-scale MD simulations of TIP5P and TIP4P/2005 to examine the effects of box shape, system density, and equilibration time, among other variables, on the apparent phase behaviour of these models. We therefore leave in-depth examination of these models for future investigations.

Chapter 3: Fluctuations near the liquid-liquid transition in a model of silica

This chapter was previously published by Royal Society of Chemistry: Guo, J.; Palmer, J. C. Fluctuations near the liquid–liquid transition in a model of silica. *Phys. Chem. Chem. Phys.* **2018**, *20*, 25195–25202

3.1 Introduction

Tetrahedral liquids, such as water and silica, exhibit a number of well-known thermophysical anomalies.^{3,16,19,112} Whereas "simple" non-network-forming liquids densify upon cooling, for example, many tetrahedral liquids exhibit density maxima, below which they expand as temperature decreases.^{3,16,112} The non-monotonic behavior of the density arises from the conversion of entropically favored high-density arrangements in the liquid into energetically favored low-density tetrahedral coordination structures.^{3,92,112} It has been proposed that this mixture-like behavior^{81,86,113} may result in a liquid-liquid phase transition (LLPT) between a high-density and a low-density phase (HDL and LDL, respectively) at deeply supercooled conditions.^{3,12,16,17,112} According to the LLPT hypothesis, anomalies such as the marked increase in the isothermal compressibility of tetrahedral liquids upon cooling^{7,16,89} arise due to fluctuations associated with the liquid-liquid critical point (LLCP) terminating the line of HDL–LDL coexistence. This intriguing scenario also posits that the high-density and low-density amorphous glasses (HDA and LDA) of water^{11,82–84} and silica¹¹⁴ are kinetically arrested forms of the two liquids.

Results from recent experiments on stretched and deeply supercooled water are consistent with the LLPT hypothesis.^{22,115} Direct observation of such a transition, however, has thus far been prevented by the rapid nucleation of ice near water's

predicted LLCP.^{3,16,17,23,62,74} Computational studies have also investigated the existence of LLPTs in models of tetrahedral liquids .^{24,26,27,29,38,40,41,60,68,85–87,116} Coarsegrained models of DNA tetramers¹¹⁶ and tetra-functional patchy particles⁴¹ have been shown to exhibit LLPTs. Similar behavior has not been confirmed in most molecular and atomic systems, however, due to the computationally demanding nature of equilibrating their liquid phases at the low temperatures where the transitions are predicted to occur. Exceptions include the ST2 model of water^{12,36} and the modified Woodcock-Angell-Cheeseman (mWAC) model of silica, ^{57,61} which both exhibit low-temperature LLPTs.^{28,30,32–35,39,42,109} Although the existence of an LLPT in ST2 has been challenged by Limmer and Chandler, ^{25,117} it was recently demonstrated that the results and conclusions presented in their studies are invalid due to a serious conceptual error in the design of their simulation code.^{112,118}

Models such as ST2 water and mWAC silica, while not quantitatively accurate in predicting properties of these substances, provide unique opportunities to elucidate the connections between LLPTs and other anomalous thermophysical behaviors, and thus guide interpretation of the growing body of experimental work on tetrahedral liquids. In this study, we perform large-scale molecular dynamics (MD) simulations of up to 216,000 atoms to investigate the scattering behavior of the mWAC model of silica near its LLPT. We find that mWAC exhibits anomalous scattering, similar to that observed in experimental and computational studies on supercooled water, in which the magnitude of the static structure factor S(Q) at low-Q increases upon cooling. This behavior arises from a significant increase in density fluctuations in the liquid. Bhatia-Thornton analysis reveals that the enhanced density fluctuations are strongly coupled to those associated with local tetrahedral order, consistent with proximity to an HDL-LDL demixing transition. Moreover, we find that the Ornstein-Zernike correlation length estimated from the anomalous scattering component exhibits power-law growth upon cooling and appears to diverge near 3300 K, which is in excellent agreement with previous estimates of mWAC's liquid-liquid critical temperature. Finally, we use MD to thermally quench systems with 4500, 36,000, and 432,000 atoms below the predicted LLCP into the two-phase region. We observed spontaneous HDL–LDL phase separation immediately following the quench, independent of system size. These findings are similar to those reported in our analogous investigation of the ST2 model of water,¹¹⁹ suggesting common signatures that may be useful in identifying LLPTs in other models and interpreting scattering experiments on tetrahedral liquids.

Methods

Molecular dynamics simulations

Large-scale MD simulations of the mWAC silica (SiO₂) model were performed using GROMACS 4.6.7.¹²⁰ Following the original model of Woodcock, Angell, and Cheeseman⁵⁷ (WAC), the mWAC model describes SiO_2 as a 1:2 binary mixture of Si⁺⁴ and O⁻² ions, which interact through a standard Coulomb-Buckingham potential.^{60,61} Parameters for mWAC are identical to those of the original WAC model.^{57,60} In addition, the mWAC model introduces a new parameter f that scales the magnitudes of the charges on the ions.^{61,109} The original WAC model (f = 1) does not exhibit an LLPT in the computationally accessible region of its phase diagram.⁶¹ Reducing f to 0.84, however, induces an LLPT under conditions that can be readily explored with simulation.^{61,109} The existence of an LLPT in the mWAC model with f = 0.84 has been confirmed by both equation of state and free energy calculations.^{61,109} Spontaneous liquid-liquid phase separation has also been directly observed in mWAC upon thermally quenching large systems containing N = 108,000atoms at fixed density into the predicted two-phase region.¹⁰⁹ These studies suggest that the liquid-liquid critical temperature, pressure, and density for mWAC with f = 0.84 are $T_{\rm C} = 3350 \pm 75$ K, $p_{\rm C} = 0.19 \pm 0.09$ GPa, and $\rho_{\rm C} = 1.8 \pm 0.1$ g cm⁻³, respectively.

We performed canonical (NVT) ensemble MD simulations to study the behavior of the mWAC model with f = 0.84 in the vicinity of its LLCP. The simulations were conducted approximately along the critical isochore ($\rho \approx 1.85$ g cm⁻³) at different temperatures (3375 K $\leq T \leq$ 3800 K) for a system with N = 36,000 atoms in a periodic, cubic cell with edge length L = 8.65 nm. This large system was chosen to minimize finite-size effects and facilitate investigation of long-wavelength fluctuations near $T_{\rm C}$. The equations of motion were integrated using the velocity-Verlet algorithm with a 2 fs time step, and the temperature was maintained using a Nosé-Hoover thermostat^{95,96} with a 2 ps time constant. Pairwise interactions were truncated at 1.0 nm, and long-range contributions to the electrostatics were treated using the particle mesh Ewald method, with parameters chosen to ensure a relative error of less than 10^{-4} in the calculated energy. Each simulation was equilibrated for 25 ns, followed by a production period of 175 ns. Configurations were saved every 20 ps during the production period for subsequent analysis. Three independent runs were performed at the lowest temperatures ($T \leq 3450$ K) to improve sampling statistics.

Fluctuation analysis

Density fluctuations in mWAC were analyzed by computing the static structure factor,

$$S(Q) = N_{\rm s}^{-1} \sum_{j,l=1}^{N_{\rm s}} \langle \exp\left[i\mathbf{Q}\cdot(\mathbf{r}_l - \mathbf{r}_j)\right] \rangle, \qquad (3.1)$$

where $\{\mathbf{r}_i\}_{i=1}^{N_s}$ are the positions of the N_s scatterers and $Q = |\mathbf{Q}|$ is the norm of the wavevector. For a homogeneous isotropic system, S(Q) can be expressed in terms of the Fourier transform of the radial distribution function, g(r):

$$S(Q) = 1 + 4\pi\rho_{\rm s} \int_0^\infty r^2 \frac{\sin(Qr)}{Qr} [g(r) - 1] dr, \qquad (3.2)$$

where $\rho_{\rm s} = N_{\rm s}/V$ and V is the volume of the system. The small-Q limit of S(Q) is related to fluctuations in $N_{\rm s}$ and the isothermal compressibility $\kappa_T^{90,104}$

$$S(0) = \lim_{Q \to 0} S(Q) = \frac{\langle N_{\rm s}^2 \rangle - \langle N_{\rm s} \rangle^2}{\langle N_{\rm s} \rangle} = \rho_{\rm s} k_{\rm B} T \kappa_T, \qquad (3.3)$$

where $k_{\rm B}$ is Boltzmann's constant.

For our analysis, we set $N_{\rm s} = N_{\rm Si} = N/3$ and only considered scattering from Si atoms, noting that similar results are obtained when contributions from O atoms are also included. In analogy to O atoms in water's tetrahedral O–H–O hydrogenbond network, Si atoms serve as "molecular" centers that are linked via Si–O–Si ionic bonds in liquid silica.¹²¹ Hence, they are suitable scattering centers for characterizing fluctuations in silica's tetrahedral network. Accordingly, we evaluated S(Q)numerically from the Si–Si radial distribution function via Eq. 3.2, using standard finite-size corrections⁹⁹ and an upper integration limit of $r_{\rm max} = L/2$.

Coupled fluctuations in density and tetrahedral order were characterized using the Bhatia-Thornton analysis proposed by Overduin and Patey⁹² for supercooled water. This approach treats the pure liquid as a binary mixture of molecular species with low (L) and high (H) tetrahedral symmetry in their first coordination shell, respectively. We adapted this analysis to study fluctuations in liquid silica by computing the local tetrahedral order parameter^{92,97}

$$q_{i} = 1 - \frac{3}{8} \sum_{j=1}^{3} \sum_{l=j+1}^{4} \left(\cos \phi_{jl} + \frac{1}{3} \right)^{2}, \qquad (3.4)$$

where ϕ_{jl} is the angle formed by the lines joining Si atom *i* with its *j*th and *l*th nearest Si neighbors ($\{j, l\} \leq 4$). Silicon atoms with q_i values smaller or larger than the median were classified as L and H species, respectively. This approach resulted in an equimolar mixture and allowed the classification of each Si atom to change with time.⁹²

Density (NN), species concentration (CC), and coupled density-concentration (NC) fluctuations were analyzed by computing the Bhatia-Thornton structure factors for isotropic binary mixtures:^{91,92}

$$S_{\rm NN}(Q) = S(Q)$$
$$= x_{\rm H}S_{\rm HH}(Q) + x_{\rm L}S_{\rm LL}(Q) + 2\sqrt{x_{\rm H}x_{\rm L}}S_{\rm HL}(Q), \qquad (3.5)$$

$$S_{\rm CC}(Q) = x_{\rm H} x_{\rm L} [x_{\rm L} S_{\rm HH}(Q) + x_{\rm H} S_{\rm LL}(Q) - 2\sqrt{x_{\rm H} x_{\rm L}} S_{\rm HL}(Q)], \qquad (3.6)$$

and

$$S_{\rm NC}(Q) = x_{\rm H} x_{\rm L} [S_{\rm HH}(Q) - S_{\rm LL}(Q) + \frac{x_{\rm L} - x_{\rm H}}{\sqrt{x_{\rm H} x_{\rm L}}} S_{\rm HL}(Q)], \qquad (3.7)$$

where $x_{\{H,L\}} = 0.5$ are the species mole fractions. Equations 3.5 –3.7 are functions of the Ashcroft-Langreth partial structure factors⁹⁸ for species α and β ,

$$S_{\alpha\beta}(Q) = \delta_{\alpha\beta} + 4\pi\rho_{\rm s}\sqrt{x_{\alpha}x_{\beta}} \int_0^\infty r^2 \frac{\sin(Qr)}{Qr} \left[g_{\alpha\beta}(r) - 1\right] dr, \qquad (3.8)$$

where $\delta_{\alpha\beta}$ is the Kronecker delta function and $g_{\alpha\beta}(r)$ is the partial radial distribution function. Because the scattering lengths are identical for both species in the present case, $S_{NN}(Q) = S(Q)$, as indicated by Eq. 3.5. Thus, $S_{NN}(Q)$ and S(Q) are used interchangeably in our analysis and discussion.

Results and Discussion

We applied the analysis methods described in the previous section to characterize the behavior of the mWAC model near its LLPT. The primary peak in the total structure factor S(Q) is split into a weak shoulder S_1 at $Q \approx 15 \text{ nm}^{-1}$ and a pronounced maximum S_2 at $Q \approx 23 \text{ nm}^{-1}$ (Fig. 3.1(a)). The separation between S_1 and S_2 in S(Q)has been shown to be a sensitive measure of the degree of local tetrahedral order in the liquid.^{23,122} In deeply supercooled water, S_1 and S_2 are more clearly resolved and become increasingly separated as T is lowered, indicating a significant enhancement in tetrahedral order upon cooling.^{23,101,102,119} In the mWAC model, by contrast, they are poorly resolved at all temperatures, suggesting that the extent of tetrahedral order and its increase upon cooling are not as significant.

The radial distribution function g(r) reveals a slight increase in short- and intermediate-range order as the mWAC model is cooled (Fig. 3.1(b)). Very weak longlength-scale correlations are also observed in g(r) (Fig. 3.1(b) inset), which increase in range from 2.8 to more than 4.0 nm as T is lowered from 3800 to 3375 K. Similar behavior has been reported in supercooled water,^{93,119} and it suggests the growth of a static correlation length. These long-length-scale correlations give rise to the sharp increase in S(Q) below $Q \approx 5$ nm⁻¹ (Fig. 3.1(a)). The magnitude of the structure factor in the small-Q limit, S(0), also increases rapidly as T decreases (Fig. 3.1(a) inset). This trend has been reported in previous experimental and computational studies on supercooled water,^{22,90,92,100,103,119} and it signifies a pronounced increase in density fluctuations upong cooling, as expected near a phase transition. As indicated by Eq. 3.3, the thermodynamic consequence of this behavior is an increase in κ_T as T decreases, which is a well-known anomaly of water and other tetrahedral liquids.^{3,16,112}

The increase in density fluctuations upon cooling is accompanied by changes in tetrahedral order. The probability density distribution for the tetrahedral order parameter q is bimodal and exhibits distinct peaks near $q \approx 0.5$ and 0.85 (Fig. 3.2(a)), suggesting the existence of two subpopulations with distinct local coordination structures. The high-q peak is located close to the value q = 1 expected for a regular tetra-



Figure 3.1: (a) Structure factor and (b) radial distribution function computed for a system of N = 36,000 atoms. The inset in (a) shows S(0) as a function of T, and the inset in (b) shows an enlargement of $r^2[g(r)-1]$ for $r \ge 2.5$ nm.

hedron, revealing that a substantial fraction of the Si atoms have highly ordered coordination environments. As T is lowered, the low- and high-q peaks decrease and increase in magnitude, respectively. This behavior signifies an increase in tetrahedral order, in agreement with the analysis of S(Q) and g(r). Similar trends have been observed in computational studies of water models such as TIP5P, TIP4P/2005, SPC/E, and ST2.^{92,93,100,119} The high-q peak for these water models is more pronounced, however, particularly at low temperatures, reflecting a higher degree of tetrahedral symmetry in the liquid. Near ST2's LLPT, for example, the low-q peak

appears as a weak shoulder in the tail of the strongly dominant high-q peak.¹¹⁹ By contrast, the q-distribution for mWAC remains bimodal with distinct low- and high-q peaks, even at low T.

