Single Water Entropy : Implementation and Analysis on Supercooled Water

A Thesis

submitted to
Indian Institute of Science Education and Research Pune
in partial fulfillment of the requirements for the
BS-MS Dual Degree Programme

by

Sarathchandran J



Indian Institute of Science Education and Research Pune
Dr. Homi Bhabha Road,
Pashan, Pune 411008, INDIA.

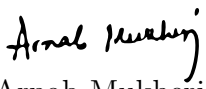
May, 2022

Supervisor: Arnab Mukherjee
© Sarathchandran J 2022

All rights reserved

Certificate

This is to certify that this dissertation entitled Single Water Entropy : Implementation and Analysis on Supercooled Watertowards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents study/work carried out by Sarathchandran Jat Indian Institute of Science Education and Research under the supervision of Arnab Mukherjee, Professor, Department of Chemistry , during the academic year 2021-2022.



Arnab Mukherjee

Committee:

Arnab Mukherjee

Rakesh S. Singh

Dedicated to all my teachers ...

Declaration

I hereby declare that the matter embodied in the report entitled Single Water Entropy : Implementation and Analysis on Supercooled Water are the results of the work carried out by me at the Department of Chemistry, Indian Institute of Science Education and Research, Pune, under the supervision of Arnab Mukherjee and the same has not been submitted elsewhere for any other degree.



Sarathchandran J

Acknowledgments

I would like to express my sincere gratitude to my thesis supervisor Prof. Arnab Mukherjee, for the encouragement and guidance he provided throughout my MS project. I am grateful for the freedom he entrusted me with to explore the topics of my interest. I would also like to thank my expert Dr.Rakesh S. Singh for the numerous discussions that helped me in my project. IISER Pune had provided me with an amazing opportunity, that helped me find my interests. I thank the institute for the resources it provided throughout this project. I am grateful to everyone at my lab who made the lab experience very enjoyable. Last but not least I would like to thank my parents who taught me the most important lessons in life, for I wouldn't be where I am today if not for them.

Abstract

The calculation of solvent entropy from a molecular dynamics simulation remains to be a challenge due to the difficulty of sampling the large phase space of even a small system of 1000 water molecules in realistic time scales. This motivates the search for good approximations for the same, Single Water Entropy (SWE) being the approach adopted in this thesis. The initial work involved the implementation of the SWE calculator in C++ and further testing on the system of supercooled water. The insight of an entropy estimate at a single particle level is used to analyze HDL and LDL phases. From the fluctuations in SWE at lower temperatures, their relation to structural polymorphism was studied with the help of structural order parameters. To validate the approach of SWE and to see if the approach conserves configurational entropy, the same approach was applied to an ideal monoatomic gas system which showed that the entropy estimate by SWE was able to capture the nature of the ideal gas entropy trends.

Contents

Abstract **xi**

1 Introduction **1**

2 Theory **5**

2.1 Molecular Dynamics Simulations 5

2.2 Single Water Entropy 6

3 Supercooled Water **11**

3.1 Simulation details 11

3.2 Trajectory validation 12

4 Supercooled Water : Results **15**

4.1 Single Water Entropy Analysis 15

4.2 Structure Parameters 17

5 Ideal Gas Entropy **25**

5.1 Simulation Details 25

5.2 Results 26

6 Conclusion **29**

List of Figures

2.1 Ordering of trajectory with distance from box center 6

2.2 Permutation of Trajectory 7

2.3 Method of Permutation 8

3.1 Variation of TIP5PE density with temperature 12

3.2 Variation of TIP5PE C_p with temperature 13

4.1 SWE of 1^{st} hundred water molecules 16

4.2 Zeta parameter example 18

4.3 Zeta for 3 molecules at different temperatures. 20

4.4 Zeta parameter variation with temperature 21

4.5 Local Structure Index Example 22

4.6 Variation of LSI with temperature 23

5.1 Ideal gas single particle trajectory with different boundary conditions 27

5.2 Variation of Single Particle Entropy with box size for ideal gas 28

List of Tables

4.1 Table of average and standard deviation of SWE 15

5.1 Table of system configuration for the sets of Ideal Gas simulations 26

Chapter 1

Introduction

Water continues to be a system of active research, due to its wide applications in biochemical sciences. The properties of water as a liquid showing anomalies when compared to other liquids also spark interest in this field. The density maximum observed at 4°C at 1atm pressure with a negative thermal expansion coefficient at temperatures lower than 4°C have critical real life implications[1]. Similar is the anomaly in the isobaric specific heat capacity (C_p) that increases below 35°C. A more direct topic of study would be the role of water in protein folding and enzyme binding to substrates. The control and characterization of hydrophobic and hydrophilic structures continue to be an active topic of research in this regard.

In the case of a protein binding, the entropic role of water as a solvent could provide a better insight into the system. While there exist good methods of calculating protein entropy[2] , the solvent entropy still remains a challenge. Unlike the atoms of a protein, the solvent molecules diffuse throughout the entire simulation box. This results in the phase space of the solvent system being very large and sampling by diffusion being a relatively slow process results in the poor sampling of the phases space. There are approaches to calculating the entropy from radial distribution functions of solvents, such as the excess entropy[3] based on the Inhomogeneous Solvation Theory (IST) and from the liquid density of states by the 2PT[4] approach. But neither of these approaches provides an estimate of entropy on a single particle level.

When using the Single Water Entropy approach, it is possible to associate an entropy

value for each solvent molecule in the system. This opens the door to studying the entropic variation along the hydration shell around multiple systems of interest such as hydrophobic particles[5] and ions[6].

Supercooled Water

Experimentally when cooling water from room temperature at atmospheric pressure, water turns to ice at 0°C. But it is possible to cool water below 0°C without undergoing a liquid to solid phase transition and the liquid water thus obtained is called supercooled water. When trying to simulate the same process it is found to be very difficult to crystallize water in an unbiased simulation. This is because, even though the crystal ice is the thermodynamically more stable structure below 0°C, the transition between liquid to solid phase encounters a potential barrier that is kinetically hindering the transition.

This provides a unique opportunity to probe the supercooled state of water through molecular dynamics simulation, which has resulted in the hypothesis regarding the second liquid-liquid critical point(LLCP) of water in the supercooled regime.[7] While this region hasn't to date been accessible experimentally, a lot of work through MD simulations have characterized this phase space region.[8] [9]. Various structure parameters have been proposed to characterize the 2 liquid phases in this region - High Density Liquid(HDL) and Low Density Liquid (LDL).[10] [11]. This work attempts to utilize the tool of SWE to obtain an entropic insight into these two phases of water.

Configurational Entropy

As will be discussed in Chapter 2, the method of single water entropy calculation involves a permutation reduction step - A relabeling of particles within a trajectory to localize individual indices to a region in space. This permutation helps increase the sampling of the phases space. But it also raises the question of whether the permutation results in *"the loss of configurational entropy of the system"*. While a proper mathematical framework must be developed to answer this question in full, a simple test on Ideal Gas was performed in Chapter 5.

