Title
Explaining Water's Surface

Author
Odendahl, Nathan L

Publication Date
2021

Peer reviewed|Thesis/dissertation

Explaining Water's Surface

by

Nathan L Odendahl

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Chemistry

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Phillip Geissler, Chair
Professor Richard Saykally
Assistant Professor Kranthi Mandadapu

Spring 2021

Explaining Water's Surface

Abstract

Explaining Water's Surface

by

Nathan L Odendahl

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor Phillip Geissler, Chair

The surface of liquid water is characterized by large fluctuations that are too complex to describe exactly but, like with other aspects of molecular liquids, can be described in a useful way when approached statistically. The water liquid/vapor interface is composed of a sharp drop in density and reverberations from that drop that extend about 1 nm from the surface. Transient but characteristic structure forms as water molecules realign to maximize hydrogen bonding without ready neighbors in vapor. Much of the structure of water's surface can be accounted for as a broadening of the bilayer structure of the basal plane of the ice crystal, particularly the layering in density and the orientational trends.

The mean field surface model predicts the average orientational behavior of structured molecular liquids, which is demonstrated in two dimensions with the water-like Mercedes-Benz (MB) model and in three dimensions with comparisons to the SPC/E and TIP5P models of water. Liquids with strong directional bonding tend to orient predictably when potential bonding neighbors are removed. The mean field surface model assumes that bonds are most likely to form in regions of high density, then uses a self-referencing iterative approach to determine whether neighbors are also likely to be oriented in a way that is favorable to bonding. Physically motivated changes to the reference density, bonding rules, and molecular structure clarify which aspects of molecular water are actually essential to the orientational structure at the interface. It is shown that moderate changes to bonding rules, like flexibility, strength, and asymmetry, have little impact on the main surface patterns. The layering in the density profile of water is by contrast necessary to accurately predict water's surface structure. All of these observations can be rationalized through liquid water's relationship to the structure of the basal plane of ice.

The dielectric response of water is also affected by the introduction of an interface. The Gaussian field model known as dielectric continuum theory[1–3] (DCT) makes predictions for the effect of an interface on the dielectric behavior of water. Using the method of images[4] and correcting for the dielectric saturation of a rigid water model, it is shown that polarization fluctuations near the air-water interface largely follow these predictions at scales as small as 5 Å despite the macroscopic character of DCT.

TO MATTHEW

```
if (Matthew.is_brother_to(Nathan))
{
        atoms.init();
        computers.init();
        Nathan.inspire(atoms);
        Nathan.inspire(computers);
}
```

Contents

List of Figures

List of Tables

Acknowledgments

The work of a PhD student is simultaneously exciting and mundane. For the people who I don't list but who supported me through the nitty-gritty and showed me how to enjoy it, thank you.

Phill Geissler was with me through every step of this research as an advisor and a mentor. I told him early that I needed a weekly meeting to succeed and, like clockwork, he was there to advise. He has a talent for communicating complicated concepts in a refreshingly straightforward way, and he showed me that a good explanation has to be built on a deep foundation of understanding. He helped me see that getting the details right is always worth it.

I am grateful to all the members of the Geissler group who I crossed with, though I'll only list a few here. Steve Cox had the patience to get me started on water's surface and Georg Menzl was with me the rest of the way. Georg is the key player behind all the work in Chapter 4. He is the kind of person that you can turn to for a great conversation that is both critical and productive. Layne Frechette also has my thanks for his genuine friendship. Layne has a real talent for science and for teaching, and I was lucky enough to benefit from both.

My family all live states away from me now, but they are the most important people in the life I have lived up to today. My dad has always been my example and has always taken my side. My mom spent every moment of her life caring for me and my siblings. My oldest brother used to sit with me and talk and teach and talk, and a conversation with him is never dull. Steve set an example for me, in his own way, and he will be missed. Thank you Annie, who gave me the boost in human knowledge that I needed to thrive.

And to Lucie, who I leaned on so heavily during graduate school. Thank you for supporting me, even when it was stressful for you. I look forward to many more days of stressing and leaning and graduating.