The bimodal nature of the q-distributions motivates application of the Bhatia-Thornton mixture analysis to characterize concentration fluctuations in the two subpopulations of Si atoms. For a binary system, the magnitude of $S_{\rm CC}(0)$ should diverge as a demixing transition is approached.^{92,119} Similarly, if the partial molar volumes of the two components differ, $S_{\rm NN}(0)$ and $S_{\rm NC}(0)$ should also exhibit divergent behavior.^{92,119} Accordingly, we observe that the magnitudes of $S_{\rm CC}(0)$ and $S_{\rm NC}(0)$ increase rapidly as T is lowered towards $T_{\rm C}$ (Fig. 3.2(b)). This behavior indicates a significant increase in species concentration and coupled density-concentration fluctuations upon cooling, consistent with approach to a demixing transition. The growing inhomogeneities in concentration can be directly observed in simulation snapshots where Si atoms are colored according to their species designation (Fig. 3.2(c)).

The low-Q region of the structure factors can also be analyzed to estimate the Ornstein-Zernike (OZ) correlation length ξ , a measure of the characteristic length scale over which static correlations in the liquid decay. The quantity ξ can be obtained from $S_{\rm CC}(Q)$ or the anomalous contribution to the structure factor,⁹³

$$S^{\mathcal{A}}(Q) \equiv S(Q) - S^{\mathcal{N}}(Q), \qquad (3.9)$$

where $S^{N}(Q)$ is the normal component. This definition assumes that "anomalous" fluctuations associated with the critical phenomena are superimposed on "normal" thermal fluctuations in the liquid. Unfortunately, defining $S^{N}(Q)$ and $S^{A}(Q)$ is to some extent arbitrary in practice. ^{101,104} Consequently, following Refs. 92,93 and our previous study of the ST2 water model, ¹¹⁹ we motivate our definitions of these func-



Figure 3.2: (a) Probability density distribution for the tetrahedral order parameter q. The median values of q at the highest (3800 K) and lowest (3375 K) temperatures are indicated by vertical dashed lines. (b) Bhatia-Thornton structure factors $S_{\rm CC}(Q)$ (Eq. 3.6) and $S_{\rm NC}(Q)$ (Eq. 3.7). (c) Simulation snapshots at 3800 K and 3375 K showing the H (red) and L (blue) species of Si atoms. Data in (a) – (c) are for a system containing N = 36,000 atoms.

tions from the Bhatia-Thornton expression for the structure factor

$$S(Q) = S_{\rm NC}^2(Q) / S_{\rm CC}(Q) + \theta(Q),$$
 (3.10)

where

$$\theta(Q) = x_{\rm H} x_{\rm L} [S_{\rm HH}(Q) S_{\rm LL}(Q) - S_{\rm HL}^2(Q)] / S_{\rm CC}(Q).$$
(3.11)

In the absence of coupling between density and species concentration fluctuations, $S_{\rm NC}(Q) = 0$ and Eq. 3.10 reduces to $S(Q) = \theta(Q)$. This behavior suggests $S^{\rm N}(Q) = \theta(Q)$ as a suitable operational definition for the normal component, which we adopted in our analysis.

The OZ correlation length ξ was estimated by fitting the low-Q regions of $S_{\rm CC}(Q)$ and $S^{\rm A}(Q)$ to Lorentzian functions^{90,104}

$$S_{\rm CC}(Q) = \frac{a(T)}{\xi^{-2} + Q^2} \tag{3.12}$$

and

$$S^{\rm A}(Q) = \frac{b(T)}{\xi^{-2} + Q^2},\tag{3.13}$$

where a(T) and b(T) are temperature-specific constants. Two criteria were used to determine the appropriate range of Q over which the fitting should be performed at each temperature. First, as indicated by Eqs. 3.12 and 3.13, $1/S_{CC}(Q)$ and $1/S^{A}(Q)$ should vary linearly with Q^2 over the fitted range. Second, the fitted range should be restricted to $Q \ge 2\pi (L/2)^{-1} \approx 1.45$ nm⁻¹¹⁰⁰ to ensure consistency with the upper limit $r_{max} = L/2$ used in evaluating g(r). In each case, a graphical approach was used to identify an appropriate range of Q satisfying both criteria (Fig. 3.3).

Along a critical isochore, ξ should exhibit power-law divergence as $T \to {T_{\rm C}}^{93}$

$$\xi = \xi_0 \epsilon^{-\nu}, \tag{3.14}$$

Figure 3.3: Ornstein-Zernike plots for (a) $S_{CC}(Q)$ and (b) $S^{A}(Q)$. Points are data from MD simulations of N = 36,000 atoms. Solid lines in (a) and (b) are fits to the Lorentzian functions given by Eqs. 3.12 and 3.13, respectively.

where $\epsilon = (T - T_{\rm C})/T_{\rm C}$, v is the critical exponent associated with ξ , and ξ_0 is a constant. In accord with this expectation, ξ estimated from $S_{\rm CC}(Q)$ and $S^{\rm A}(Q)$ increases rapidly upon cooling (Fig. 3.4). Liquid-liquid transitions in tetrahedral fluids are thought to belong to the 3D Ising universality class.¹²³ Indeed, Bhatia-Thornton analysis¹¹⁹ and histogram reweighting studies^{33,39} indicate that the behavior of the LLPT in ST2 is consistent with the 3D Ising universality class. Consequently, the scaling exponent v should equal approximately $0.63^{94,106-108}$ for small ϵ and eventually cross over to the mean-field value of 0.5 as ϵ increases.¹⁰⁶⁻¹⁰⁸ Uncertainty in the

Figure 3.4: Ornstein-Zernike correlation length ξ estimated from $S_{\rm CC}(Q)$ (squares) and $S^{\rm A}(Q)$ (circles) for a system of N = 36,000 atoms. Lines are powerlaw fits using Eq. 3.14

values of ξ and $T_{\rm C}$, however, typically prevent the asymptotic and crossover regimes from being clearly distinguished in computational investigations. In fitting the data to Eq. 3.14, we therefore followed previous studies^{93,94,119} and set v = 0.63, while leaving $T_{\rm C}$ and ξ_0 as adjustable parameters. The data for $S_{\rm CC}(Q)$ and $S^{\rm A}(Q)$ are clearly captured by fits using $T_{\rm C} = 3280 \pm 30$ and 3300 ± 40 K, respectively, demonstrating that the increase in ξ upon cooling is consistent with the power-law divergence expected near a critical singularity. The large black symbols in Fig. 3.4 are estimates of ξ at 3375 and 3350 K for a system of N = 216,000 atoms. Uncertainties were estimated using block analysis. These estimates for $T_{\rm C}$ are also in excellent agreement with the value 3350 ± 75 K obtained from previous equation of state and free energy calculations.¹⁰⁹

We also performed simulations of a significantly larger system (L = 15.72 nm) containing N = 216,000 atoms at 3375 and 3350 K to assess the magnitude of finitesize effects on the values of ξ calculated at low temperatures. The estimates of ξ are in statistical agreement at 3375 K, the common temperature examined for both systems (Fig. 3.4). Further, the values of ξ computed from $S_{CC}(Q)$ and $S^{A}(Q)$ at 3350 K are consistent with extrapolation of the power-law fits to the data for the smaller system. These results demonstrate that the system with N = 36,000 atoms is large enough to accurately estimate ξ over the range of temperatures examined. They also suggest that the power-law fits based on Eq. 3.14 reasonably capture the divergent behavior of ξ near mWAC's LLCP.

The low-Q behavior of the Bhatia-Thornton structure factors and the power-law divergence of ξ upon cooling are consistent with the existence of an LLCP in the mWAC model. The associated LLPT in mWAC can also be directly observed in simulation at $T < T_{\rm C}$. In our previous study of mWAC,¹⁰⁹ we used canonical ensemble MD simulations to study liquid-liquid phase separation in a large system with N = 108,000 atoms. The simulations were performed in a rectangular simulation cell with an aspect ratio of $L_x:L_y:L_z = 1:1:3$ to promote the formation of planar liquid-liquid interfaces, and the mean density was fixed at 1.87 g cm^{-3} , which is in between the characteristic values of the HDL and LDL phases (2.10 and 1.55 g cm⁻³, respectively). The system was equilibrated at $T_{high} = 5000$ K for 1 ns and then thermally quenched to a lower temperature T_{low} by instantaneously changing the set point of the thermostat. Small HDL- and LDL-like domains developed in the system after quenching to $T_{\rm low}$ = 3450 and 3300 K, but no LLPT was observed.¹⁰⁹ At $T_{\rm low}$ = 3150 and 3000 K, by contrast, the system spontaneously phase separated into large HDL and LDL domains partitioned by planar interfaces that spanned the xand y-dimensions of the simulation cell.¹⁰⁹ The MD simulations thus showed clear evidence of an LLPT in mWAC below the critical temperature $T_{\rm C}\approx 3300~{\rm K}$ estimated from analysis of ξ in this study (Fig. 3.4).

Yagasaki et al.¹¹⁰ performed similar simulations to study the possibility of liquidliquid phase separation in the TIP5P, TIP4P/2005, and ST2 water models. Upon thermally quenching small systems with N = 4000 molecules into the predicted liquidliquid coexistence region of each model, they observed the formation of HDL and LDL domains separated by diffuse interfaces. Overduin and Patey⁹³ noted similar behavior for TIP5P and TIP4P/2005 in small systems with N = 4000 molecules, but found that the HDL and LDL domains gradually vanished as N was increased and observed no phase separation in systems with N = 32,000 molecules. Consequently, they posited that the behavior reported by Yagasaki et al. was a finite-size artifact rather than a genuine LLPT. Overduin and Patey did not investigate the behavior of the ST2 water model in their study, but they speculated that the liquid-liquid separation observed by Yagasaki et al. in this model may also be a finite-size artifact.⁹³ Recently, however, we observed HDL–LDL immiscibility in large-scale MD simulations of N = 256,000 ST2 water molecules, in accord with Yagasaki et al.'s study of significantly smaller systems.¹¹⁹ These findings are fully consistent with the wellestablished existence of an LLPT in ST2,^{12,28,30,32–35,39,42} and they demonstrate that its liquid-liquid separation behavior is not a finite-size artifact.

We investigated the influence of system size effects on mWAC's LLPT by performing canonical ensemble MD simulations, analogous to those described above, using N = 4500, 36,000, and 432,000 atoms. To facilitate comparison with our previous study,¹⁰⁹ the simulations were conducted by fixing the mean density at 1.87 g cm⁻³ and using elongated rectangular simulation cells with aspect ratios of $L_x:L_y:L_z = 1:1:3$. Following the quench from $T_{high} = 5000$ to $T_{low} = 3000$ K, each system rapidly phase separated into HDL and LDL. Density profiles computed over the last 25 ns of the ~ 200 ns MD trajectories demonstrate that the interfaces separating HDL and LDL are oriented normal to the z-axis of each simulation cell (Fig. 3.5). As expected, they also show that the HDL and LDL domains have densities of ca. 2.10 and 1.55 g cm⁻³, respectively. These findings are in excellent agreement with our previous MD simulations of N = 108,000 atoms and our free energy calculations using N = 1500.¹⁰⁹ The only qualitative difference between the three systems is that four HDL-LDL interfaces, rather than two, are observed when N = 432,000 (Fig.

Figure 3.5: Density profiles at $T_{low} = 3000$ K computed along the major axes (zaxes) of rectangular simulation cells containing (a) N = 4500, (b) 36,000, and (c) 432,000 atoms. The abscissa in each panel is normalized by the length of the simulation cell, L_z . (d) Liquid-liquid interfaces are observed in snapshots of the largest system (N = 432,000) in which Si atoms are colored according to the local density profile in (c).

3.5). This behavior indicates that the condition $L_z \ll \xi_z$ is no longer satisfied, where ξ_z is the correlation length in the *z*-direction;¹²⁴ a system with only two liquid-liquid interfaces can thus be recovered by simply reducing the aspect ratio of the simulation cell. Hence, our simulations demonstrate that mWAC's LLPT can be directly observed, independent of system size, over the computationally accessible range of N spanning more than two orders of magnitude.

Conclusions

We conducted large-scale MD simulations of systems up to N = 216,000 atoms in size to study static fluctuations in the mWAC model of silica near its LLPT. Previous equation of state and free energy studies show that mWAC exhibits an LLPT below the critical temperature $T_{\rm C} \approx 3350$ K.^{61,109} In accord with these studies, we found that density fluctuations in the liquid increase rapidly upon cooling, as evidenced by the anomalous increase in the magnitude of the static structure factor at low-Qwith decreasing temperature. Similar anomalous scattering behavior, signifying an increase in the liquid's isothermal compressibility upon cooling, has been observed in experimental and computational studies of supercooled water. Moreover, Bhatia-Thornton mixture analysis revealed that fluctuations in density and local tetrahedral order are strongly coupled in the liquid and that this coupling contributes to anomalous scattering at low-Q. The Ornstein-Zernike correlation length extracted using the Bhatia-Thornton analysis exhibits power-law growth upon cooling and is predicted to diverge at $T \approx 3300$ K, consistent with previous estimates of $T_{\rm C}$.^{61,109} Hence, the fluctuation analysis presented in this study supports the existence of a low-temperature LLCP in mWAC, in agreement with previous computational studies of significantly smaller systems.^{61,109}

Finally, we demonstrated that spontaneous liquid-liquid phase separation is observed when systems with N = 4500, 36,000, and 432,000 atoms are thermally quenched below $T_{\rm C}$ into the two-phase region. These simulations show that mWAC's qualitative behavior is insensitive to system size over a range of N spanning more than two orders of magnitude and thus not strongly influenced by finite-size effects. These findings are similar to those reported in our investigation of the analogous LLPT in the ST2 water model,¹¹⁹ but in sharp contrast to the strong system-size dependent behavior that has been recently observed for TIP5P and TIP4P/2005.⁹³ We anticipate that future investigations will be needed to elucidate the origin of the finite-size effects in TIP5P and TIP4P/2005 and reveal whether these models of water exhibit genuine LLPTs similar to those observed in ST2 and mWAC.