Since ideal gas systems have only configurational entropy and are exactly solvable, the results of the SWE approach could be compared to see if the trends in configurational entropy are conserved.

Chapter 2

Theory

2.1 Molecular Dynamics Simulations

A Molecular Dynamics simulation (MD) refers to a computer simulation of atoms or molecules treated as point particles with their interactions amongst themselves modelled according to various types of force fields. Unlike the quantum mechanical approach of solving the Schrödinger wave equation, the systems are treated classically, with the parameters used in force fields obtained from experimental and quantum calculations. Thus the trajectory is obtained by integrating the Newton's equations of motion for the system of particles in the box. For an n particle system the force equation would have been follows

$$\mathbf{F_i} = -\nabla_i V(\mathbf{X}_1, ..., \mathbf{X}_n) = m_i \frac{d^2\mathbf{X_i}}{dt^2} = m_i \mathbf{a}_i(t) \tag{2.1}$$

where $\mathbf{X}_i$ represents the position of the i^{th} atom and V the potential due to the bonding and non-bonding interactions between atoms parameterized according to the forcefield used.

Since the equation of motion is a second order partial differential equation, the leap frog integrator of GROMACS was used for solving the trajectories. Further simulation details are provided in the simulation section.

2.2 Single Water Entropy

The calculation of Single water entropy from the the MD trajectory is done in 3 steps.

- Trajectory ordering
- Trajectory permutation
- Entropy calculation

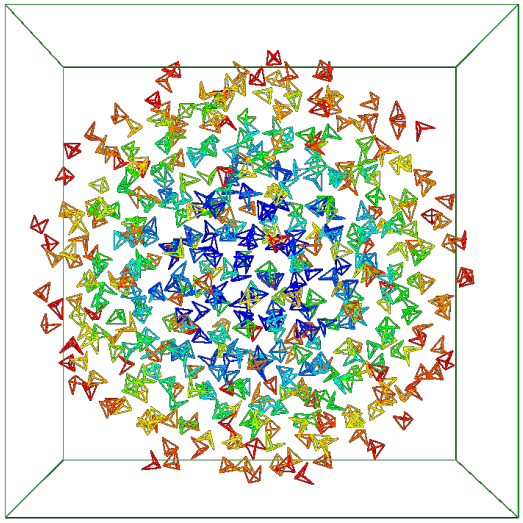


Figure 2.1: *Figures shows a single frame of a simulation after* `trjordern` *was performed with distance to center of the box*

The previous implementation of SWE by Dr.Arnab Mukherjee[5] utilized GROMACS-4.0.7 trajectory inputs. It was written in C with no parallelization. This section discusses the theory and some implementation details of the new SWE calculator written in C++ for GROMACS 2021.2.

2.2.1 Trajectory Ordering

When studying a large system the region of interest for entropy calculation would sometimes be small. Performing permutation calculation on the entire system would be computationally

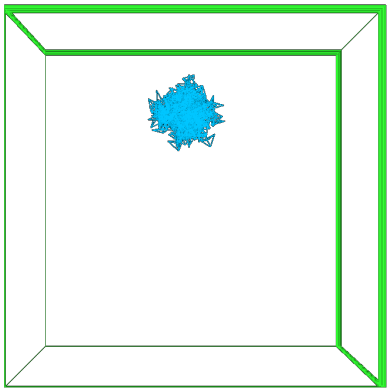


Figure 2.2: *Figure shows the trajectory of a single particle, after it underwent permutation. The single particle is localized to a region in space*

expensive, so there is a need to control the region into which the permutation could be performed. For this, the GROMACS module `trjorder` could be utilized to re-index the molecules in a trajectory based on their distance to a target residue. In the case of the bulk water system, there is no reference residue to calculate the distance from to reorder, so `trjordern` was created by modifying the original `trjorder` module to accept a vector representing a point in the simulation box about which reordering could be performed.

2.2.2 Permutation

Due to the large phase space available for even a small system of 1000 water molecules, the sampling achieved by diffusion of solvent is very poor in realistic timescales of simulations. To improve the sampling the method of permutation of solvent molecules as suggested by Grubmüller,et al[12] is utilized.

Theory

In a real world system water molecules could be thought of as indistinguishable particles. Whereas in the case of molecular dynamics simulations, due to the indexing of molecules they are distinguishable thus requiring each molecule to sample the entire simulation volume to sample the entire phase space. By invoking the indistinguishability of water molecules

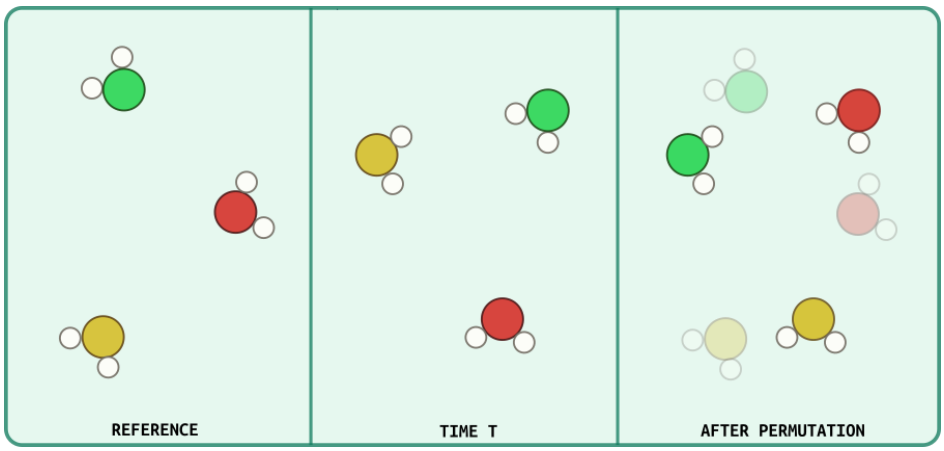


Figure 2.3: *The first panel shows the reference frame, while the second a general frame at time t. The third panel shows the reindexing(re-colouring) of the molecules based on the distance to molecules in the reference frame.*

each water molecule can contribute to the sampling of all other water molecules.

For the case of an N particle system, let the trajectory be represented by the $3N$ dimensional vectors $\mathbf{c}(t)$. A reference configuration $\mathbf{c_0}$ is chosen. For each $\mathbf{c}(t_i)$ in trajectory a permuted configuration is obtained by minimizing the root mean squared deviation between the reference configuration and the permuted configuration.

$$|\pi(\mathbf{c}(t_i)) - \mathbf{c_0}| \stackrel{!}{=} min \tag{2.2}$$

where π represents the permutation operation. This particular permutation problem can be rephrased as a Linear Assignment Problem (LAP) for which solutions already exist. After the permutation is performed the trajectory of a single particle becomes localised to a smaller region in space.