Chapter 1

Interfaces

Overview

Understanding the water-vapor boundary is an important step in the creation of a greater theory of liquid water. The Willard-Chandler Instantaneous Interface[5] method is a robust tool for time-resolved, molecular-level interfacial detail.

1.1 The importance of defining an interface

The molecular study of a liquid starts in the "bulk" region, where the edges of the liquid can be safely ignored. In physical experiments, bulk usually means that the volume of a liquid is much larger than its surface area and so surface effects are unimportant. In computational experiments, bulk usually implies periodic boundary conditions, but other computational boundaries are used.

Bulk is a useful starting point to conceptualize molecular liquids. However, an interface necessarily forms around any inhomogeneity and the "bulk" assumption needs modification. It is therefore essential to define an interface and to study the patterns around the newly formed surface.

An enormous number of liquids are available to study, but water is a particularly attractive choice. Water is essential to life and plays a role in nearly every aspect of biology, is ubiquitous on planet earth, is among the most common solvents used in chemistry, has a number of unique properties as a liquid, and already has a significant body of research around it. Water will be the main liquid examined throughout this work.

An even greater number of potential interfaces exists. Any liquid may form a liquid-solid, liquid-liquid, or liquid-vapor phase interface with any other chemical. Solutes of varying size, shape, and charge can interact with a liquid solvent. It would be impossible to study every conceivable liquid interface between every type of molecule, but it is possible to generate a set of principles that predict most interfacial behaviors. As a general case-study, the water-vapor interface is an excellent starting point and will be the main focus of this discussion.

Bulk water forms a 3-dimensional tetrahedral network of transient hydrogen bonds. The network is disrupted by a 2-dimensional plane to form a liquid-vapor interface. A quantitative measurement of interfacial behavior requires a definition of where water ends and vapor begins.

1.2 The Gibbs Dividing Surface

It would be convenient if all phases had a perfectly flat, perfectly well-distinguished boundary. In reality, there tends to be a region containing a mixture of phases–that is, an interfacial region which does not strictly resemble either phase, shown in the transition in density in Figure 1.1. The boundary is ambiguous in this framing and a line separating the two phases is hard to draw, but the region of ambiguity is precisely where the least familiar behavior will occur.

For a single component system, like liquid water and water vapor, the Gibbs dividing surface is effectively a flat line placed at the height z_{GDS} where density is halfway between its bulk liquid (ρ_ℓ) and bulk vapor (ρ_v) values,

$$\rho(z_{GDS}) = \rho^*, \tag{1.1}$$

$$\rho^* = \frac{1}{2}(\rho_\ell + \rho_v) \approx \frac{\rho_\ell}{2}, \tag{1.2}$$

$$\rho^*_{water} \approx 0.0166 \text{ Å}^{-3} \tag{1.3}$$

This definition is unfortunately dependent on a macroscopic measurement: the average density at each depth, $\rho(z)$, which is found by averaging over some combination of space and time. The need to average also limits the resolution of the interface. To create a more detailed interface, one could sub-sample across spatial dimensions x and y or across time t.