Chapter 4: Hybrid Monte Carlo with LAMMPS

This chapter was previously published by World Scientific: Guo, J.; Haji-Akbari, A.; Palmer, J. C. Hybrid Monte Carlo with LAMMPS. *Journal of Theoretical and Computational Chemistry* **2018**, *17*, 1840002

4.1 Introduction

Molecular dynamics (MD) has become the dominant tool of modern computational statistical mechanics in recent decades due to the emergence of efficient and highly parallel open-source simulation software packages.^{120,125} There exist a number of scenarios, however, in which Monte Carlo (MC) sampling techniques can offer significant advantages over conventional MD.¹²⁶⁻¹²⁹ Slow relaxation processes that frustrate sampling with MD, for example, can be accelerated or even bypassed with cleverly designed MC moves.^{126–129} Whereas the discrete time integration schemes used in MD introduce sampling errors, MC methods are exact in principle and asymptotically sample from the correct statistical ensemble. Hence, MC methods are useful for applications (e.g., free energy calculations) where control over sampling errors is important.^{127,129} There is also greater flexibility in implementing biased sampling schemes with MC. Because the gradient of the bias potential is needed to compute forces in MD, the biased order parameter (OP) must be continuously differentiable with respect to the particle coordinates. By contrast, arbitrary OPs can be biased in MC, including discrete variables (e.g., largest cluster size) such as those widely used in studying nucleation processes.^{127,130}

Comparatively few open-source package have been developed to perform Monte Carlo simulations. Although noteworthy efforts have recently been made to address this issue, ^{131–133} the majority of MC studies are still performed using bespoke codes that are not made publicly available. A number of hybrid methods have been developed,¹³⁴ however, that can leverage the efficiency and flexibility of open-source MD software packages to perform MC sampling. The most common method is hybrid Monte Carlo (HMC),^{135–138} which uses short microcanonical MD trajectories as trial MC moves. Because these moves efficiently relax collective degrees of freedom, HMC sampling has been successfully applied to investigate processes ranging from nucleation^{139–146} to conformational rearrangements in biomolecules and polymers.^{137,147}

The simplest approach to implement HMC with an existing MD package is to use an external script to handle the workflow. In this case, the external script launches the MD software to generate trial trajectories, evaluates the Metropolis acceptance criterion, and performs the necessary file manipulation and bookkeeping tasks before and after each HMC move. Although the simplicity of this approach is appealing, it is generally inefficient because it requires frequent file manipulation and reinitialization of the MD software at the beginning of each trial trajectory. Alternatively, the HMC acceptance-rejection test and supporting routines may be implemented directly in the MD software. This approach is efficient, but it requires significant knowledge of the MD software's inner workings. Additionally, further modification of the MD source code will be needed to implement biased HMC sampling schemes.

Here, we demonstrate a strategy for performing HMC with LAMMPS,^{125,148} one of the most widely used and flexible open-source MD software packages. We describe how HMC can be implemented straightforwardly in both the canonical and isothermal-isobaric ensembles by using a simple Python driver script to execute LAMMPS commands through its built-in library interface and accompanying wrappers.¹⁴⁸ This approach retains the simplicity of using a high-level language to drive the simulations, while avoiding the computational overhead associated with frequent external file manipulation and reinitialization of setup routines. We document a number of rigorous consistency checks that have been used to validate our approach, and we also illustrate how biased HMC sampling schemes can be implemented efficiently by invoking fast Fortran routines from the Python script to compute complex OPs. The codes and sample input files from the documented examples are freely available on the web (https://github.com/palmergroup/hmcwithlammps).

4.2 Hybrid Monte Carlo

4.2.1 Algorithm

A number of variations on the standard HMC algorithm have been proposed in the literature. We focus on the original HMC algorithm proposed by Duane et al.,¹³⁵ which is widely used to simulate condensed matter systems.^{136–147,149,150} This algorithm consists of three basic steps:^{134,135,144}

Step 1. Draw a set of initial velocities from the Maxwell-Boltzmann (MB) distribution

$$P_{\rm MB}(\mathbf{v}) = \prod_{i=1}^{N} \left(\frac{\beta m_i}{2\pi}\right)^{3/2} e^{-\frac{\beta}{2}m_i(\mathbf{v}_i \cdot \mathbf{v}_i)},\tag{4.1}$$

where $\beta = (k_B T)^{-1}$, N is the number of particles in the system, $\mathbf{v} = \{\mathbf{v}_i\}_{i=1}^N$, and \mathbf{v}_i and m_i are the velocity vector and mass of particle *i*, respectively.

Step 2. Propagate a microcanonical MD trajectory to take the system from state $\{\mathbf{v}^{\text{old}}, \mathbf{r}^{\text{old}}\}$ to $\{\mathbf{v}^{\text{new}}, \mathbf{r}^{\text{new}}\}$, where $\mathbf{r} \equiv \{\mathbf{r}_i\}_{i=1}^N$ is the complete set of particle coordinates. The length of the trajectory $\Delta t = n_{\text{steps}} \times \delta t$ is specified by the number of integration steps n_{steps} and time step δt .

Step 3. Accept or reject the new configuration \mathbf{r}^{new} according to the Metropolis criterion

$$P_{\rm acc}^{\rm old \to new} = \min\left(1, e^{-\beta \Delta \mathscr{H}}\right),$$
 (4.2)

where $\Delta \mathcal{H} = \mathcal{H}^{\text{new}} - \mathcal{H}^{\text{old}}$ is the change in the system's Hamiltonian during the MD trajectory. The Hamiltonian is defined as the sum of kinetic and potential energy $\mathcal{H}(\mathbf{v}, \mathbf{r}) \equiv K(\mathbf{v}) + U(\mathbf{r})$, with $K(\mathbf{v}) \equiv \sum_{i=1}^{N} \frac{\beta}{2} m_i (\mathbf{v}_i \cdot \mathbf{v}_i)$.

Repeated application of Steps 1–3 will generate a Markov chain that asymptotically samples configurations from the canonical (NVT) ensemble (i.e., $\rho(\mathbf{r}) \propto e^{-\beta U(\mathbf{r})}$).^{134,135} Sampling from the isothermal-isobaric (NpT) ensemble can be achieved by simply combining the HMC algorithm with standard volume change MC moves.^{144,151}

4.2.2 Implementation notes

We have implemented HMC using the LAMMPS library interface and its accompanying Python wrapper (LAMMPS-17Nov16 release¹⁴⁸). The wrapper provides a low-level Python interface to the library that allows users to create instances of LAMMPS, invoke LAMMPS commands, perform MD, extract results, and modify simulation parameters. Thus, all of the HMC workflow is handled through a custom Python driver script, which invokes LAMMPS commands to generate trial MD trajectories, performs the Metropolis acceptance-rejection test, and intermittently dumps simulation output to file. The user only needs to provide a LAMMPS data file containing an initial configuration for the system and a standard input file for performing a microcanonical MD simulation that specifies the potential functions, integrator, etc. The LAMMPS commands and Python wrapper used by our HMC code are well-documented.¹⁴⁸ Accordingly, here we provide only a short description of the major component of our code.

Velocity initialization (Step 1) is performed in the Python driver script by assigning components of \mathbf{v}_i using

$$v_{i,j} = \left(\beta m_i\right)^{-1/2} \mathcal{N}(0,1)$$
 (4.3)

for j = x, y, z, where $\mathcal{N}(0, 1)$ is a standard Gaussian random number generated via Python's random module. The resulting velocities obey MB statistics (Eq. 4.1). The particle masses m_i needed for this step are imported into Python from LAMMPS, along with other system parameters such as N. After initialization, the velocities are passed to LAMMPS using the scatter.atoms command¹⁴⁸ provided by the Python wrapper.

Propagation of each trial MD trajectory (Step 2) is performed by using the Python wrapper's command function to pass the run directive to LAMMPS. This directive invokes a time-reversible and volume-preserving velocity-Verlet integrator (fix nve in LAMMPS¹⁴⁸). These integrator attributes are required to perform HMC sampling using the standard acceptance criterion (Eq. 4.2). The choice of integrator is specified in the user-generated input file, whereas parameters δt and n_{steps} are set in the Python driver.

The Metropolis acceptance-rejection test (Step 3; Eq. 4.2) is performed in the Python driver script. The Hamiltonian \mathcal{H} is evaluated from values of K and U extracted from LAMMPS using the Python wrapper's extract_compute command.¹⁴⁸ The acceptance-rejection step also requires passing particle coordinates to and from LAMMPS. These exchange operations are performed in the driver script using the Python wrapper commands scatter.atoms and gather.atoms,¹⁴⁸ respectively.

Sampling in the isothermal-isobaric ensemble is performed by combining the HMC algorithm with MC moves that attempt to change the logarithm of the simulation cell volume from $\ln V$ to $\ln V + \Delta(\ln V)$ (see Algorithm 11 in Ch. 5 of Ref. 151). For each *MC sweep*, the driver script executes one attempted HMC or log-volume change move, the type of which is selected at random. The HMC and log-volume moves are attempted with probabilities P_{HMC} and $1 - P_{\text{HMC}}$, respectively. The parameter P_{HMC} , the size of the maximum allowed log-volume change $\Delta(\ln V)_{\text{max}}$, and the number of MC sweeps n_{sweeps} are specified in the driver script. Adjustments to the cell dimensions for the log-volume moves are performed by using the Python wrapper's command function to pass the change_box directive¹⁴⁸ to LAMMPS. Particle coordinate exchanges with LAMMPS are executed using the same commands

employed in the HMC moves. The current implementation of the log-volume moves can only be used with cubic simulation cells.

4.2.3 Model and sampling protocol

Code validation tests were performed using the Rowley, Nicholson, and Parsonage Lennard-Jones model for argon ($\epsilon_{Ar}/k_B = 119.8$ K and $\sigma_{Ar} = 0.3405$ nm).¹⁵² Interactions between argon molecules were calculated using a force-shifted Lennard-Jones potential (pair_style lj/smooth/linear in LAMMPS¹⁴⁸) with a cut-off distance of $2.5\sigma_{Ar}$. Canonical ensemble HMC simulations of a system containing N = 512 molecules were performed using $\delta t = 30$ fs and $n_{\text{steps}} = 10$. For isothermalisobaric HMC simulations, these parameters were used along with $P_{\text{HMC}} = 0.8$ and $\Delta(\ln V)_{max} = 0.04$. Depending on the state conditions, this sampling protocol yields typical acceptance rates between 60 and 80 % for the HMC and log-volume moves. For comparison, we also performed MD simulations of this system with LAMMPS using $\delta t = 8$ fs. For canonical ensemble MD simulations, the system temperature is maintained using a Nosé-Hoover (NH) thermostat^{95,96} (fix nvt in LAMMPS¹⁴⁸). Simulations in the isothermal-isobaric ensemble were performed by combining the NH thermostat with a Martyna-Tuckerman-Tobias-Klein (MTTK) barostat¹⁵³ (fix npt in LAMMPS¹⁴⁸). The relaxation time constants for the thermostat and barostat were set to 1 and 5 ps, respectively.

The duration of each simulation was measured in terms of the statistical inefficiency $g_A \equiv 1 + 2\tau_A$, where $\tau_A \equiv \int_0^\infty C_A(t)dt$ is the integrated auto-correlation time for property A computed from the normalized auto-correlation function $C_A(t) = (\langle A(t)A(0) \rangle - \langle A \rangle^2)/(\langle A^2 \rangle - \langle A \rangle^2)$. The value of g_A quantifies the number of correlated time series samples required to generate a single effectively uncorrelated sample for an observable A.¹⁵⁴ For sampling in the canonical ensemble, we use time series data for the potential energy U from the post-equilibrium, production phase of each simulation to compute g. For isothermal-isobaric simulations, we report $g = \max(g_U, g_\rho)$,
where ρ is the mean density of the system. All simulations were equilibrated for at least $3 \times 10^3 g$, following by a production period $\geq 2 \times 10^4 g$ in duration. To compute statistical uncertainties, we estimate the number of effectively uncorrelated samples using $n_{\text{samp}} = t_{\text{sim}}/g$, where t_{sim} is the duration of the simulation's production phase.

4.2.4 Validation

We performed a number of consistency checks to validate our HMC implementation. In the first test, we compare liquid-state properties along three isotherms (T = 89.84, 119.80, and 179.69 K) computed using HMC with the equation of state data generated from MD simulations performed with LAMMPS (Fig. 4.1). Average liquid densities $\langle \rho \rangle$ from the isothermal-isobaric ensemble HMC and MD simulations agree within statistical uncertainty at all conditions examined. Similar agreement between HMC and MD is observed in comparing structural properties computed for the liquid, such as the radial distribution function (not shown). Further, the predicted equation of state agrees with published reference data for the Lennard-Jones fluid.¹⁵⁵

Although consistency checks based on equation of state comparisons should always be conducted, they can be misleading if not performed carefully over a broad range of state conditions. Further, comparisons with literature data are ambiguous if statistical uncertainties are not reported. Consequently, these checks may fail to detect subtle implementation errors that can significantly affect free energy calculations, for example. Fortunately, there are often rigorous quantitative consistency checks that can be performed to ensure that a code is functioning properly. For pedagogical purposes, we describe two of these checks.

The HMC method has a built-in consistency check.¹³⁶ As discussed in Section 4.2.2, the standard HMC algorithm requires using a time-reversible and volume-preserving integrator to propagate trial MD trajectories. Invoking the volume-preserving property, it is straightforward to show that an HMC sampling algorithm



Figure 4.1: Equation of state data for liquid argon calculated using HMC and MD (filled and open symbols, respectively). Excellent agreement between HMC and MD is observed at all conditions examined. Statistical uncertainties are smaller than the symbols.

should satisfy the normalization condition:¹³⁶

$$\langle e^{-\beta \Delta \mathcal{H}} \rangle = 1, \tag{4.4}$$

where the average is taken over the attempted HMC moves. This rigorous statistical mechanical relationship should be obeyed independent of the chosen sampling parameters. Accordingly, we find that our HMC implementation satisfies the normalization condition, within statistical uncertainty, for all choices of the integration time step δt (Fig. 4.2). Similarly, we find that Eq. 4.4 is satisfied when other sampling parameters (e.g., n_{steps}) are varied.