Implementation

The permutation module `g_permutemn` was written in C++, utilising the `libgmx.so` library. The menu driven program asked the user for the reference structure as well as the selection of particles to permute from and into. The last option enables the program to permute into a smaller population from a much larger population, reducing the computational cost. By default, the reference structure was taken to be the first frame of the input trajectory.

For solving the LAP, the `lap` module[13] provided by R. Jonker was utilized. The permuted trajectory along with the optional center of mass trajectory could be outputted into *xtc* or *trr* file formats. The most computationally expensive step of this implementation was the LAP solver, but the program has been written to be parallelizable across frames of the trajectory hence speeding up the calculation.

2.2.3 Entropy Calculation

The single water entropy is calculated from the permuted trajectory.

Theory

As described in the subsection 2.2.2 the trajectory of a single particle becomes localized to a region in space after permutation. The configurational volume available for each particle will vary from particle to particle. The translational entropies of particles could now be approximated by the N particle Quasi Harmonic (QH) approximation[2].

Given the QH approximation, the probability density of atoms can be approximated by a multivariate gaussian distribution given as

$$p(x) = (2\pi)^{-\frac{3N}{2}} \det \mathbf{Q} \exp\Big(-\frac{1}{2}(x - \langle x \rangle)^T \mathbf{Q} (x - \langle x \rangle)\Big) \tag{2.3}$$

where $\mathbf{Q}$ is the precision matrix of the particles trajectory calculated as

$$\mathbf{Q} = \big[(x - \langle x \rangle)(x - \langle x \rangle)^T\big]^{-1}$$

For each particle , a 3x3 covariance matrix (inverse of the precision matrix) is thus constructed, from which 3 eigen values($\lambda_i, i \in 1, 2, 3$) are obtained. They are mapped to their corresponding frequencies given by

$$\omega_i = (k_B T/\lambda_i)^{1/2}$$

The translational entropy is calculated as

$$S_{trans}^{QH} = k_B \sum_{i=1}^{3} \frac{\hbar\omega_i}{k_B T \left(e^{\hbar\omega_i/k_B T} - 1\right)} - ln\left(1 - e^{-\hbar\omega_i/k_B T}\right) \tag{2.4}$$

with T being the temperature of the system, k_B the Boltzmann constant and $\hbar = h/2\pi$ where h is the Planck's constant. For the calculation of the non-Gaussian component contribution to the entropy, the method of k^{th} Nearest Neighbour (kNN)[14] distances as done by Numata, et al[15] was implemented.

For the rotational entropy of the system about a reference particle, the standard equation of entropy based on the angular distribution of probabilities can be calculated[16] for each of the permuted water molecules. The rotational entropy would be given as

$$S_{rot}^{QH} = -k_B \int \rho(\theta,\chi) cln\{\rho(\theta,\chi)c\} sin\theta d\theta d\chi \tag{2.5}$$

where θ is the angle between the water oxygen to reference particle vector and the dipole vector of water. χ is defined as the angle between the vector joining the 2 hydrogens of water and the normal to the plane defined by the reference particle-oxygen vector and the dipole vector.

Implementation

The final calculation was implemented as a single module `g_transent`, which calculates the covariance matrix for each of the selected molecules in the permuted trajectory. Its eigenvalues are calculated using the *eigen* library which is then used to calculate the Quasi-harmonic translational entropy. The k^{th} nearest neighbour calculations were implemented using the `distances` package provided by Fredrik Savje[17]. The same module is capable of calculating the rotation entropy along with the translational entropy in a single read of the trajectory.

Chapter 3

Supercooled Water

3.1 Simulation details

The supercooled water simulations were performed using GROMACS-2021.2[18]. The water model used for the simulation was TIP5P-E[19] using the OPLS-AA forcefield as it has been shown to provide accurate density measurements during simulations. An initial box of 1000 water molecules was generated in a cubic box of side length 3.14nm using the `solvate` module of GROMACS. The system was then energy minimized using the steepest descent followed by the conjugate gradient descent algorithm. The energy minimised system was then brought up to 300K using the Berendsen thermostat[20] with a coupling constant of 0.2ps. Afterwards, an NPT equilibration at 300K temperature and 1 bar pressure was performed for 2ns utilizing the Nose-Hoover thermostat and the Parrinello-Rahman barostat with coupling constants of 0.4 and 0.8 ps respectively.

The equilibrated system thus obtained was subject to heating/cooling to reach its target temperatures using the Nose-Hoover thermostat with 0.4ps temperature coupling. In the current study 190K, 200K, 210K, 220K, 230K, 240K, 250K, 260K, 273K, 285K, 300K, 315K and 330K were the temperatures under consideration. Next, an NPT equilibration was performed for each of the 13 systems for 30ns at 1 bar pressure and respective temperature using the Nose-Hoover thermostat and the Parrinello-Rahman barostat with 0.4ps and 0.8ps coupling respectively.

The final production run was an NPT simulation 1μs long using the above mentioned thermostat and barostat using the same coupling constant with an integrator timestep of 2fs. The trajectory coordinates were outputted every 0.4ps. The electrostatics were calculated using Particle Mesh Ewald(PME) electrostatics with a 10 Å cutoff for both van der Walls and electrostatics interactions.

3.2 Trajectory validation

To validate the trajectories generated few characteristic properties of the system were calculated

3.2.1 Density

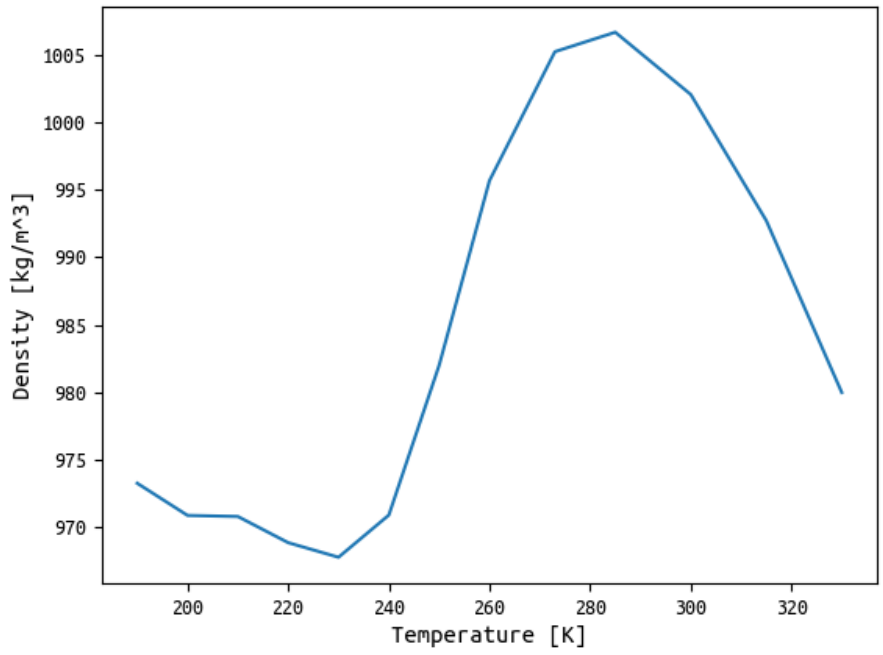


Figure 3.1: *Figure shows the variation of density with temperature*

The density plot was calculated using the `energy` module of GROMACS. A density maximum at 285K can be observed along with the decreasing density with temperature. A

density minima can be observed for 230K beyond which a decrease in temperature results in a slight increase in density.