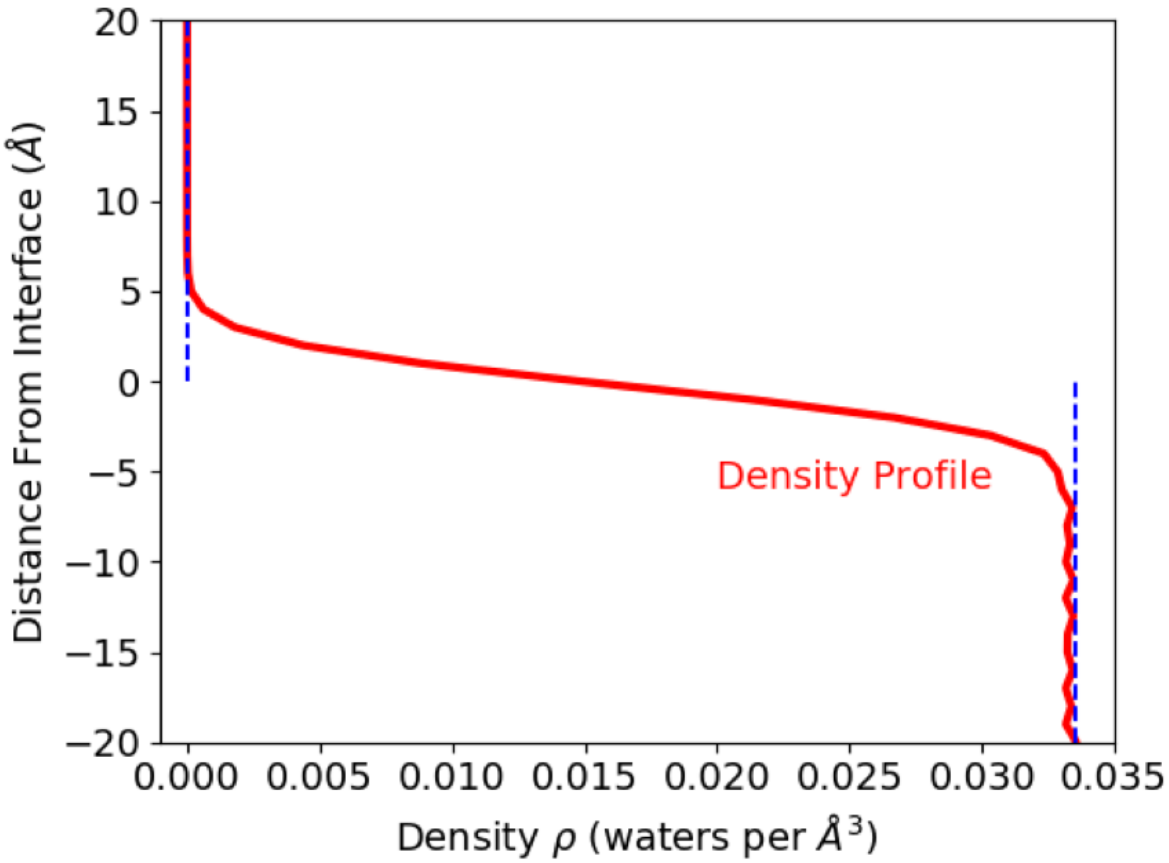

Figure 1.1: Average density profile of the water liquid-vapor boundary region. The density value is generated numerically by binning the z-coordinate of simulated water molecules. A liquid-vapor interface was initialized in the xy-plane and remained stable throughout the trajectory. Data was generated in a molecular dynamics simulation of SPC/E molecules under standard conditions.

1.3 The instantaneous interface

A molecular-level study of an interface would benefit from a rigorous, repeatable, and transferable definition of an interface that is continuous in both time and space. In 2010, Willard and Chandler proposed a useful method[5] for defining an instantaneous interface of water with a generic second phase.

Conceptually, the Willard-Chandler (WC) interface coarsens out the density of a water molecule i from its position centered at $\vec{r}_i = (x_i, y_i, z_i)$ to a spherically-symmetric region of density that decays with distance as a 3D Gaussian function. That is, the contribution at

$\vec{r} = (x, y, z)$ to a density field by molecule i is traditionally represented

$$\rho(\vec{r}, \vec{r}_i) = \delta(\vec{r} - \vec{r}_i) \tag{1.4}$$

and is replaced with a 3D Gaussian kernel with a decay length specified by length scale parameter ξ,

$$\rho_{WC}(\vec{r}, \vec{r}_i) = (2\pi\xi^2)^{-3/2} \exp\left(\frac{-|\vec{r} - \vec{r}_i|^2}{2\xi^2}\right). \tag{1.5}$$

These definitions of single-molecule density both have unit integrals,

$$\int d\vec{r}\rho(\vec{r}, \vec{r}_i) = \int d\vec{r}\rho_{WC}(\vec{r}, \vec{r}_i) = 1 \tag{1.6}$$

so no transformation of units or renormalization is implied.