Systematic sampling errors can also be detected through fluctuation analysis. In the canonical ensemble, for example, the log ratio of the potential energy distributions at temperatures $\beta_1 = (k_B T_1)^{-1}$ and $\beta_2 = (k_B T_2)^{-1}$ should vary linearly with U:

$$\ln\left[\frac{P(U|\beta_2)}{P(U|\beta_1)}\right] = -(\beta_2 - \beta_1)U + \text{const.}.$$
(4.5)



Figure 4.2: Test of the HMC normalization condition (Eq. 4.4). Normalized deviations $(\langle e^{-\beta\Delta \mathscr{H}} \rangle - 1) \times \sigma^{-1}$ computed using the estimated uncertainty σ standard error) are less than 1 and thus illustrate that Eq. 4.4 is satisfied within statistical error.

Hence, if canonical sampling is performed at two nearby temperatures, the slope of a linear fit line to the log ratio of the resulting potential energy distributions should be equal to $-(\beta_2 - \beta_1)$ within statistical uncertainty. The challenge with this test lies in determining whether deviations from the expected slope are statistically significant. To address this issue, Shirts¹⁵⁶ has developed the checkensemble package,¹⁵⁷ which provides Python routines to implement this consistency check and estimate statistical uncertainties.

We used the canonical ensemble consistency test (Eq. 4.5) implemented in checkensemble to validate our HMC sampling algorithm. Data for this test were generated from HMC simulations performed at $T_1 = 132.0$ K and $T_2 = 145.5$ K and a fixed density of 0.451 g cm⁻³. Following the recommendation of Shirts, the temperature gap was chosen such that $\Delta T^{\text{true}} = T_2 - T_1 \approx T_1(2k_B/C_V^R)^{1/2}$, where C_V^R is the residual heat capacity estimated at T_1 . For comparison, the same consistency check was also conducted using data from canonical ensemble MD simulations performed with LAMMPS. The production phases of the HMC and MD simulations were long enough to generate $\ge 3 \times 10^4$ uncorrelated samples at each thermodynamic condition.

The effective temperature gap $\Delta T^{\text{obs.}} \pm \sigma = 13.421 \pm 0.059$ K estimated from the slope of the fit to the HMC data is within statistical uncertainty of $\Delta T^{\text{true}} = 13.5$ K (Fig. 4.3). In Fig. 4.3, the HMC simulation data, fit to the data, and theoretical prediction based on Eq. 4.5 are shown as points, a dash lined, and solid line, respectively. Here, σ is the standard deviation computed from the covariance matrix for the fit parameters. It provides a measure of the certainty in the error between $\Delta T^{\text{obs.}}$ and ΔT^{true} , ¹⁵⁶ with $\Delta T^{\text{obs.}} \pm 2\sigma$ signifying approximately the 95% confidence interval for $\Delta T^{\text{obs.}}$ (assuming the fitting errors are normally distributed). Deviations $|\Delta T^{\text{obs.}} - \Delta T^{\text{true}}|$ consistently larger than $2 - 3\sigma$ would therefore indicate the presence of systematic sampling errors. For our HMC code, however, we find that $|\Delta T^{\text{obs.}} - \Delta T^{\text{true}}|$ is only 0.079 K or equivalently 1.34 σ . The value of $\Delta T^{\text{obs.}}$ is thus statistically indistinguishable from ΔT^{true} , which suggests that our HMC algorithm correctly samples fluctuations in U in accord with Eq. 4.5. Similar agreement is found when analyzing the MD simulations data (not shown), which yields $\Delta T^{\text{obs.}} \pm \sigma = 13.504 \pm 0.068$ K or a deviation of only 0.06 σ .

The test reported above uses histogram data to check for ensemble consistency (Fig. 4.3). Although histograms are convenient for visualizing data, the results may be sensitive to the choice of bin width. As a result, we also conducted ensemble consistency tests using the maximum likelihood estimation (MLE) approach in checkensemble. The MLE approach performs parameter estimation directly from the time series data and thereby avoids potential systematic errors associated with analyzing histograms. We used MLE to perform the canonical ensemble consistency test (Eq. 4.5). We also conducted two additional MLE consistency checks to validate our isothermal-isobaric HMC sampling algorithm. The first check is based on analyzing fluctuations in enthalpy, $H \equiv U + pV$, at temperatures β_1 and β_2 , which should



Figure 4.3: Histogram-based canonical ensemble consistency test (Eq. 4.5) performed on HMC simulation data for Lennard-Jones argon.

obey

$$\ln\left[\frac{P(H|\beta_2,p)}{P(H|\beta_1,p)}\right] = -(\beta_2 - \beta_1)H + \text{const.}$$
(4.6)

for $p_1 = p_2 = p$. The second check tests whether volume fluctuations at pressures p_1 and p_2 follow

$$\ln\left[\frac{P(V|\beta, p_2)}{P(V|\beta, p_1)}\right] = -\beta(p_2 - p_1)V + \text{const.}$$

$$(4.7)$$

for $\beta_1 = \beta_2 = \beta$. These expressions have the same linear form as the relationship presented for the canonical ensemble (Eq. 4.5). Accordingly, they can be applied in an analogous manner to validate isothermal-isobaric sampling algorithms.

Our HMC implementation satisfies all three ensemble consistency checks performed using the MLE approach in checkensemble (Table 1). In agreement with the histogram based analysis, the MLE test demonstrates that the potential energy distributions generated by HMC simulations at 132.0 and 145.5 K are consistent with the canonical ensemble (Table 1; Eq. 4.5). Similarly, the enthalpy fluctuations sampled at $T_1 = 125.0$ K and $T_2 = 127.8$ K and p = 100 atm using HMC are consistent with the isothermal-isobaric distribution function (Table 1; Eq. 4.6); the temperature gap for this test was chosen by substituting the constant-pressure residual heat capacity C_p^R for C_v^R in the criterion presented above. The distributions of U and H from MD simulations with LAMMPS also satisfy the MLE based ensemble consistency tests (Table 1).

The volume distribution test (Eq. 4.7) was performed at 125 K using $p_1 = 100$ atm and $p_2 = 120$ atm. The pressure gap was chosen using the criterion suggested by Shirts: ¹⁵⁶ $\Delta p^{\text{true}} = p_2 - p_1 \approx (2k_BT/V\kappa_T)^{1/2}$, where κ_T is the isothermal compressibility at the lower pressure p_1 . The effective pressure gap estimated from the HMC data deviates from the true value by ca. 5σ (Table 1). This deviation is more statistically significant than those observed when analyzing potential energy and enthalpy fluctuations. We have confirmed that similar deviations are observed when identical tests are performed with other MC codes that employ standard single-particle moves and volume displacements in V rather than $\ln V$ (see Algorithms 2 and 10 in Ref. 151, respectively). For comparison, a deviation of 8.44 σ is found when MD simulations of argon are performed using the MTTK barostat (Table 1). Shirts¹⁵⁶ conducted similar tests with this model of argon at nearby conditions to analyze the ensemble consistency of the MTTK, Parrinello-Rahman, and Berendsen pressure control algorithms for MD. Deviation of 9.0, 9.5, and 17.1 σ were reported for the three barostats, respectively. The MTTK barostat was therefore found to be the most ensemble consistent algorithm, with a reported deviation similar to the value obtained from our analysis (Table 1). Hence, in practice, all of the pressure control algorithms seem to be slightly off in sampling volume fluctuations consistent with the isothermalisobaric ensemble. The MLE tests reveal, however, that our HMC sampling scheme is generally comparable to or better than state-of-the-art MD algorithms in sampling from the canonical and isothermal-isobaric ensembles, which indicates that the code functions as intended.

We have validated our HMC sampling algorithm using several rigorous quan-

titative consistency checks. The fluctuation analysis tests performed with the checkensemble package are general and can be applied to scrutinize any MD or MC code. Our experience suggests that they are extremely effective in detecting subtle implementation or user errors that affect the ensemble consistency of sampling algorithms. We therefore recommend that users also conduct these tests as an initial starting point for learning how to use open-source simulation software packages.

$U^a (\Delta T^{\text{true}} = 13.5 \text{ K})$	$\Delta T^{\text{obs.}} \pm \sigma$ (K)	Deviation ^{b} (σ)
HMC	13.502 ± 0.058	0.03
MD	13.580 ± 0.067	1.19
H^c ($\Delta T^{\text{true}} = 2.80 \text{ K}$)	$\Delta T^{\text{obs.}} \pm \sigma$ (K)	Deviation (σ)
HMC	2.793 ± 0.025	0.27
MD	2.763 ± 0.130	2.91
V^d ($\Delta p^{\text{true}} = 20 \text{ atm}$)	$\Delta p^{\text{obs.}} \pm \sigma \text{ (atm)}$	Deviation (σ)
HMC	19.249 ± 0.149	5.04
MD	19.333 ± 0.079	8.44

Table 1: Results of ensemble consistency tests

Note: Results are reported as output by checkensemble, without modifying the number of significant figures

 $^a~T_1$ = 132.0 K, T_2 = 145.5 K, and ρ = 0.451 g cm $^{-3}$

 b Deviation is defined as $|{\rm true-observed}|\times\sigma^{-1}$ and has units of σ

 c T_{1} = 125.0 K, T_{2} = 127.8 K, and p = 100 atm

 d $p_{1} = 100$ atm, $p_{2} = 120$ atm, and T = 125 K

4.3 Free Energy Calculations

4.3.1 Biased sampling

Our HMC code can also be easily adapted to implement biased sampling schemes for performing free energy calculations. Rare events such as chemical reactions, phase transitions, and protein folding, for example, are characterized by infrequent transitions between long-lived metastable states. They are commonly analyzed in simulation by computing the Landau free energy:¹³⁰

$$\beta F(\phi) = -\ln P(\phi) + \text{const.}, \tag{4.8}$$

where *P* is the equilibrium probability density function associated with the order parameter (OP) $\phi(\mathbf{r})$. The parameter ϕ is chosen to distinguish between the metastable states of interest; if ϕ also describes the dynamical progress along the rare event transition pathway, then it is a special type of OP known as a reaction coordinate.¹³⁰

Direct calculation of $F(\phi)$ through unbiased simulations is almost always impractical due to the presence of large free energy barriers ($\gg k_B T$) along ϕ , which separate the metastable basins of interest. Consequently, biased sampling schemes are typically required to converge $F(\phi)$ calculations over the full range of ϕ relevant to the characterization of the rare event. These schemes can be readily implemented with HMC by evaluating $\Delta \mathcal{H}$ in the Metropolis acceptance-rejection test using a modified potential energy function:

$$U^{\mathsf{w}} \equiv U + \eta(\phi). \tag{4.9}$$

The weight function $\eta(\phi)$ is chosen to enhance sampling along ϕ . An optimal choice of $\beta\eta(\phi) = -\beta F(\phi)$ will offset the underlying free energy landscape and result in uniform sampling. Of course, other choices are necessary in practice because $F(\phi)$ is not known *a priori*. Irrespective of the choice of $\eta(\phi)$, however, the free energy can always be obtained using

$$\beta F(\phi) = -\ln P^{W}(\phi) - \beta \eta(\phi) + \text{const.}, \qquad (4.10)$$

where $P^{w}(\phi)$ is the probability density function sampled during the biased simulation.

The OP, ϕ , is often a complex function of the particle coordinates. Consequently, it

can be computationally expensive to compute, particularly with Python, which is not as natively efficient as languages such as Fortran and C. To address this issue, our HMC implementation invokes routines from a custom Fortran library to compute ϕ from the particle coordinates passed by the Python driver script. The value of ϕ returned to the driver script is used to calculate the weight function η for the Metropolis acceptance-rejection test. The Python interface to the Fortran library is generated using the package F2PY.¹⁵⁸ A similar approach can be employed to invoke fast C or C++ routines from Python to compute ϕ .

4.3.2 Nucleation of mW water

To illustrate the approach described above, we use biased simulations to drive homogeneous ice nucleation in the coarse-grained mW water model.¹⁵⁹ The mW model treats water molecules as point particles that interact through a three-body Stillinger-Weber (SW) potential¹⁶⁰ (pair_style sw in LAMMPS¹⁴⁸), which is short ranged and fast to compute even for large systems. Despite its coarse-grained nature, the mW model is remarkably accurate in predicting bulk properties of liquid water near ambient conditions.¹⁵⁹ Because mW crystallizes rapidly compared to atomistic models,^{74,161–163} it has also been widely used to study ice nucleation in supercooled water.^{26,77,141,164,165} Thus mW is a well-studied and computationally efficient model for demonstrating the capabilities of biased sampling algorithms.

We use a biased HMC sampling scheme to compute the free energy $F(n^*)$, where n^* is the size of the largest crystalline cluster. Reinhardt and Doye⁷⁷ performed similar calculations for mW using umbrella sampling¹⁶⁶ with standard single-particle MC moves. Following their study, we compute $F(n^*)$ for a system of N = 1400 mW molecules at 220 K and 1 bar, which corresponds to a supercooling of ca. 20 %. The parameter n^* is evaluated by using bond order parameters to classify molecules as either ice- or liquid-like based on their local environment. A molecule is considered ice-like if it forms a sufficient number of crystalline connections with its neighbors.

The number of connections for molecule *i* is determined using

$$n_{\rm con}(i) = \sum_{j=1}^{N_b(i)} \Gamma(d_l(i,j)), \tag{4.11}$$

where $N_b(i)$ is the number of nearest neighbors and

$$d_l(i,j) \equiv \frac{\mathbf{q}_l(i) \cdot \mathbf{q}_l(j)}{|\mathbf{q}_l(i)| |\mathbf{q}_l(j)|} \tag{4.12}$$

is the scalar product of the complex vectors $\mathbf{q}_l \in \mathbb{C}^{2l+1}$, the components of which are given by:

$$q_{l,m}(i) \equiv \frac{1}{N_b(i)} \sum_{i=1}^{N_b(i)} Y_{l,m}(\mathbf{r}_{ij}), \quad -l \le m \le l$$
(4.13)

where $Y_{l,m}$ are the spherical harmonics and $\mathbf{r}_{ij} = \mathbf{r}_j - \mathbf{r}_i$ is the displacement vector from particle *i* to particle *j*. As in Ref. 77, we use l = 3. Similarly, $\Gamma(\cdot)$ is given by:

$$\Gamma(x) = \begin{cases} 1 & \text{if } [(x < -0.82) \lor (-0.145 < x < -0.065)] \\ 0 & \text{otherwise.} \end{cases}$$
(4.14)

Nearest neighbors are identified with a distance cutoff of $r_c = 0.36$ nm. Water molecules with $n_{\rm con} \ge 3$ are classified as ice-like. Clusters of ice-like molecules are identified using a single-linkage algorithm based on the neighbor cutoff distance. A detailed justification of these choices is given in Ref. 77.