3.2.2 Specific heat capacity

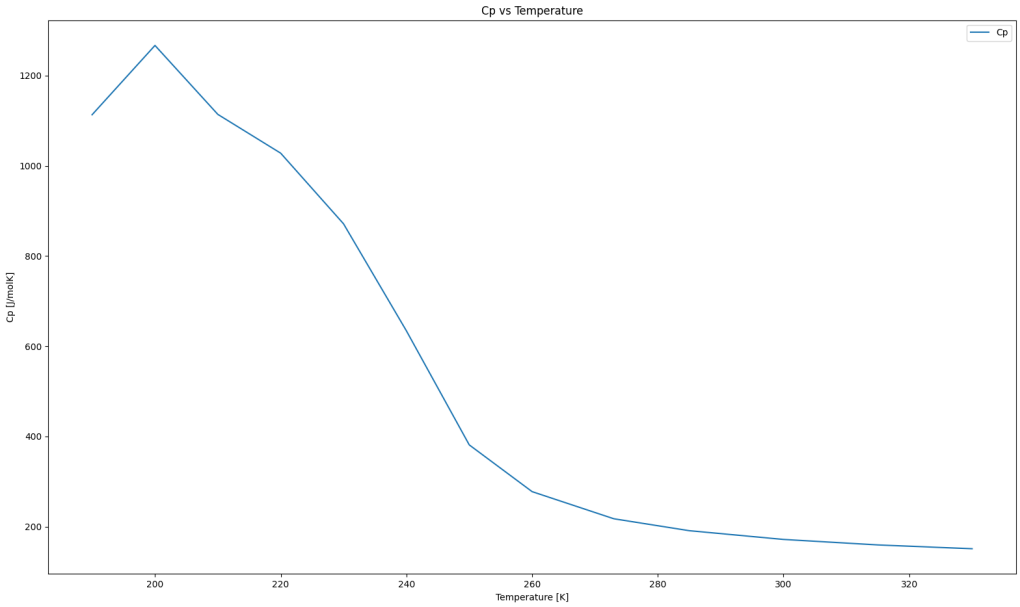


Figure 3.2: *Figure shows the variation of isobaric specific heat capacity with temperature*

The specific heat capacity was calculated by initially obtaining the enthalpy fluctuations of the simulation using the `energy` module from which σ_H was calculated. Isobaric specific heat capacity was calculated as

$$C_p = \frac{\sigma_H^2}{k_B T^2} \tag{3.1}$$

From the figure a peak at 200K can be observed. This corresponds to the crossing of the widom line.

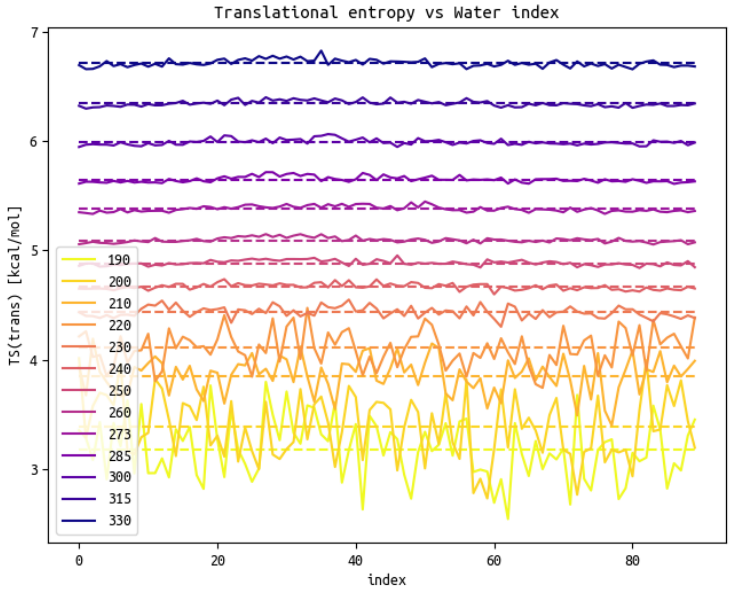
Chapter 4

Supercooled Water : Results

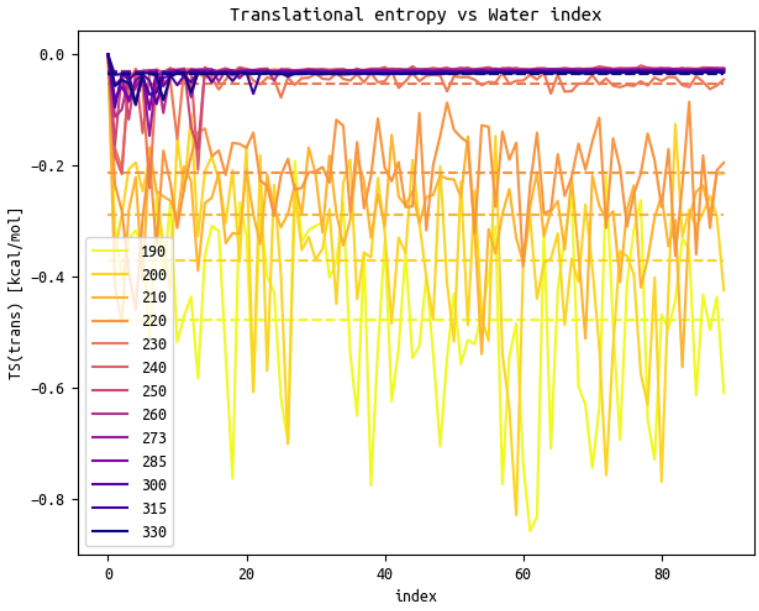
4.1 Single Water Entropy Analysis

Table 4.1: *Average and standard deviation of SWE of the* 1^{st} *100 water molecules for various temperatures*

Temperature	Translational Entropy		Rotational Entropy	
	Mean	Std dev	Mean	Std dev
190	3.182	0.287	-0.482	0.171
200	3.384	0.276	-0.373	0.153
210	3.847	0.170	-0.291	0.085
220	4.112	0.156	-0.214	0.073
230	4.434	0.048	-0.052	0.031
240	4.663	0.029	-0.034	0.032
250	4.882	0.023	-0.031	0.024
260	5.092	0.022	-0.030	0.015
273	5.376	0.025	-0.031	0.014
285	5.646	0.026	-0.032	0.011
300	5.987	0.025	-0.033	0.009
315	6.341	0.024	-0.034	0.009
330	6.708	0.030	-0.035	0.008



(a)



(b)

Figure 4.1: *Figure a shows the variation of S_{Tr} of the 1^{st} hundred water molecules while Figure b shows the variation of S_{Rot}*

The Translational single water entropy for the 1^{st} 90 water molecules in the system is shown in Figure 4.1a for various temperatures from which a general trend of decrease in translational entropy with decrease in temperature can be observed. Similarly, in figure 4.1b Rotational entropy is found to be less negative at higher temperatures. The mean and standard deviation of the entropy values are provided in Table 4.1.