A snapshot of N point-molecules in a liquid-like configuration can be multiplied (convolved) with the Gaussian density kernel to generate a smooth density field that varies in both space and time t,

$$\rho_{tot}(\vec{r}, t) = \sum_{i=1}^{N} \rho_{WC}(\vec{r}, \vec{r}_i, t). \tag{1.7}$$

An arbitrary density ρ^* specifies a transition from liquid to vapor. For liquid water under standard conditions, one-half of bulk liquid density is an effective choice,

$$\rho^* = \frac{1}{2}(\rho_\ell + \rho_v) \approx 0.0166 \text{ Å}^{-3}. \tag{1.8}$$

The smooth and continuous interface in Figure 1.2 is generated through the Willard-Chandler scheme for a water liquid-vapor simulation. Figure 2.1 displays the density as a function of depth from the instantaneous interface.

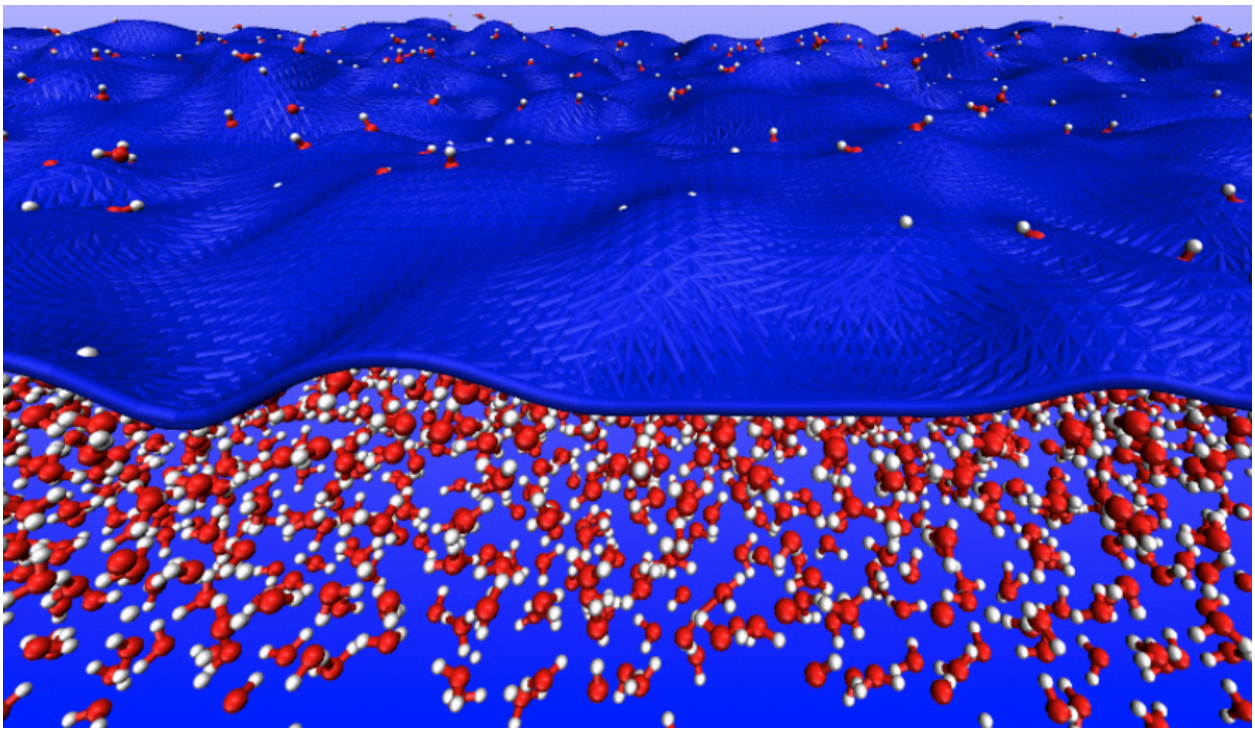

Figure 1.2: Instantaneous interface (dark blue surface) generated according to the Willard-Chandler method for a liquid-vapor simulation of SPC/E water (red and white spheres represent oxygen and hydrogen atoms, respectively).

Robustness of the WC interface

Part of what makes the instantaneous interface useful is its stability across a range of reasonable parameters. There are three parameters to fit: the density cutoff ρ^*, the effective length of density coarsening ξ, and the kernel for density coarsening, in this case the Gaussian function.