The free energy $F(n^*)$ is computed by performing independent simulations in overlapping regions or windows along n^* . Rather than specifying a fixed functional form for the weight function, we used the Wang-Landau (WL) algorithm¹⁶⁷ to optimize $\eta(n^*)$ in each sampling window. Our implementation of the WL algorithm closely follows the scheme described in Appendix A of Ref. 168. After each MC trial move, a histogram representation of the weight function is updated using $\eta(k) = \eta(k) + \ln f$, where $\ln f$ is an update factor and k is the histogram bin corresponding to the value of n^* after the move. In addition, a states-visited histogram, h, is also updated after each move via h(k) = h(k) + 1. This procedure continues until h is sufficiently flat or uniform. The criterion min $[h] = s\bar{h}$ is used to assess flatness, which requires that the minimum value in h be within a factor s of the histogram mean \bar{h} . Once the flatness criterion is satisfied, h is set to zero and the update factor is decreased using $\ln f_{\text{new}} = \alpha \ln f_{\text{old}}$, where α is a multiplicative factor between zero and one. This process is repeated until $\ln f$ drops below a predetermined value. Harmonic walls were also placed at the upper and lower bound of each window $(n_{\text{L}}^*$ and n_{U}^* , respectively) to restrict sampling to the targeted region. Consequently, if the system moved beyond the specified bounds of the sampling window, it was subjected to a bias potential of the form $\eta_{\text{L},\text{U}}(n^*) = k_{\text{wall}}(n^* - n_{\text{L},\text{U}}^*)^2/2$, where k_{wall} is the spring constant.

The WL simulations were conducted using HMC parameters $\delta t = 6$ fs, $n_{\text{steps}} = 15$, $P_{\text{HMC}} = 0.8$, and $\Delta(\ln V)_{\text{max}} = 0.025$. These conservative choices of δt and $\Delta(\ln V)_{\text{max}}$ were necessary to ensure that acceptance rates above ca. 25 % were maintained for the HMC and log-volume moves during the biased simulations. Parameters s = 0.8 and $\alpha = 0.5$, along with an initial value of $\ln f = 0.2$, were chosen for the WL algorithm, and a spring constant of $k_{\text{wall}} = 300 \ k_B T$ was used for the harmonic walls. Updates to the bias potential were performed using the WL algorithm until the value of $\ln f$ fell below 5×10^{-3} . Once this threshold was reached, the WL updates were stopped and the simulations were allowed to continue sampling using the converged bias potential. Time series data from this final portion of each simulation were analyzed using the multistate Bennett acceptance ratio method ¹⁶⁹ (MBAR) to obtain an unbiased estimate of $F(n^*)$.

Example simulation trajectories show that the WL algorithm promotes nearuniform sampling in each window (Fig. 4.4A). In Figure 4.4A, time is reported in units of MC sweeps, where a sweep is defined as one attempted HMC or log-volume MC move. Harmonic walls were applied at $n_{\rm L}^* = 1$ and $n_{\rm U}^* = 25$ to restrain the simulations to the targeted sampling region. The adaptive bias compensates for barriers in the underlying free energy landscape and drives the system towards under-explored regions of the OP space. After several WL iterations, the bias converges such that $\beta\eta(n^*) \approx -\beta F(n^*)$. Under these conditions, the free energy landscape is almost completely offset by the bias, and the system can freely diffuse across the range of n^* defined by the sampling window. By contrast, unbiased HMC simulations are only able to explore the lower region of the sampling window due to the presence of free energy barriers along n^* (Fig. 4.4A). Adaptive biasing using the WL algorithm thus significantly enhances sampling of the OP space.

The free energy $F(n^*)$ computed from the WL HMC simulations is in good agreement with independent calculations performed by Reinhardt and Doye⁷⁷ (Fig. 4.4B). Both estimates of $F(n^*)$ exhibit a maximum at approximately $n^* = 114$ molecules, which is the critical nucleus size. The barrier to nucleation, or height of the maximum, predicted by our calculations is $\sim 22 \pm 1 \ k_B T$. The shaded region in Fig. 4.4B indicates statistical uncertainties in $F(n^*)$. Although this value is slightly lower than the barrier of $\sim 24 \ k_B T$ predicted by Reinhardt and Doye,⁷⁷ the two estimates are in statistical agreement if one assumes uncertainties of $\sim 1 \ k_B T$ for their calculations.

As discussed by Reinhardt and Doye,¹⁷⁰ convergence of the biased simulations can be verified by computing the enthalpy as a function of the OP,

$$\langle H(n^*)\rangle = \frac{\langle H(n^*)e^{\beta\eta(n^*)}\rangle_{\rm w}}{\langle e^{\beta\eta(n^*)}\rangle_{\rm w}},\tag{4.15}$$

where the ensemble averages on the right side of the equation are computed using data from the biased simulations. Because enthalpy is an absolute quantity, the values of $\langle H(n^*) \rangle$ computed in adjacent sampling windows should be consistent in the regions of overlap. Significant hysteresis in $\langle H(n^*) \rangle$ would therefore indicate poor



Figure 4.4: A. Trajectory from an HMC simulation that was adaptively biased using the WL algorithm (dashed line). A sample trajectory from an unbiased HMC simulation (solid line) is also shown for comparison. B. Free energy $F(n^*)$ computed from biased WL HMC simulations (dashed line) are in good agreement with previous calculations performed by Reinhardt and Doye⁷⁷ (solid line). C. Average enthalpy $\Delta \langle H(n^*) \rangle = \langle H(n^*) \rangle - \langle H(2) \rangle$ computed as a function of n^* using data from the biased WL HMC simulations.

convergence of the simulations. As expected, the enthalpy computed from the WL HMC simulations is consistent across the sampling windows (Fig. 4.4C). The data in Fig. 4.4C have been shifted such that the maximum value at $n^* = 2$ is zero. Data from different sampling windows are denoted using different symbols. Consistency in regions of overlap illustrates that the order parameter space is well-sampled. The agreement in regions of overlap is slightly better than observed in similar calculations performed by Reinhardt and Doye (see Fig. 6.6 in Ref. 170), which may explain the minor discrepancies between our respective free energy calculations (Fig. 4.4C). We anticipate that the HMC moves used in our sampling algorithm may be more efficient in equilibrating slow, collective degrees of freedom in the system than the standard single-particle MC moves used in their study. Nevertheless, the agreement between the two studies is satisfactory, considering the challenging nature of these calculations.

4.4 Conclusions

We have documented a successful strategy for performing canonical and isothermal-isobaric ensemble HMC simulations with the widely-used LAMMPS MD software package. The HMC simulations are driven by an external Python script that manages the overall workflow and invokes commands through LAMMPS' built-in library interface to propagate trial MD trajectories. This approach is general and also computationally efficient because it utilizes LAMMPS' highly optimized routines to perform numerically intensive tasks such as force calculations. We have also shown that our HMC implementation satisfies several stringent consistency tests based on rigorous statistical mechanical relationships. These tests are generally useful for validating molecular dynamics and Monte Carlo sampling algorithms, and they can also help detect even subtle user errors that may affect the ensemble consistency of simulations. Finally, we have demonstrated that our code can be easily adapted to perform biased HMC simulations in which complex OPs are calculated using custom Fortran or C/C++ routines. As an illustrative example, we used an adaptive bias scheme based on the Wang-Landau algorithm to successfully compute the free energy barrier to homogeneous ice nucleation for the mW water model at a supercooling of ca. 20 %. The results from these calculations were found to be in good agreement with those previously reported in the literature. The codes and input files from the examples described in this study are freely available on the web.¹⁷¹

Although the algorithms described in this study are appropriate for point particles, we note that some modifications must be made to simulate polyatomic molecules. For flexible polyatomic molecules, with no internal bond or angle constraints (e.g., bead-spring polymers with harmonic bonds), only the algorithm for performing volume change MC moves needs to be modified. Specifically, as described in standard texts,¹⁵¹ the volume change moves should be conducted by rescaling the position of each molecule's center of mass, rather than independently rescaling the position of each atomic site. For rigid polyatomic molecules, Steps 1 and 2 of the HMC algorithm must also be modified. Care must be taken to account for internal constrains within rigid molecules when generating initial velocities from the Maxwell-Boltzmann distribution in Step 1 (see Supporting Information of Ref. 143). Initial velocities may be drawn from other distributions, but the HMC acceptance criterion (Eq. 4.2) should be reformulated accordingly to prevent detailed balance violations.¹⁷² A volume-preserving and time-reversible integrator for rigid bodies must also be used to propagate the trial MD trajectories in Step 2. An appropriate integrator¹⁷³ is distributed with LAMMPS (fix rigid/nve¹⁴⁸), but the current implementation does not allow the rigid body velocities to be modified via LAMMPS' built-in library interface. As a result, LAMMPS' source code must be modified to perform HMC simulations using this integrator. Alternatively, holonomic constraint $algorithms^{174,175}$ (e.g., fix shake and fix rattle in LAMMPS¹⁴⁸) may be used to integrate the equations of motion for rigid polyatomic molecules. Unfortunately,

these algorithms are not strictly time reversible¹⁷³ because of the inexact nature of the iterative numerical procedures that are used to solve the constraint equations. Destroying time reversibility will lead to detailed balance violations and introduce sampling errors. In many cases, these errors may be negligible if the solutions to the constraint equations are fully converged. Nevertheless, we urge caution and suggest that users apply the consistency tests in Sec. 4.2.4 to detect potential sampling errors and identify specific cases where these algorithms may be used to perform HMC sampling.

Chapter 5: Conclusions and Future Work

5.1 Conclusions

In first two publications, we investigated the low-temperature scattering phenomena of two computer models of tetrahedral substances: the ST2 model of water and the mWAC model of silica. Both models have previously been verified to exhibit a genuine LLPT between a LDL and a HDL through free energy calculations.^{28,55} In order to accommodate enhanced fluctuations near the LLCP, large-scale molecular dynamics simulations (N = 32,000 ST2 water molecules and N = 12,000 mWAC ions of SiO₂) were performed along critical isochore approaching the estimated critical po int. Because both liquid phases are distinguished by their densities and local tetrahedral order, we performed the fluctuation analysis by first labeling each molecule based on the extent of local tetrahedral order. Then, we analyzed the density, local tetrahedral order and their coupled fluctuations resulting from the interplay between two different molecular motifs.

The static structure factor in both mWAC model of silica and ST2 model of water become enhanced at low wavenumber region upon cooling. Such enhancement corresponds to the increase of density fluctuation, and hence the increases of isothermal compressibility, which is a well-known anomaly of tetrahedral liquids. This behavior is in accordance with the experimental observation of static structure factor of supercooled water at the low wavenumber region. The distributions of local tetrahedral order parameter for both models are bimodal, suggesting the existence of two populations of molecular species. Bhatia-Thornton mixture analysis indicated that fluctuations in density and local tetrahedral order are strongly coupled, and are responsible for the anomalous scattering behaviors at the low wavenumber region. The length scale with which these fluctuations are correlated were analyzed in the framework of Ornstein-Zernike theory. The Ornstein-Zernike correlation length corresponding to the anomalous scattering components exhibit power-law growth upon cooling and is predicted to diverge at $T \approx 245K$ and $T \approx 3300K$ for ST2 and mWAC, respectively. These estimates of T_c are consistent with previous estimates of T_c from equation of state and free energy calculations. Consequently, current fluctuation analysis is consistent with the existence of a low-temperature LLCP in both mWAC and ST2. Finally, for both models, we also demonstrated that the LLPT can be directly observed in simulations by quenching the system into the two-phase coexistence region below the estimated critical point. Different system sizes were used to check finite-size effects, and for all system size investigated, the LLPT can be observed and the liquids remain free of signs of crystallization.

In the third publication, we presented a successful strategy for performing HMC simulations by interfacing Python with the LAMMPS. The overall workflow is managed by a Python script and the numerically intensive calculations such as force calculations are handled by LAMMPS's bulit-in library interface, which is optimized by low-level languages such as C++. In general, our implementation combines easy-to-use syntax from high-level language, and numerically efficiency afforded by low-level languages. To validate our implementation, we performed several stringent consistency checks based on rigorous statistical mechanical principles. These tests are generally useful and able to detect subtle user errors that may cause sampling problems. To further demonstrate the capabilities of our approach, we perform biased HMC simulations of mW water model, in which complex OPs are calculated using custom Fortran or C/C++ routines. Simulations were driven by an adaptive bias scheme based on the Wang-Landau algorithm to compute the free energy barrier of homogeneous ice nucleation at a supercooling of ca. 20 %. The resulting free energy surface is in good agreement with those previously reported in the literature.

We note that some modifications must be made in our implementations of HMC code to simulate polyatomic molecules.

5.2 Future Work

5.2.1 Future Work 1

As a plausible thermodynamic scenarios to resolve water's low temperature anomalies, the LLCP hypothesis was proposed and subsequently validated in ST2 model of water.^{12,28} Ever since, many atomistic models of water have been investigated for the existence of LLCP.^{24,29,31,176} Recent evidence from fluctuation analysis, similar to that performed here for ST2 and mWAC, suggests that two realistic models of water TIP4P/2005 and TIP4P/Ice, exhibit a low-temperature LLPT. This finding indicates that LLCPs may be general features of classical water models.¹⁷⁶ Current search for the evidence of LLCP in an atomistic models of water is still computationally expensive even with the assistance of advanced sampling methods.^{28,176} Coarse-grained (CG) models of water are usually less computationally expensive than atomistic model of water due to the reduced number of interaction sites and short-ranged nature of their interaction.⁷¹ Moreover, the study of coarse-grained water models can provide important molecular insights crucial for understanding the LLPT phenomena.^{41,177} The mW model of water is a widely used coarse-grained water model. Previous investigations suggested that the mW water becomes unstable with respect to the ice crystallization at deeply supercooled conditions, revealing no signs of the metastable LLPT.²⁶ Recently, Chan et al. introduced a machine learning framework to parameterize a three-body potential for water that effectively incorporates bond-order dependence on the local environment.^{177,178} Both the experimental and simulation data generated from on-the-fly molecular dynamic simulations of TIP4P/2005 water are used to train the bond-order potentials. The resulting family of machine-learned bond-order (ML-BOP) CG models of water quantitatively predict

many experimental properties of water such as TMD and melting point at ambient pressure.¹⁷⁷ The ML-BOP models were found to be more accurate than the mW model and comparable to the best classical atomistic water models, thus promising to further bridge the length and time scales between the simulation and the experiment.¹⁷⁷ Consequently, ML-BOP models of water will play an important role in understanding the peculiar properties of water. To this end, a potential future avenue for study is to investigate the low-temperature phase behavior of the ML-BOP models by first performing the equation of state calculations at a wide range of temperature and pressure conditions, and tracing the locus of stability of limit with respect to the gas and ice nucleation. The LLCP hypothesis can also be checked by determining the locus of thermodynamic extrema of isothermal compressibility and isobaric heat capacity, which will be asymptotically converged to an LLCP if it exists.