Another observation made is the larger fluctuations in translational and rotational entropy values at the lower temperatures captured by the larger standard deviation values. This opens up the possibility of categorising water the water molecules into high entropy and low entropy water molecules and to observe of any correlation between these entropy values and HDL/LDL states.

4.2 Structure Parameters

To characterize the 2 different phases, structure parameters have to be utilized. Two such structure parameters are used to analyze the trajectory.

4.2.1 Zeta Parameter

Theory and implementation

Consider the case of 4 water molecules hydrogen bonded with water indexed 0. The ζ parameter[10] for water 0 is defined as the difference between the closest non-hydrogen bonded water molecule and the farthest hydrogen bonded water molecule. A larger value of ζ suggests a region of low local density whereas a smaller value points to regions of high density.

`g_zeta` was a C++ program utilizing the GROMACS library `libgmx.so` written to calculate the ζ parameter of the water molecules provided as a selection for the entire trajectory. It would also calculate the number of hydrogen bonds of the water molecules under consideration and print to files.

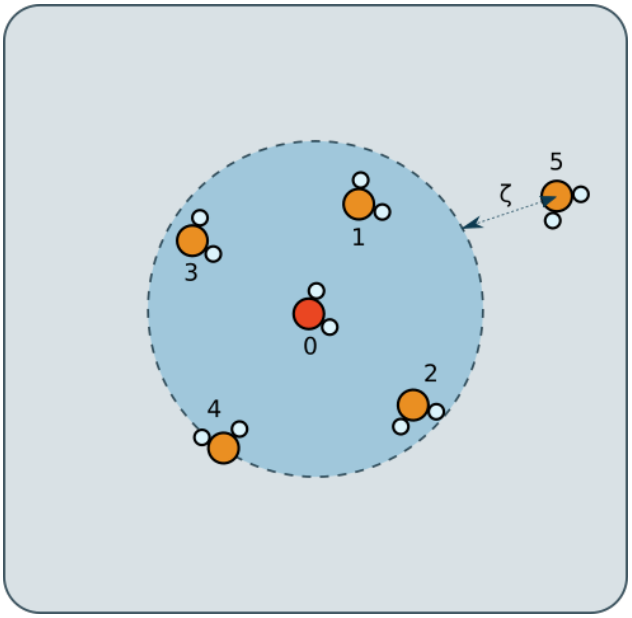
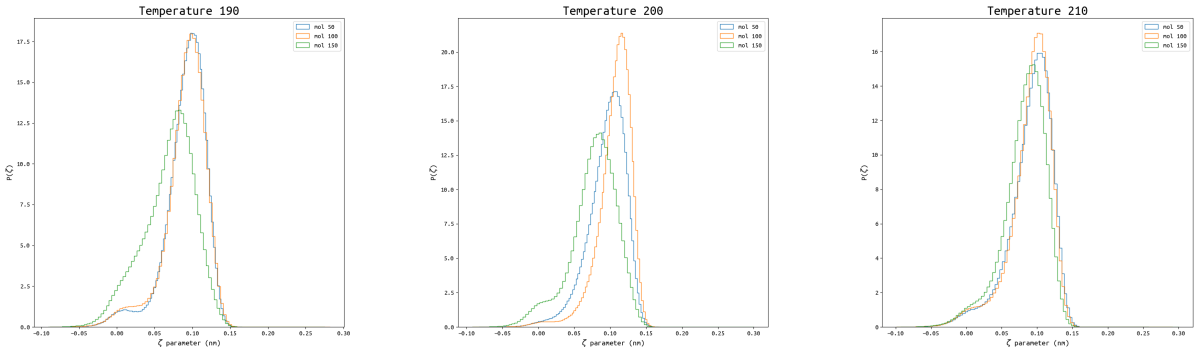


Figure 4.2: *Figure shows the ζ parameter among 5 water molecules. Molecules 1 to 4 are hydrogen bonded to water 0 with 4 being the farthest from 0. Molecule 5 is the closest non-hydrogen bonded water.*

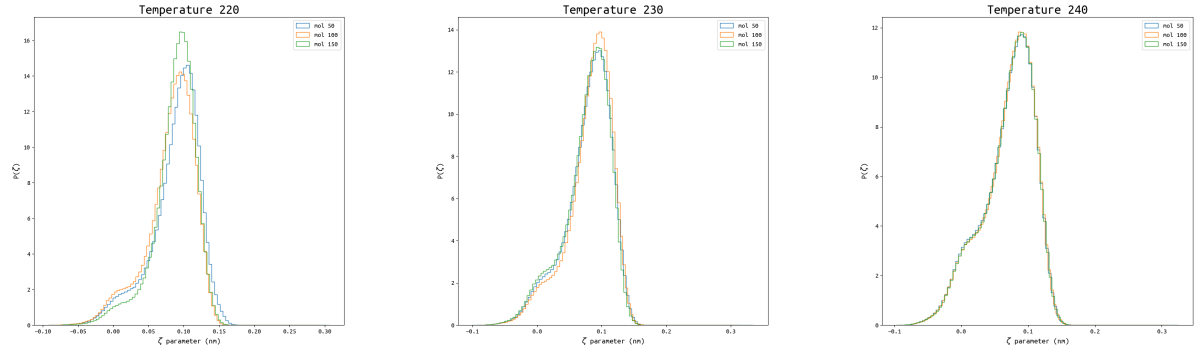
Results

From 4.3 it is seen that the distribution of zeta isn't much varied upon the choice of water molecule. So with a random reference molecule Figure 4.4 was constructed. At a lower temperature, a peak is observed the ζ parameter peaks at 1 Å. While as the temperature is increased a secondary peak at a negative ζ value is observed with a corresponding fall in probability for the 1 Å peak. This indicates that at higher temperatures water tends to remain in the HDL state while at lower temperatures there is a shift in preference to the LDL state. Interestingly even with the broader distribution of ζ at higher temperatures, the entropy doesn't fluctuate as much. Whereas in the case of lower temperature where we observe primarily an LDL region water molecules show different entropic behaviour. To check for the existence of a correlation between the fluctuations in entropy and structural polymorphism, 100 water molecules were classified into either High entropy or Low entropy waters and their distribution of zeta was calculated. This procedure was performed for 2 different temperatures - 190K and 240K.