A variety of kernels could be chosen, although none are as apparently stable as the Gaussian. A handful of examples are shown in Figure 1.3. The Lorentzian kernel has both a longer decay length and a more narrow coarsening; this leads to a softer, broader interface when compared to a Gaussian kernel. The modified Gaussian kernels, G_4 and G_6, may actually generate smoother interfaces in the presence of surface roughness, as in Figure 1.4. Here, G_n is a shorthand for

$$G_n(x;\eta) = \frac{1}{N_n}\exp(-x^n/\eta) \tag{1.9}$$

where N_n normalizes the distribution,

$$N_n^{-1} = \int_{-\infty}^{\infty} \exp -x^n/\eta \tag{1.10}$$

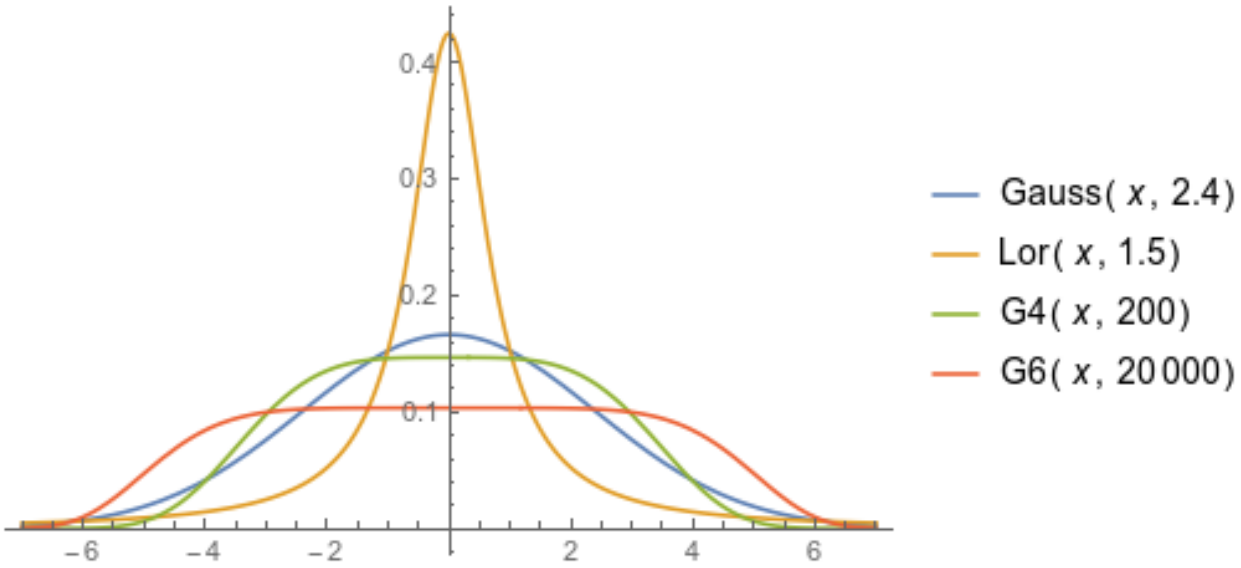

Figure 1.3: Comparison of four coarsening functions: a Gaussian, a Lorentzian, and two modifications to the exponent of a gaussian-like exponential. The legend indicates kernel and η. The kernels are all similar in width, but the rate of decay is different.

The Willard-Chandler choice of Gaussian function, for example, would be represented $G_2(x; \eta = 2(\xi_{WC})^2)$.

The wider kernels, however, are non-standard. Using them may introduce a set of unexpected changes, including differences in density profile, surface structures, and other measurable details. There is no clear justification for choosing a different kernel.