5.2.2 Future Work 2

Recent work shows that machine-learning methods can be used to develop improved models for water and other substances. Nonetheless, robust and easy to use software to develop these models is not yield broadly available. Thus, a fruitful avenue for future could be to develop an open-source Python package to help with the construction of such models. The accuracy of a classical molecular force field is essential to the applicability and predictive power of molecular simulations.¹⁷⁹ A successful development of a new molecular force field often requires the selection of appropriate functional form, and derivation of its associated parameters.^{180,181} Conventional construction of a molecular force field often relies on the physical intuition of a developer and careful choice of reference data obtained from the experiment.¹⁸¹ Thus, it is a non-trivial task to create an accurate force field for the substance of interest. To facilitate the development of reliable and accurate molecular force fields, software should be developed with the following in mind. First, the software should be able to interface with popular open-source MD package such as LAMMPS to deal with a diverse range of intermolecular potentials. Parallel on-the-fly MD simulations must be able to launch concurrently for efficiency. In order to develop a force-field that reproduces certain desired physical properties, objective functions should also be readily customizable in a consistent manner. To satisfy these design objectives, we have begun to develop a reproducible and systematic force-field parameterization workflow. The force-field parameterization starts with selecting a potential functional form that is compatiable with the MD packages of the user's choice. The initial guess of its associated parameters are provided. The desired physical properties are selected as reference properties, and corresponding objective function forms for minimizing the difference between the reference and predicted properties should also be defined. Then, the rest of tasks of force field development can be boiled down to solving a non-linear least squared problem, in which the force field parameters are iteratively optimized by minimizing difference between reference and predicted properties of interest. This minimization is achieved by implementing a gradient-free local optimization algorithm, which avoids expensive gradient evaluations. Finally, we have validated the current workflow by recovering the force field parameters and reproducing reference properties of mW model of water.

To demonstrate the validity of the parameterization workflow, we believe that developing a coarse-grained model of water that can exhibit the LLPT would be a good starting point not just for validating our workflow but also for understanding the role of molecular potentials on the LLPT phenomena. For the first stage, we will use ST2 water model, an atomistic model of water with a well-established LLPT, to generate extensive reference physical properties including intermolecular forces, radial distribution and isobars. Owing to the accuracy and efficiency of ML-BOP models of water, we choose bond-order potential form and ML-BOP parameters as the initial guess. It would be interesting to see if the ML-BOP model of water will be able to exhibit an LLPT after being trained against the reference data of ST2. The difference between the resulting and original ML-BOP force-field parameters can reveal crucial insights into the molecular origins of LLPTs in single-component systems.

5.2.3 Future Work 3

Previous equation of state, free energy and fluctuation analysis^{55,61,182} suggested that the family of mWAC models of silica exhibits LLCPs at supercooled conditions, with the location of the LLCP depending sentiently on the model parameters. To accurately locate the LLCPs in this family of models and understanding how they are affected by model parameters, future studies could apply the finite size scaling theory and histogram reweighting techniques to study the low-temperature criticality of mWAC model of silica.¹⁸³⁻¹⁸⁵ Such techniques require the fitting to 3D Ising universal distribution of magnetization, and in the case of single-component fluid, this order parameter is linear combination of density and configurational energy $\rho + sE$, in which s is a mixing-field parameter. The joint probability distribution of density and energy can be obtained by using the open-source Free Energy and Advanced Sampling Simulation Toolkit (FEASST) software to perform grand-canonical ensemble Monte Carlo (GCMC) with flat-histogram method to enhance the sampling at low temperature.¹⁸⁶ This package includes an expanded ensemble technique to perform the gradual insertion and deletion of an ion, as opposed to the conventional insertion and deletion of neutral ion pair. This method was shown to improve the equilibration of ionic liquids at relatively high density and low temperatures. ¹⁸⁶ The approach was validated and benchmarked by studying the restricted primitive model (RPM) of ions and was compared with the previous studies.

With this software, one can run a series of GCMC simulations near the estimated critical point to obtain the two-dimensional probability distribution of density and configurational energy. The chemical potential, temperature and mixing field parameters can then be adjusted so that the resulting joint distribution can be reweighted to match the 3D Ising distribution. By examining the quality of fits between the order parameter distribution and the asymptotic form for 3D Ising model, the universality of class of metastable critical point can be confirmed.

References

- [1] Franks, F.; of Chemistry (Great Britain), R. S. Water : a matrix of life, 2nd ed.;
 RSC paperbacks; Royal Society of Chemistry: Cambridge, UK, 2000; pp xii, 225 p.
- [2] Angell, C. A. Supercooled Water. Annual Review of Physical Chemistry 1983, 34, 593–630.
- [3] Debenedetti, P. G. Supercooled and glassy water. J. Phys.: Condens. Matter 2003, 15, R1669–R1726.
- [4] Stillinger, F. H. Advances in Chemical Physics; John Wiley & Sons, Ltd, 2007; pp 1–101.
- [5] Langham, E. J.; Mason, B. J. N.; Bernal, J. D. The heterogeneous and homogeneous nucleation of supercooled water. Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences 1958, 247, 493–504.
- [6] Mason, B. J. The Supercooling and Nucleation of Water. Adv. Phys. 1958, 7, 221–234.
- [7] Speedy, R. J.; Angell, C. A. Isothermal Compressibility of Supercooled Water and Evidence for a Thermodynamic Singularity at -45degreesc. J. Chem. Phys. 1976, 65, 851–858.
- [8] Angell, C. A.; Oguni, M.; Sichina, W. J. Heat-Capacity of Water at Extremes of Supercooling and Superheating. J. Phys. Chem. 1982, 86, 998–1002.

- [9] Angell, C. A.; Shuppert, J.; Tucker, J. C. Anomalous Properties of Supercooled Water - Heat-Capacity, Expansivity, and Proton Magnetic-Resonance Chemical-Shift from 0 to -38 Degrees. J. Phys. Chem. 1973, 77, 3092–3099.
- [10] Speedy, R. J. Stability-limit conjecture. An interpretation of the properties of water. J. Phys. Chem. 1982, 86, 982–991.
- [11] Amann-Winkel, K.; Bohmer, R.; Fujara, F.; Gainaru, C.; Geil, B.; Loerting, T. Colloquium: Water's controversial glass transitions. *Rev. Mod. Phys.* 2016, 88, 011002.
- [12] Poole, P. H.; Sciortino, F.; Essmann, U.; Stanley, H. E. Phase Behavior of Metastable Water. *Nature* 1992, 360, 324–328.
- [13] Angell, C. A. Insights into phases of liquid water from study of its unusual glass-forming properties. *Science* 2008, 319, 582–587.
- [14] Sastry, S.; Debenedetti, P. G.; Sciortino, F.; Stanley, H. E. Singularity-free interpretation of the thermodynamics of supercooled water. *Phys. Rev. E* 1996, 53, 6144–6154.
- [15] Stanley, H.; Cruz, L.; Harrington, S.; Poole, P.; Sastry, S.; Sciortino, F.; Starr, F.; Zhang, R. Cooperative molecular motions in water: The liquid-liquid critical point hypothesis. *Physica A: Statistical Mechanics and its Applications* 1997, 236, 19 – 37, Proceedings of the Workshop on Current Problems in Complex Fluids.
- [16] Gallo, P.; Amann-Winkel, K.; Angell, C. A.; Anisimov, M. A.; Caupin, F.; Chakravarty, C.; Lascaris, E.; Loerting, T.; Panagiotopoulos, A. Z.; Russo, J.; Sellberg, J. A.; Stanley, H. E.; Tanaka, H.; Vega, C.; Xu, L. M.; Pettersson, L. G. M. Water: A Tale of Two Liquids. *Chem. Rev.* **2016**, *116*, 7463–7500.

- [17] Gallo, P.; Stanley, H. E. Supercooled water reveals its secrets. Science 2017, 358, 1543–1544.
- [18] Giovambattista, N. Liquid Polymorphism; John Wiley & Sons, Ltd, 2013; pp 113–138.
- [19] Handle, P. H.; Loerting, T.; Sciortino, F. Supercooled and glassy water: Metastable liquid(s), amorphous solid(s), and a no-man's land. *Proc. Natl. Acad. Sci. U.S.A.* 2017, 114, 13336–13344.
- [20] Xu, L.; Kumar, P.; Buldyrev, S. V.; Chen, S.-H.; Poole, P. H.; Sciortino, F.; Stanley, H. E. Relation between the Widom line and the dynamic crossover in systems with a liquid-liquid phase transition. *Proc. Natl. Acad. Sci. U.S.A.* 2005, 102, 16558–16562.
- [21] Mishima, O.; Stanley, H. E. The relationship between liquid, supercooled and glassy water. *Nature* 1998, 396, 329–335.
- [22] Kim, K. H.; Späh, A.; Pathak, H.; Perakiotherss, F.; Mariedahl, D.; Amann-Winkel, K.; Sellberg, J. A.; Lee, J. H.; Kim, S.; Park, J.; Hyun Nam, K.; Katayama, T.; Nillson, A. Maxima in the thermodynamic response and correlation functions of deeply supercooled water. *Science* **2017**, *358*, 1589–1593.
- [23] Sellberg, J. A.; Huang, C.; McQueen, T. A.; Loh, N. D.; Laksmono, H.; Schlesinger, D.; Sierra, R. G.; Nordlund, D.; Hampton, C. Y.; Starodub, D.; DePonte, D. P.; Beye, M.; Chen, C.; Martin, A. V.; Barty, A.; Wikfeldt, K. T.; Weiss, T. M.; Caronna, C.; Feldkamp, J.; Skinner, L. B.; Seibert, M. M.; Messerschmidt, M.; Williams, G. J.; Boutet, S.; Pettersson, L. G. M.; Bogan, M. J.; Nilsson, A. Ultrafast X-ray probing of water structure below the homogeneous ice nucleation temperature. *Nature* 2014, *510*, 381–384.

- [24] Abascal, J. L. F.; Vega, C. Widom line and the liquid-liquid critical point for the TIP4P/2005 water model. J. Chem. Phys. 2010, 133, 234502.
- [25] Limmer, D. T.; Chandler, D. The putative liquid-liquid transition is a liquidsolid transition in atomistic models of water. J. Chem. Phys. 2011, 135, 134503.
- [26] Moore, E. B.; Molinero, V. Structural transformation in supercooled water controls the crystallization rate of ice. *Nature* 2011, 479, 506–508.
- [27] Ni, Y. C.; Skinner, J. L. Evidence for a liquid-liquid critical point in supercooled water within the E3B3 model and a possible interpretation of the kink in the homogeneous nucleation line. J. Chem. Phys. 2016, 144, 214501.
- [28] Palmer, J. C.; Martelli, F.; Liu, Y.; Car, R.; Panagiotopoulos, A. Z.; Debenedetti, P. G. Metastable liquid-liquid transition in a molecular model of water. *Nature* 2014, 510, 385–388.
- [29] Paschek, D. How the liquid-liquid transition affects hydrophobic hydration in deeply supercooled water. *Phys. Rev. Lett.* 2005, 94, 217802.
- [30] Poole, P. H.; Bowles, R. K.; Saika-Voivod, I.; Sciortino, F. Free energy surface of ST2 water near the liquid-liquid phase transition. J. Chem. Phys. 2013, 138, 034505.
- [31] Yamada, M.; Mossa, S.; Stanley, H. E.; Sciortino, F. Interplay between Time-Temperature Transformation and the Liquid-Liquid Phase Transition in Water. *Phys. Rev. Lett.* 2002, 88, 195701.
- [32] Liu, Y.; Palmer, J. C.; Panagiotopoulos, A. Z.; Debenedetti, P. G. Liquid-liquid transition in ST2 water. J. Chem. Phys. 2012, 137, 214505.
- [33] Liu, Y.; Panagiotopoulos, A. Z.; Debenedetti, P. G. Low-temperature fluid-phase behavior of ST2 water. J. Chem. Phys. 2009, 131, 104508.

- [34] Palmer, J. C.; Car, R.; Debenedetti, P. G. The liquid-liquid transition in supercooled ST2 water: a comparison between umbrella sampling and welltempered metadynamics. *Faraday Discuss.* 2013, 167, 77–94.
- [35] Sciortino, F.; Saika-Voivod, I.; Poole, P. H. Study of the ST2 model of water close to the liquid-liquid critical point. *Phys. Chem. Chem. Phys.* 2011, 13, 19759– 19764.
- [36] Stillinger, F. H.; Rahman, A. Improved Simulation of Liquid Water by Molecular Dynamics. J. Chem. Phys. 1974, 60, 1545–1557.
- [37] Harrington, S.; Poole, P. H.; Sciortino, F.; Stanley, H. E. Equation of state of supercooled water simulated using the extended simple point charge intermolecular potential. *The Journal of Chemical Physics* **1997**, *107*, 7443–7450.
- [38] Cuthbertson, M. J.; Poole, P. H. Mixturelike Behavior Near a Liquid-Liquid Phase Transition in Simulations of Supercooled Water. *Phys. Rev. Lett.* 2011, 106, 115706.
- [39] Kesselring, T. A.; Lascaris, E.; Franzese, G.; Buldyrev, S. V.; Herrmann, H. J.; Stanley, H. E. Finite-size scaling investigation of the liquid-liquid critical point in ST2 water and its stability with respect to crystallization. J. Chem. Phys. 2013, 138, 244506.
- [40] Poole, P. H.; Saika-Voivod, I.; Sciortino, F. Density minimum and liquid-liquid phase transition. J. Phys.: Condens. Matter 2005, 17, L431–L437.
- [41] Smallenburg, F.; Filion, L.; Sciortino, F. Erasing no-man's land by thermodynamically stabilizing the liquid-liquid transition in tetrahedral particles. *Nat. Phys.* 2014, 10, 653–657.