A secondary peak around 2.7 Å is observed which increases in height with an increase in temperature. To explain the reason for this secondary peak the number of hydrogen bonds

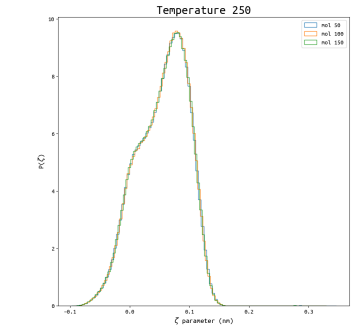


(a) (b) (c)

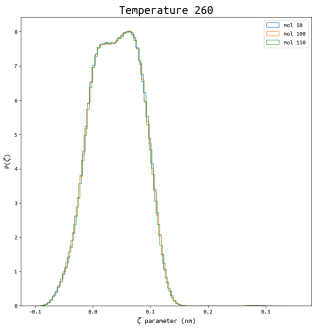


(d) (e) (f)

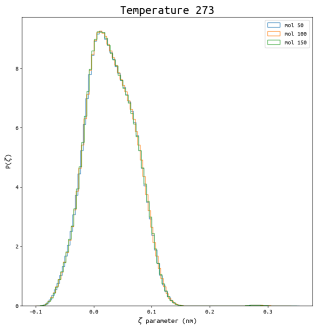
for molecules when ζ was greater than 2.7 Å was checked and found to be 0. If a water molecule doesn't form hydrogen bonds, the largest hydrogen bond distance collapses to 0. So the ζ parameter becomes just the distance to the closest non-hydrogen bonded water which is essentially the closest water molecule in this case. At higher temperatures, the water molecules have higher kinetic energy, enabling them to form non-hydrogen bonded structures more frequently, which is captured by the secondary peak.



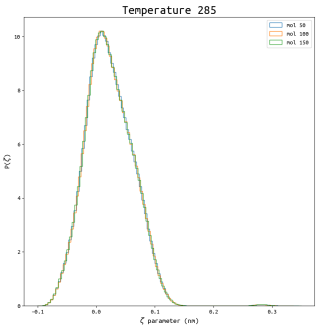
(g)



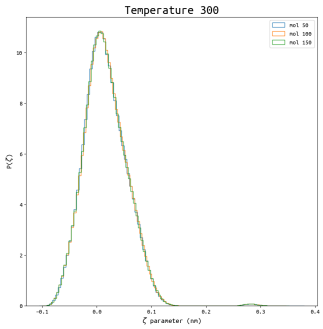
(h)



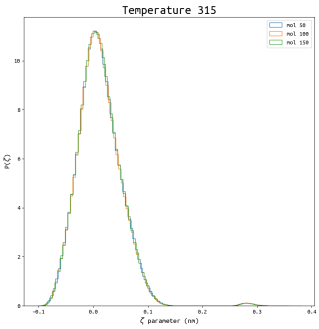
(i)



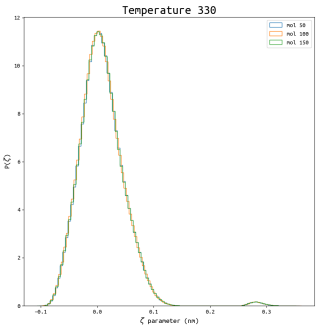
(j)



(k)



(l)



(m)

Figure 4.3: *The distribution of ζ parameters for molecules 50 , 100 and 150 showing similar distribution regardless of molecules chosen for different temperatures*

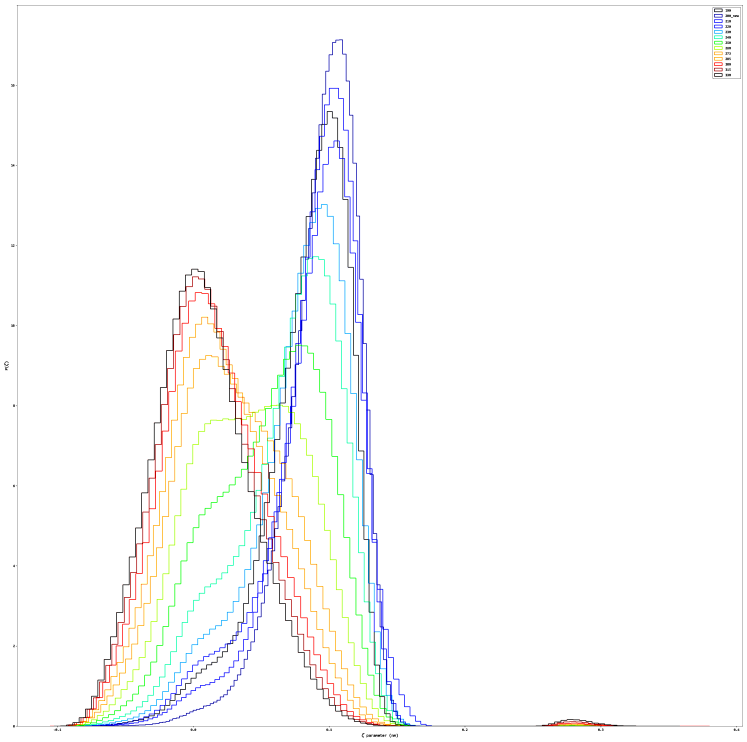


Figure 4.4: *Figure shows the variation of the ζ parameter with temperature for a molecule chosen at random*

4.2.2 Local Structure Index

.

Theory and Implementation

Local structure index (LSI)[11] is a structure factor that measures the extent of the gap between the second and first hydration shell of a water molecule. Consider the case of a water molecule indexed k with n water molecules at an oxygen-oxygen distance lesser than 3.7 Å from it. Let r_j represent the distance of the j^{th} water molecule from the central