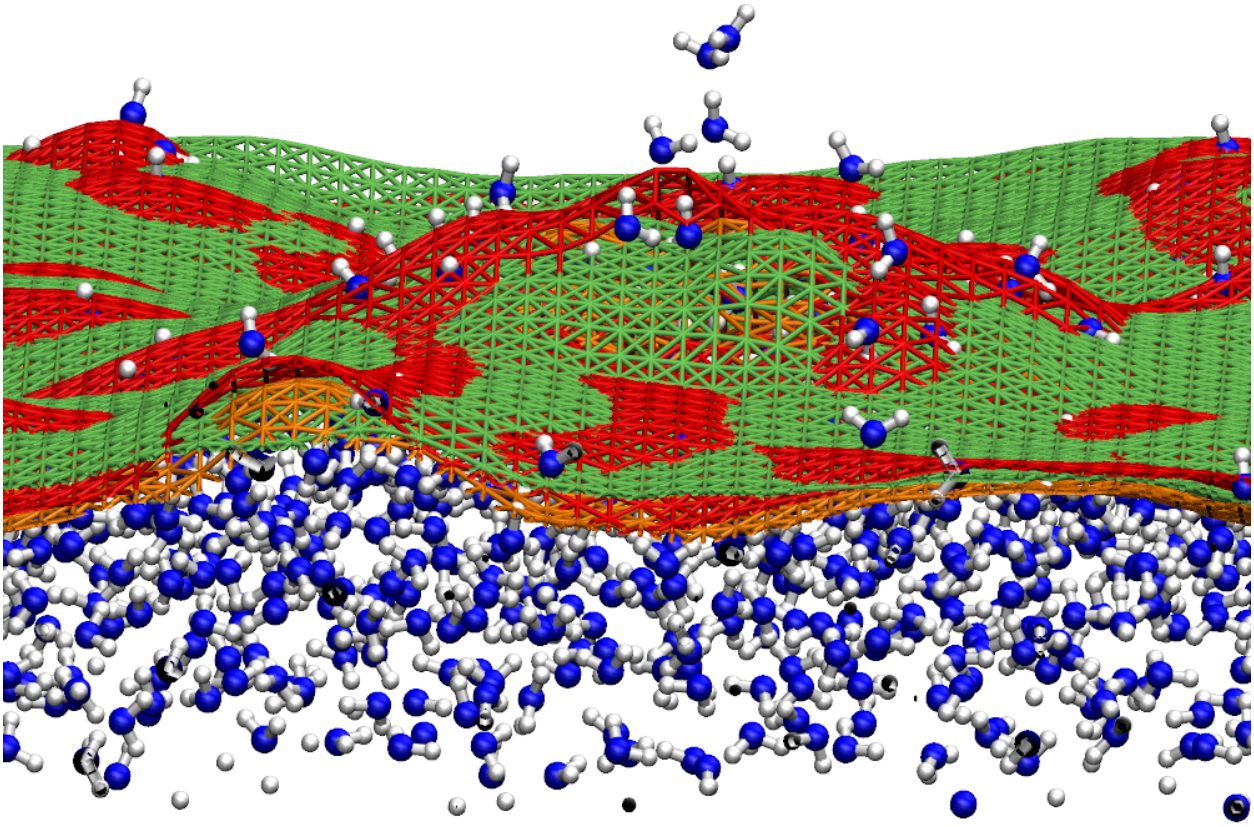

Figure 1.4: Comparison of interfaces generated with three parameter sets: WC (Red), $G_6(\eta = 2e4)$ (Orange), and $G_6(\eta = 1e5)$ (Green). Interfaces shown for SPC/E water (oxygen-blue, hydrogen-white) exposed to an electric field, $\mathcal{E}_n = 0.1$ V/Å, which produces an abnormally rough interface that the WC parameters may not capture effectively.

Because the total density in the system does not change with a normalized kernel, the density at any point in the bulk region of a liquid will roughly match the average density $\bar{\rho}$, and the density profile for any given cross section using the kernelized scheme will be

similar in shape and scale to an ensemble average density cross section like in Figure 1.1. A different choice of ρ^* will effectively shift the interface by a uniform amount. Extreme changes will emphasize or wash out structural features. Figure 1.5 sets the interfacial cutoff to $\rho^* = 0.6\bar{\rho}$, which amounts to a small shift in the depth coordinate and a slight change in surface roughness. More extreme choices of ρ^* would lead to more extreme washout of features.