- [42] Smallenburg, F.; Sciortino, F. Tuning the Liquid-Liquid Transition by Modulating the Hydrogen-Bond Angular Flexibility in a Model for Water. *Phys. Rev. Lett.* 2015, 115, 015701.
- [43] Aasland, S.; McMillan, P. F. Density-driven liquid-liquid phase separation in the system AI2O3-Y2O3. Nature 1994, 369, 633-636.
- [44] Katayama, Y.; Mizutani, T.; Utsumi, W.; Shimomura, O. A first-order liquidliquid phase transition in phosphorus. *Nature* 2000, 403, 1998–2001.
- [45] Sastry, S.; Austen Angell, C. Liquid-liquid phase transition in supercooled silicon. Nat. Mater. 2003, 2, 739–743.
- [46] Tanaka, H.; Kurita, R.; Mataki, H. Liquid-Liquid Transition in the Molecular Liquid Triphenyl Phosphite. *Phys. Rev. Lett.* 2004, 92, 025701.
- [47] Angell, C. A.; Kanno, H. Density Maxima in High-Pressure Supercooled Water and Liquid Silicon Dioxide. *Science* 1976, 193, 1121–1122.
- [48] Henry, L.; Mezouar, M.; Garbarino, G.; Sifré, D.; Weck, G.; Datchi, F. Liquid-liquid transition and critical point in sulfur. *Nature* 2020, 584, 382–386.
- [49] Flörke, O. W.; Graetsch, H.; Brunk, F.; Benda, L.; Paschen, S.; Bergna, H. E.; Roberts, W. O.; Welsh, W. A.; Chapman, D. M.; Ettlinger, M.; Kerner, D.; Maier, M.; Meon, W.; Schmoll, R.; Gies, H.; Schiffmann, D. Ullmann's Encyclopedia of Industrial Chemistry; American Cancer Society, 2000; Chapter 1, p 1.
- [50] Angell, C. A. Forty Years of Silica Simulations. Which Way Now? International Journal of Applied Glass Science 2015, 6, 3–14.
- [51] Angell, C. A.; Cheeseman, P. A.; Tamaddon, S. Pressure Enhancement of Ion

Mobilities in Liquid Silicates from Computer Simulation Studies to 800 Kilobars. *Science* **1982**, *218*, 885–887.

- [52] Vatalis, K. I.; Charalambides, G.; Benetis, N. P. Market of High Purity Quartz Innovative Applications. *Procedia Economics and Finance* 2015, 24, 734 – 742, International Conference on Applied Economics (ICOAE) 2015, 2-4 July 2015, Kazan, Russia.
- [53] Wenk, H.-R.; Bulakh, A. Minerals: Their Constitution and Origin; Cambridge University Press, 2004.
- [54] Shi, R.; Tanaka, H. Impact of local symmetry breaking on the physical properties of tetrahedral liquids. *Proceedings of the National Academy of Sciences* 2018, 115, 1980–1985.
- [55] Chen, R.; Lascaris, E.; Palmer, J. C. Liquid-liquid phase transition in an ionic model of silica. *The Journal of Chemical Physics* 2017, 146, 234503.
- [56] Saika-Voivod, I.; Poole, P. H. Liquid Polymorphism; John Wiley & Sons, Ltd, 2013; pp 373–384.
- [57] Woodcock, L. V.; Angell, C. A.; Cheeseman, P. Molecular-Dynamics Studies of Vitreous State - Simple Ionic Systems and Silica. J. Chem. Phys. 1976, 65, 1565–1577.
- [58] van Beest, B. W. H.; Kramer, G. J.; van Santen, R. A. Force-Fields for Silicas and Aluminophosphates Based on Abinitio Calculations. *Phys. Rev. Lett.* 1990, 64, 1955–1958.
- [59] Saika-Voivod, I.; Sciortino, F.; Poole, P. H. Computer simulations of liquid silica: Equation of state and liquid-liquid phase transition. *Phys. Rev. E* 2000, 63, 011202.

- [60] Lascaris, E.; Hemmati, M.; Buldyrev, S. V.; Stanley, H. E.; Angell, C. A. Search for a liquid-liquid critical point in models of silica. J. Chem. Phys. 2014, 140, 224502.
- [61] Lascaris, E. Tunable Liquid-Liquid Critical Point in an Ionic Model of Silica. Phys. Rev. Lett. 2016, 116, 125701.
- [62] Sosso, G. C.; Chen, J.; Cox, S. J.; Fitzner, M.; Pedevilla, P.; Zen, A.; Michaelides, A. Crystal Nucleation in Liquids: Open Questions and Future Challenges in Molecular Dynamics Simulations. *Chem. Rev.* 2016, 116, 7078– 7116.
- [63] Vekilov, P. G. Nucleation. Crystal Growth & Design 2010, 10, 5007-5019.
- [64] Russell, L. M. Sea-spray particles cause freezing in clouds. Nature 2015, 525, 194–195.
- [65] Herbert, R. J.; Murray, B. J.; Dobbie, S. J.; Koop, T. Sensitivity of liquid clouds to homogenous freezing parameterizations. *Geophysical Research Letters* 2015, 42, 1599–1605.
- [66] Gibbs, J. W. The collected works of J. Willard Gibbs.; Longmans, Green: New York, 1928.
- [67] Karthika, S.; Radhakrishnan, T. K.; Kalaichelvi, P. A Review of Classical and Nonclassical Nucleation Theories. *Crystal Growth and Design* 2016,
- [68] Russo, J.; Romano, F.; Tanaka, H. New metastable form of ice and its role in the homogeneous crystallization of water. *Nature Materials* 2014, 13, 733–739.
- [69] Matsumoto, M.; Saito, S.; Ohmine, I. Molecular dynamics simulation of the ice nucleation and growth process leading to water freezing. *Nature* 2002, 416, 409–413.

- [70] Sanz, E.; Vega, C.; Espinosa, J. R.; Caballero-Bernal, R.; Abascal, J. L. F.; Valeriani, C. Homogeneous Ice Nucleation at Moderate Supercooling from Molecular Simulation. *Journal of the American Chemical Society* **2013**, *135*, 15008– 15017.
- [71] Molinero, V.; Moore, E. B. Water Modeled As an Intermediate Element between Carbon and Silicon. J. Phys. Chem. B 2009, 113, 4008–4016.
- [72] Moore, E. B.; Molinero, V. Is it cubic? Ice crystallization from deeply supercooled water. Phys. Chem. Chem. Phys. 2011, 13, 20008–20016.
- [73] Moore, E. B.; de la Llave, E.; Welke, K.; Scherlis, D. A.; Molinero, V. Freezing, melting and structure of ice in a hydrophilic nanopore. *Phys. Chem. Chem. Phys.* 2010, 12, 4124–4134.
- [74] Haji-Akbari, A.; Debenedetti, P. G. Direct calculation of ice homogeneous nucleation rate for a molecular model of water. *Proc. Natl. Acad. Sci. U.S.A.* 2015, 112, 10582–10588.
- [75] Reinhardt, A.; Doye, J. P. K. Free energy landscapes for homogeneous nucleation of ice for a monatomic water model. *The Journal of Chemical Physics* 2012, 136, 054501.
- [76] Li, T.; Donadio, D.; Russo, G.; Galli, G. Homogeneous ice nucleation from supercooled water. *Phys. Chem. Chem. Phys.* 2011, 13, 19807–19813.
- [77] Reinhardt, A.; Doye, J. P. K.; Noya, E. G.; Vega, C. Local order parameters for use in driving homogeneous ice nucleation with all-atom models of water. *The Journal of Chemical Physics* 2012, 137, 194504.
- [78] Kanno, H.; Angell, C. A. Water Anomalous Compressibilities to 1.9-Kbar and Correlation with Supercooling Limits. J. Chem. Phys. 1979, 70, 4008–4016.

- [79] Oguni, M.; Angell, C. A. Anomalous Components of Supercooled Water Expansivity, Compressibility, and Heat-Capacity (Cp and Cv) from Binary Formamide+Water Solution Studies. J. Chem. Phys. 1983, 78, 7334–7342.
- [80] Holten, V.; Bertrand, C.; Anisimov, M.; Sengers, J. Thermodynamics of supercooled water. J. Chem. Phys. 2012, 136, 094507.
- [81] Holten, V.; Sengers, J. V.; Anisimov, M. A. Equation of state for supercooled water at pressures up to 400 MPa. J. Phys. Chem. Ref. Data 2014, 43, 043101.
- [82] Mishima, O.; Calvert, L. D.; Whalley, E. An apparently first-order transition between two amorphous phases of ice induced by pressure. *Nature* 1985, 314, 76–78.
- [83] Amann-Winkel, K.; Gainaru, C.; Handle, P. H.; Seidl, M.; Nelson, H.;
 Bohmer, R.; Loerting, T. Water's second glass transition. *Proc. Natl. Acad. Sci.* U.S.A. 2013, 110, 17720–17725.
- [84] Perakis, F.; Amann-Winkel, K.; Lehmkühler, F.; Sprung, M.; Mariedahl, D.; Sellberg, J. A.; Pathak, H.; Späh, A.; Cavalca, F.; Schlesinger, D.; Ricci, A.; Jain, A.; Massani, B.; Aubree, F.; Benmore, C. J.; Loerting, T.; Grübel, G.; Pettersson, L. G. M.; Nilsson, A. Diffusive dynamics during the high-to-low density transition in amorphous ice. *Proc. Natl. Acad. Sci. U.S.A.* 2017, 114, 8193–8198.
- [85] Li, Y. P.; Li, J. C.; Wang, F. Liquid-liquid transition in supercooled water suggested by microsecond simulations. *Proc. Natl. Acad. Sci. U.S.A.* 2013, 110, 12209–12212.
- [86] Holten, V.; Palmer, J. C.; Poole, P. H.; Debenedetti, P. G.; Anisimov, M. A. Twostate thermodynamics of the ST2 model for supercooled water. J. Chem. Phys. 2014, 140, 104502.

- [87] Singh, R. S.; Biddle, J. W.; Debenedetti, P. G.; Anisimov, M. A. Two-state thermodynamics and the possibility of a liquid-liquid phase transition in supercooled TIP4P/2005 water. J. Chem. Phys. 2016, 144, 144504.
- [88] Palmer, J. C.; Singh, R. S.; Chen, R.; Martelli, F.; Debenedetti, P. G. Density and bond-orientational relaxations in supercooled water. *Mol. Phys.* 2016, 114, 2580–2585.
- [89] Pathak, H.; Palmer, J. C.; Schlesinger, D.; Wikfeldt, K. T.; Sellberg, J. A.; Pettersson, L. G. M.; Nilsson, A. The structural validity of various thermodynamical models of supercooled water. J. Chem. Phys. 2016, 145, 134507.
- [90] Huang, C.; Wikfeldt, K. T.; Tokushima, T.; Nordlund, D.; Harada, Y.; Bergmann, U.; Niebuhr, M.; Weiss, T. M.; Horikawa, Y.; Leetmaa, M.; Ljungberg, O., M. P. yand Takahashi; Lenz, A.; Ojamäe, L.; Lyubartsev, A. P.; Shin, S.; Pettersson, L. G. M.; Nilsson, A. The inhomogeneous structure of water at ambient conditions. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 15214– 15218.
- [91] Bhatia, A.; Thornton, D. Structural aspects of the electrical resistivity of binary alloys. *Phys. Rev. B* 1970, 2, 3004–3012.
- [92] Overduin, S. D.; Patey, G. N. Understanding the structure factor and isothermal compressibility of ambient water in terms of local structural environments. J. Phys. Chem. B 2012, 116, 12014–12020.
- [93] Overduin, S. D.; Patey, G. N. Fluctuations and local ice structure in model supercooled water. J. Chem. Phys. 2015, 143, 094504.
- [94] Das, S. K.; Horbach, J.; Binder, K.; Fisher, M. E.; Sengers, J. V. Static and dynamic critical behavior of a symmetrical binary fluid: A computer simulation. *J. Chem. Phys.* 2006, 125, 024506.

- [95] Nosé, S. A unified formulation of the constant temperature molecular dynamics methods. J. Chem. Phys. 1984, 81, 511–519.
- [96] Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* 1985, 31, 1695–1697.
- [97] Errington, J. R.; Debenedetti, P. G. Relationship between structural order and the anomalies of liquid water. *Nature* 2001, 409, 318–321.
- [98] Ashcroft, N.; Langreth, D. C. Structure of binary liquid mixtures. I. *Phys. Rev.* 1967, 156, 685.
- [99] Salacuse, J.; Denton, A.; Egelstaff, P. Finite-size effects in molecular dynamics simulations: Static structure factor and compressibility. I. Theoretical method. *Phys. Rev. E* 1996, *53*, 2382.
- [100] Sedlmeier, F.; Horinek, D.; Netz, R. R. Spatial correlations of density and structural fluctuations in liquid water: A comparative simulation study. J. Am. Chem. Soc. 2011, 133, 1391–1398.
- [101] Pettersson, L. G.; Nilsson, A. The structure of water; from ambient to deeply supercooled. J. Non-Cryst. Solids 2015, 407, 399–417.
- [102] Soper, A.; Bruni, F.; Ricci, M. Site-site pair correlation functions of water from 25 to 400 C: Revised analysis of new and old diffraction data. J. Chem. Phys. 1997, 106, 247–254.
- [103] Moore, E. B.; Molinero, V. Growing correlation length in supercooled water. J. Chem. Phys. 2009, 130, 244505.
- [104] Clark, G. N. I.; Hura, G. L.; Teixeira, J.; Soper, A. K.; Head-Gordon, T. Smallangle scattering and the structure of ambient liquid water. *Proc. Natl. Acad. Sci. U.S.A.* 2010, 107, 14003–14007.