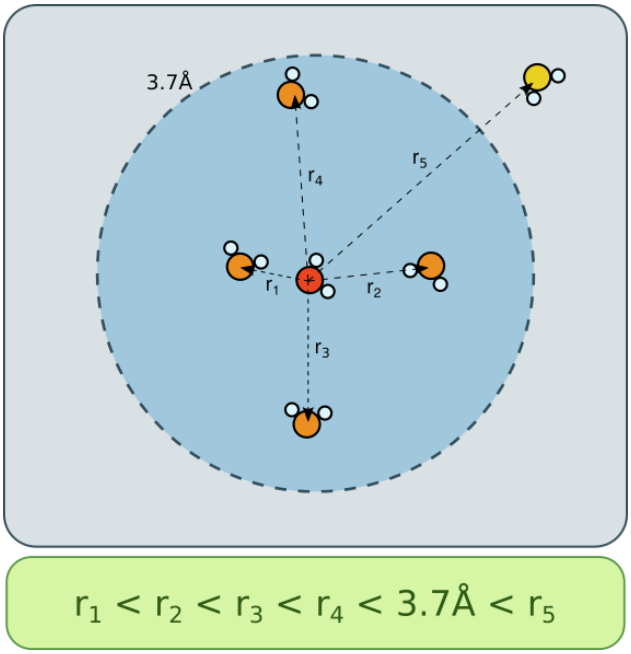


Figure 4.5: *Figure shows the arrangement of molecules according to distance from the reference molecule coloured red for LSI calculation*

molecule. If the distances are ordered as

$$r_1 < r_2 < r_3 < ..r_j < ..r_n < 3.7\,\text{Å} < r_{n+1}$$

the LSI is defined as

$$LSI = \frac{1}{n}\sum_{i=1}^{n}(\Delta(i) - \overline{\Delta})^2 \tag{4.1}$$

where

$$\Delta(i) = r_{i+1} - r_i \tag{4.2}$$

and

$$\overline{\Delta} = \frac{1}{n}\sum_{i=1}^{n}\Delta(i) \tag{4.3}$$

A lower value of LSI corresponds to a more tightly packed structure much like the HDL water, whereas a larger LSI will correspond to the existence of LDL water. `g_LSI` was implemented in C++ to calculate the LSI for a user provided selection of water molecules utilizing the `libgmx.so` library.

Results

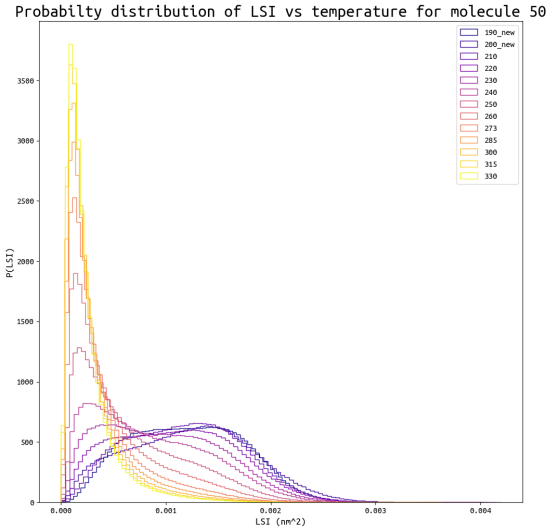


Figure 4.6: *Figure shows the variation of LSI with temperature for a single particle*

From Figure 4.6 it is observed that at higher temperatures there is a sharp peak close to 0 that decays, complementing the results obtained from the zeta parameter analysis that water mostly exists in the HDL form at high temperatures. With the lowering of temperatures the sharpness of this peak decreases with a shift to a new peak at 0.0015 nm^2 having an almost bimodal distribution. A similar analysis to that of the zeta parameter, involving the calculation of the zeta parameter for the high and low entropy populations of water separately are being performed.

Chapter 5

Ideal Gas Entropy

5.1 Simulation Details

The ideal gas simulation was performed using LAMMPS[21], in real units. The atoms were modelled to have no interactions between themselves. Two sets of simulations were performed, first varying the volume but keeping the number of particles a constant and the other varying the number of particles at constant volume. The details of choices of volume and number of particles are provided in table 5.1a. All simulations were performed at 300K temperature with the help of Nose-Hoover thermostat for 500ps with a timestep of 1fs.

5.1.1 The choice of boundary conditions

Unlike the case of the water simulations, Periodic boundary conditions were not applied to the system, but rather walls were applied. Since the ideal gas particles do not interact with each other, they will continue to move with the initial velocity throughout the simulation without walls. This results in repeating vacancies inside the simulation box that aren't visited as can be seen in Figure 5.1a. To avoid this walls were used as boundaries to reflect the particles at the same speed. The resulting trajectory of a single particle in time is shown in Figure 5.1b

Table 5.1: *Table of simulation parameters for the 2 set of ideal gas simulations performed*

(a) SET 1

Temperature (K)	Volume(nm^3)	Number of particles
300	27	1000
300	27	2000
300	27	3000
300	27	5000
300	27	7000
300	27	10000

(b) SET 2

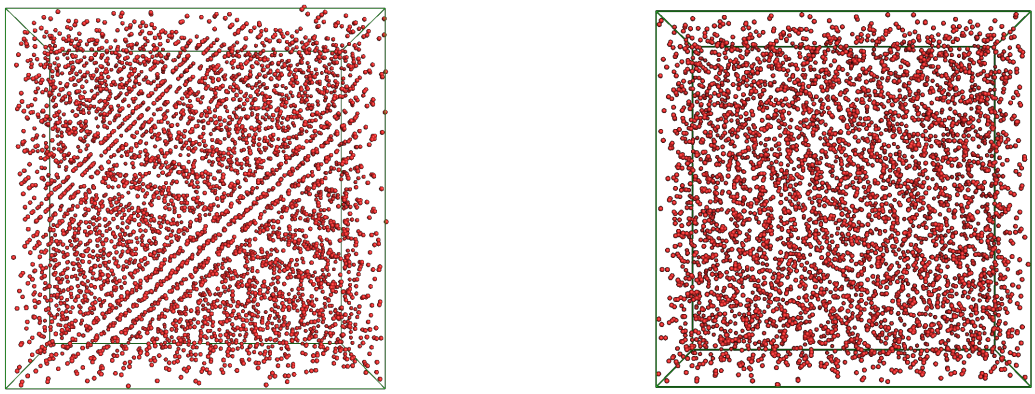
Temperature (K)	Volume(nm^3)	Number of particles
300	27	1000
300	27	2000
300	27	3000
300	27	5000
300	27	7000
300	27	10000

5.1.2 Trajectory post-processing

Since the simulations were performed in LAMMPS, the resulting trajectory was obtained as a LAMMPS dump file. It is an ASCII file of tabular position and velocity data of all particles for all times. Since SWE module was implemented to read GROMACS 2021.2 trajectory, the LAMMPS dump had to be converted to an *xtc* file. For this a new module `dumpconverter` was written in C++, which took the dump file as input and outputted the trajectory in either *xtc* or *trr* format along with the first frame of the trajectory in a *gro*/*pdb* format according to user preference.

5.2 Results

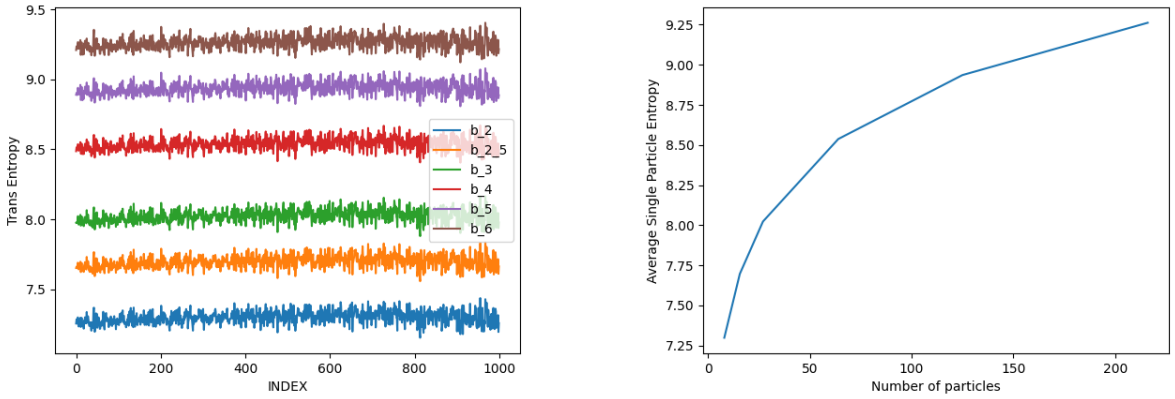
From the simulation the ideal gas entropy values calculated showed that for a single system there is very little fluctuation in translational entropy from particle to particle. The next clear observation was the trend of increasing entropy with the volume of the simulation box. This can be more easily observed from the Figure5.2b. The mean values are valid due to the



(a) (b)

Figure 5.1: *Ideal gas single particle trajectory with different boundary conditions*

small fluctuations in entropy values. The results obtained correlate with the expected $ln(V)$ trend from the Sackur-Tetrode equation. The analysis of the second set of simulations is being performed, but the computationally intensive permutation on the entire system makes the progress slow.



(a) (b)

Figure 5.2: *The Single Particle Entropy calculated for ideal gas. Figure 5.2a shows the Single particle entropy for 1000 particles for different box sizes. Figure 5.2b shows the variation of the average single particle entropy with volume*

Chapter 6

Conclusion

The thesis work involved the implementation of an automated pipeline for the calculation of Single Water Entropy, after a GROMACS simulation from start to finish. Small parts of the implementations were CPU parallelized using *OpenMP*. Further tools were developed for the conversion of trajectories from LAMMPS to GROMACS and vice versa. From the simulation of the supercooled liquid, the populations of high and low entropy water were observed at very low temperatures, which could be mapped to their spacial heterogeneity using structure parameters. Further statistical analysis based on these order parameters are being studied. The next steps would involve determining the solvent entropy in the hydration shell of ions in supercooled water.

From the simulations of the ideal gas system, the SWE approach was observed to capture the logarithmic increase of entropy with the volume of the system. While a single system is insufficient to validate the general statement that *"configurational entropy is conserved during the single water entropy calculation"*, it could be used as a starting point to test out more complicated systems with exact solutions of entropy known for the conservation of configurational entropy. Testing on newer, exactly solvable systems could also provide an opportunity to introduce a correction factor to the current system to better calculate more numerically accurate values.

Finally, with the current implementation of the entropy pipeline, calculations of water entropy along model surfaces of varying curvature and surface charge densities could be tested. Extending this as a template for single water entropy along protein surface patches

is an avenue warranting further studies.

Bibliography

[1] Pablo G. Debenedetti. Supercooled and glassy water. *Journal of Physics Condensed Matter*, 15(45):R1669–R1726, November 2003.

[2] Jürgen Schlitter. Estimation of absolute and relative entropies of macromolecules using the covariance matrix. *Chemical Physics Letters*, 215(6):617–621, 1993.

[3] Ruchi Sharma, Somendra Nath Chakraborty, and Charusita Chakravarty. Entropy, diffusivity, and structural order in liquids with waterlike anomalies. *The Journal of Chemical Physics*, 125(20):204501, nov 2006.

[4] Shiang-Tai Lin, Mario Blanco, and William A. Goddard. The two-phase model for calculating thermodynamic properties of liquids from molecular dynamics: Validation for the phase diagram of lennard-jones fluids. *The Journal of Chemical Physics*, 119(22):11792–11805, 2003.

[5] Wilbee D. Sasikala and Arnab Mukherjee. Single water entropy: Hydrophobic crossover and application to drug binding. *The Journal of Physical Chemistry B*, 118(36):10553–10564, 2014. PMID: 25133988.

[6] Debasis Saha and Arnab Mukherjee. Impact of ions on individual water entropy. *The Journal of Physical Chemistry B*, 120(30):7471–7479, 2016. PMID: 27404917.

[7] Peter H. Poole, Francesco Sciortino, Ulrich Essmann, and H. Eugene Stanley. "phase behaviour of metastable water". *Nature Science*, 360(6402):324–328, November 1992.

[8] Francesco Sciortino, Emilia La Nave, and Piero Tartaglia. Physics of the liquid-liquid critical point. *Phys. Rev. Lett.*, 91:155701, Oct 2003.

[9] V. Holten and M. A. Anisimov. Entropy-driven liquid–liquid separation in supercooled water. *Scientific Reports*, 2, Oct 2012.

[10] Tanaka Hajime Russo John. Understanding water's anomalies with locally favoured structures. *Nature Communications*, 5(1):3556, 2014.

[11] Eli Shiratani and Masaki Sasai. Growth and collapse of structural patterns in the hydrogen bond network in liquid water. *The Journal of Chemical Physics*, 104(19):7671–7680, 1996.

[12] Friedemann Reinhard and Helmut Grubmüller. Estimation of absolute solvent and solvation shell entropies via permutation reduction. *The Journal of Chemical Physics*, 126(1):014102, 2007.

[13] Volgenant A Jonker R. A shortest augmenting path algorithm for dense and sparse linear assignment problems. *Computing*, 38(4):325–340, 1987.

[14] Harshinder Singh, Neeraj Misra, Vladimir Hnizdo, Adam Fedorowicz, and Eugene Demchuk. Nearest neighbor estimates of entropy. *American Journal of Mathematical and Management Sciences*, 23(3-4):301–321, 2003.

[15] Knapp EW Numata J, Wan M. Conformational entropy of biomolecules: beyond the quasi-harmonic approximation. *Genome informatics*, 18:192–205, 2007.

[16] Themis Lazaridis and Michael E. Paulaitis. Entropy of hydrophobic hydration: a new statistical mechanical formulation. *The Journal of Physical Chemistry*, 96(9):3847–3855, 1992.

[17] Fredrik Savje. Tools for distance metrics, 2019.

[18] Lindahl, Abraham, Hess, and van der Spoel. Gromacs 2021.2 source code, May 2021.

[19] Steven W. Rick. A reoptimization of the five-site water potential (tip5p) for use with ewald sums. *The Journal of Chemical Physics*, 120(13):6085–6093, 2004.

[20] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, 81(8):3684–3690, 1984.

[21] A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. in 't Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trott, and S. J. Plimpton. LAMMPS - a flexible simulation

tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Comp. Phys. Comm.*, 271:108171, 2022.