- [105] Conde, O.; Teixeira, J.; Papon, P. Analysis of sound velocity in supercoled H2O,
 D2O, and water-ethanol mixtures. J. Chem. Phys. 1982, 76, 3747–3753.
- [106] Fisher, M. E. The renormalization group in the theory of critical behavior. *Rev. Mod. Phys.* 1974, 46, 597.
- [107] Deutsch, H.; Binder, K. The mean field to Ising crossover in the critical behavior of polymer mixtures: a finite size scaling analysis of Monte Carlo simulations. J. Phys. II 1993, 3, 1049–1073.
- [108] Jacob, J.; Kumar, A.; Anisimov, M.; Povodyrev, A.; Sengers, J. Crossover from Ising to mean-field critical behavior in an aqueous electrolyte solution. *Phys. Rev. E* 1998, 58, 2188.
- [109] Chen, R.; Lascaris, E.; Palmer, J. C. Liquid-liquid phase transition in an ionic model of silica. J. Chem. Phys. 2017, 146, 234503.
- [110] Yagasaki, T.; Matsumoto, M.; Tanaka, H. Spontaneous liquid-liquid phase separation of water. *Phys. Rev. E* 2014, *89*, 020301.
- [111] Nguyen, A. H.; Molinero, V. Identification of clathrate hydrates, hexagonal ice, cubic ice, and liquid water in simulations: The CHILL+ algorithm. J. Phys. Chem. B 2015, 119, 9369–9376.
- [112] Palmer, J. C.; Poole, P. H.; Sciortino, F.; Debenedetti, P. G. Advances in Computational Studies of the Liquid-Liquid Transition in Water and Water-Like Models. *Chem. Rev.* 2018, DOI: 10.1021/acs.chemrev.8b00228.
- [113] Anisimov, M. A.; Duška, M.; Caupin, F.; Amrhein, L. E.; Rosenbaum, A.; Sadus, R. J. Thermodynamics of Fluid Polyamorphism. *Phys. Rev. X* 2018, 011004.
- [114] Mukherjee, G.; Vaidya, S.; Sugandhi, V. Direct observation of amorphous to amorphous apparently first-order phase transition in fused quartz. *Phys. Rev. Lett.* 2001, 87, 195501.
- [115] Holten, V.; Qiu, C.; Guillerm, E.; Wilke, M.; Ricka, J.; Frenz, M.; Caupin, F. Compressibility anomalies in stretched water and their interplay with density anomalies. J. Phys. Chem. Lett. 2017, 8, 5519–5522.
- [116] Starr, F. W.; Sciortino, F. "Crystal-clear" liquid-liquid transition in a tetrahedral fluid. Soft Matter 2014, 10, 9413–9422.
- [117] Limmer, D. T.; Chandler, D. The putative liquid-liquid transition is a liquidsolid transition in atomistic models of water. II. J. Chem. Phys. 2013, 138, 214504.
- [118] Palmer, J. C.; Haji-Akbari, A.; Singh, R. S.; Martelli, F.; Car, R.; Panagiotopoulos, A. Z.; Debenedetti, P. G. Comment on "The putative liquid-liquid transition is a liquid-solid transition in atomistic models of water" [I and II: J. Chem. Phys. 135, 134503 (2011); J. Chem. Phys. 138, 214504 (2013)]. J. Chem. Phys. 2018, 148, 137101.
- [119] Guo, J.; Singh, R. S.; Palmer, J. C. Anomalous scattering in supercooled ST2 water. Mol. Phys. 2018, 116, 1953–1964.
- [120] Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **2013**, 29, 845–854.
- [121] Shell, M. S.; Debenedetti, P. G.; Panagiotopoulos, A. Z. Molecular structural order and anomalies in liquid silica. *Phys. Rev. E* 2002, 66, 011202.

- [122] Munaó, G.; Costa, D.; Sciortino, F.; Caccamo, C. Simulation and theory of a model for tetrahedral colloidal particles. J. Chem. Phys. 2011, 134, 194502.
- [123] Hsu, C. W.; Largo, J.; Sciortino, F.; Starr, F. W. Hierarchies of networked phases induced by multiple liquid–liquid critical points. *Proc. Natl. Acad. Sci. U.S.A.* 2008, 105, 13711–13715.
- [124] Schmitz, F.; Virnau, P.; Binder, K. Logarithmic finite-size effects on interfacial free energies: Phenomenological theory and Monte Carlo studies. *Phys. Rev. E* 2014, 90, 012128.
- [125] Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. J. Comput. Phys. 1995, 117, 1–19.
- [126] De Pablo, J. J.; Escobedo, F. A. Molecular simulations in chemical engineering: Present and future. *AIChE J.* 2002, 48, 2716–2721.
- [127] Theodorou, D. N. Progress and outlook in Monte Carlo simulations. Ind. Eng. Chem. Res. 2010, 49, 3047–3058.
- [128] Dubbeldam, D.; Torres-Knoop, A.; Walton, K. S. On the inner workings of Monte Carlo codes. *Mol. Simul.* 2013, 39, 1253–1292.
- [129] Palmer, J. C.; Debenedetti, P. G. Recent advances in molecular simulation: A chemical engineering perspective. AIChE J. 2015, 61, 370–383.
- [130] Peters, B. Reaction Rate Theory and Rare Events Simulations; Elsevier: Amsterdam, 2017.
- [131] Dubbeldam, D.; Calero, S.; Ellis, D. E.; Snurr, R. Q. RASPA: molecular simulation software for adsorption and diffusion in flexible nanoporous materials. *Mol. Simul.* 2016, 42, 81–101.

- [132] Mick, J. R.; Soroush Barhaghi, M.; Potoff, J. J. Prediction of radon-222 phase behavior by Monte Carlo simulation. J. Chem. Eng. Data 2016, 61, 1625–1631.
- [133] Shah, J. K.; Marin-Rimoldi, E.; Mullen, R. G.; Keene, B. P.; Khan, S.; Paluch, A. S.; Rai, N.; Romanielo, L. L.; Rosch, T. W.; Yoo, B., et al. Cassandra: An open source Monte Carlo package for molecular simulation. *J. Comput. Chem.* 2017, 38, 1727—1739.
- [134] Allen, M. P.; Quigley, D. Some comments on Monte Carlo and molecular dynamics methods. *Mol. Phys.* 2013, 111, 3442–3447.
- [135] Duane, S.; Kennedy, A. D.; Pendleton, B. J.; Roweth, D. Hybrid Monte Carlo. *Phys. Lett. B* 1987, 195, 216–222.
- [136] Mehlig, B.; Heermann, D.; Forrest, B. Hybrid Monte Carlo method for condensed-matter systems. *Phys. Rev. B* 1992, 45, 679–685.
- [137] Forrest, B.; Suter, U. Hybrid Monte Carlo simulations of dense polymer systems. J. Chem. Phys. 1994, 101, 2616–2629.
- [138] Berne, B. J.; Straub, J. E. Novel methods of sampling phase space in the simulation of biological systems. *Curr. Opin. Struct. Biol.* **1997**, 7, 181–189.
- [139] Reinhardt, A.; Doye, J. P.; Noya, E. G.; Vega, C. Local order parameters for use in driving homogeneous ice nucleation with all-atom models of water. J. Chem. Phys. 2012, 137, 194504.
- [140] Reinhardt, A.; Doye, J. P. Note: Homogeneous TIP4P/2005 ice nucleation at low supercooling. J. Chem. Phys. 2013, 139, 096102.
- [141] Reinhardt, A.; Doye, J. P. Effects of surface interactions on heterogeneous ice nucleation for a monatomic water model. J. Chem. Phys. 2014, 141, 084501.

- [142] Geiger, P.; Dellago, C. Neural networks for local structure detection in polymorphic systems. J. Chem. Phys. 2013, 139, 164105.
- [143] Joswiak, M. N.; Duff, N.; Doherty, M. F.; Peters, B. Size-dependent surface free energy and Tolman-corrected droplet nucleation of TIP4P/2005 water. J. Phys. Chem. Lett. 2013, 4, 4267–4272.
- [144] Gonzalez, M.; Sanz, E.; McBride, C.; Abascal, J.; Vega, C.; Valeriani, C. Nucleation free-energy barriers with Hybrid Monte Carlo/Umbrella Sampling. *Phys. Chem. Chem. Phys.* 2014, 16, 24913–24919.
- [145] Menzl, G.; Dellago, C. Effect of entropy on the nucleation of cavitation bubbles in water under tension. J. Chem. Phys. 2016, 145, 211918.
- [146] Menzl, G.; Singraber, A.; Dellago, C. S-shooting: a Bennett-Chandler-like method for the computation of rate constants from committor trajectories. *Faraday Discuss.* 2017, 195, 345-364.
- [147] Zhong, E. D.; Shirts, M. R. Thermodynamics of coupled protein adsorption and stability using hybrid Monte Carlo simulations. *Langmuir* 2014, 30, 4952– 4961.
- [148] http://lammps.sandia.gov/, LAMMPS: Large-scale Atomic/Molecular Massively Parallel Simulator, http://lammps.sandia.gov (2017).
- [149] Desgranges, C.; Delhommelle, J. Phase equilibria of molecular fluids via hybrid Monte Carlo Wang-Landau simulations: Applications to benzene and nalkanes. J. Chem. Phys. 2009, 130, 244109.
- [150] Rane, K. S.; Murali, S.; Errington, J. R. Monte Carlo simulation methods for computing liquid-vapor saturation properties of model systems. J. Chem. Theory Comput. 2013, 9, 2552–2566.

- [151] Frenkel, D.; Smit, B. Understanding Molecular Simulation (Second Edition);Academic Press: San Diego, 2002.
- [152] Rowley, L.; Nicholson, D.; Parsonage, N. Monte Carlo grand canonical ensemble calculation in a gas-liquid transition region for 12-6 argon. *J. Comput. Phys.* 1975, *17*, 401–414.
- [153] Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant pressure molecular dynamics algorithms. J. Chem. Phys. 1994, 101, 4177–4189.
- [154] Chodera, J. D. A simple method for automated equilibration detection in molecular simulations. J. Chem. Theory Comput. 2016, 12, 1799–1805.
- [155] https://www.nist.gov/mml/csd/chemical-informatics-research-group/ lennard-jones-fluid-reference-calculations, D. Siderius, National Institute of Science and Technology, Lennard-Jones Fluid Reference Calculations, https://www.nist.gov/mml/csd/chemical-informatics-researchgroup/lennard-jones-fluid-reference-calculations (2017).
- [156] Shirts, M. R. Simple quantitative tests to validate sampling from thermodynamic ensembles. J. Chem. Theory Comput. 2013, 9, 909–926.
- [157] https://github.com/shirtsgroup/checkensemble/, checkensemble, https://github.com/shirtsgroup/checkensemble (2017).
- [158] https://github.com/pearu/f2py, F2PY, https://github.com/pearu/f2py (2017).
- [159] Molinero, V.; Moore, E. B. Water modeled as an intermediate element between carbon and silicon. J. Phys. Chem. B 2008, 113, 4008–4016.
- [160] Stillinger, F. H.; Weber, T. A. Computer simulation of local order in condensed phases of silicon. *Phys. Rev. B* 1985, 31, 5262–5271.

- [161] Espinosa, J.; Sanz, E.; Valeriani, C.; Vega, C. Homogeneous ice nucleation evaluated for several water models. J. Chem. Phys. 2014, 141, 18C529.
- [162] Sosso, G. C.; Li, T.; Donadio, D.; Tribello, G. A.; Michaelides, A. Microscopic Mechanism and Kinetics of Ice Formation at Complex Interfaces: Zooming in on Kaolinite. J. Phys. Chem. Lett. 2016, 7, 2350–2355.
- [163] Haji-Akbari, A.; Debenedetti, P. G. Computational investigation of surface freezing in a molecular model of water. *Proc. Natl. Acad. Sci. U.S.A.* 2017, 114, 3316–3321.
- [164] Haji-Akbari, A.; DeFever, R. S.; Sarupria, S.; Debenedetti, P. G. Suppression of sub-surface freezing in free-standing thin films of a coarse-grained model of water. *Phys. Chem. Chem. Phys.* **2014**, *16*, 25916–25927.
- [165] Lupi, L.; Peters, B.; Molinero, V. Pre-ordering of interfacial water in the pathway of heterogeneous ice nucleation does not lead to a two-step crystallization mechanism. J. Chem. Phys. 2016, 145, 211910.
- [166] Torrie, G.; Valleau, J. Nonphysical sampling distributions in Monte Carlo freeenergy estimation: Umbrella sampling. J. Comput. Phys. 1977, 23, 187 – 199.
- [167] Wang, F.; Landau, D. Efficient, multiple-range random walk algorithm to calculate the density of states. *Phys. Rev. Lett.* 2001, 86, 2050–2053.
- [168] Shen, V. K.; Siderius, D. W. Elucidating the effects of adsorbent flexibility on fluid adsorption using simple models and flat-histogram sampling methods. J. Chem. Phys. 2014, 140, 244106.
- [169] Shirts, M. R.; Chodera, J. D. Statistically optimal analysis of samples from multiple equilibrium states. J. Chem. Phys. 2008, 129, 124105.

- [170] Reinhardt, A. Computer simulation of the homogeneous nucleation of ice.Ph.D. thesis, University of Oxford, 2013.
- [171] https://github.com/palmergroup/hmcwithlammps, Code and examples available at https://github.com/palmergroup/hmcwithlammps (2017).
- [172] Palmer, J. C.; Haji-Akbari, A.; Singh, R. S.; Martelli, F.; Car, R.; Panagiotopoulos, A. Z.; Debenedetti, P. G. Notes on the Hybrid Monte Carlo Method. arXiv preprint arXiv:1712.08278 2017,
- [173] Miller III, T.; Eleftheriou, M.; Pattnaik, P.; Ndirango, A.; Newns, D.; Martyna, G. Symplectic quaternion scheme for biophysical molecular dynamics. J. Chem. Phys. 2002, 116, 8649.
- [174] Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. 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.
- [175] Andersen, H. C. RATTTLE: A "velocity" version of the SHAKE algorithm for molecular dynamics calculations. J. Comput. Phys. 1983, 52, 24–34.
- [176] Debenedetti, P. G.; Sciortino, F.; Zerze, G. H. Second critical point in two realistic models of water. *Science* 2020, 369, 289–292.
- [177] Chan, H.; Cherukara, M. J.; Narayanan, B.; Loeffler, T. D.; Benmore, C.; Gray, S. K.; Sankaranarayanan, S. K. R. S. Machine learning coarse grained models for water. *Nature Communications* **2019**, *10*, 379.
- [178] Tersoff, J. New empirical approach for the structure and energy of covalent systems. *Phys. Rev. B* 1988, 37, 6991–7000.
- [179] Michael P. Allen, D. J. T. Computer Simulation of Liquids, 2nd ed.; RSC paperbacks; Oxford University Press: Oxford, UK, 2017; p 20.

- [180] Wang, L.-P.; Chen, J.; Van Voorhis, T. Systematic Parametrization of Polarizable Force Fields from Quantum Chemistry Data. *Journal of Chemical Theory* and Computation 2013, 9, 452–460.
- [181] Wang, L.-P.; Martinez, T. J.; Pande, V. S. Building Force Fields: An Automatic, Systematic, and Reproducible Approach. *The Journal of Physical Chemistry Letters* 2014, 5, 1885–1891.
- [182] Guo, J.; Palmer, J. C. Fluctuations near the liquid-liquid transition in a model of silica. *Phys. Chem. Chem. Phys.* **2018**, 20, 25195–25202.
- [183] Wilding, N. B.; Bruce, A. D. Density fluctuations and field mixing in the critical fluid. *Journal of Physics: Condensed Matter* **1992**, *4*, 3087–3108.
- [184] Wilding, N. B. Critical-point and coexistence-curve properties of the Lennard-Jones fluid: A finite-size scaling study. *Phys. Rev. E* 1995, 52, 602–611.
- [185] Ferrenberg, A. M.; Swendsen, R. H. New Monte Carlo technique for studying phase transitions. *Phys. Rev. Lett.* 1988, 61, 2635–2638.
- [186] Hatch, H. W.; Hall, S. W.; Errington, J. R.; Shen, V. K. Improving the efficiency of Monte Carlo simulations of ions using expanded grand canonical ensembles. *The Journal of Chemical Physics* **2019**, *151*, 144109.