Small variations in coarse-graining length ξ have a washout and depth-shifting effect similar to variations in ρ^*. A large ξ will mostly stretch the density profile across a wider range of depths. A very small value for ξ will lead to occasional bubbles in bulk as normal fluctuations in density appear as cavities. An infinitesimal ξ is effectively equivalent to treating molecules as point-densities, as they were without before the density coarsening.

Section 1.5 examines the effect of changing these parameters in a model fluid.

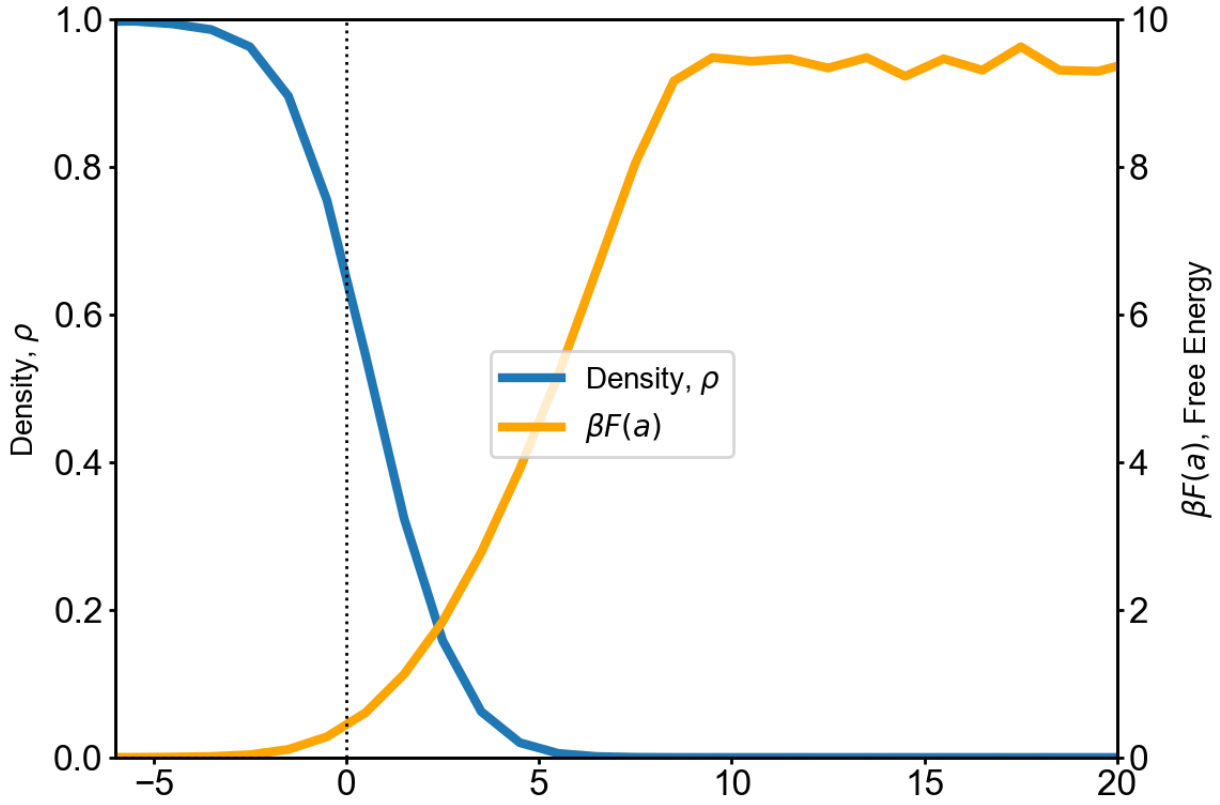

Figure 1.5: Density profile of SPC/E with interface set at $\rho^* = 0.6\bar{\rho}$. Free energy profile of a single water inferred using $\beta F = -\ln(\rho) + C$.

1.4 The implications

As a rigorous method for generating an interface that can be made arbitrarily smooth in space and time, the Willard-Chandler instantaneous interface offers access to microscopic details that were previously inaccessible–which will be the focus of Chapter 2.

Other methods for generating an instantaneous interface have been proposed. In particular, the ITIM (Identification of Truly Interfacial Molecules) method[6] has been studied[7, 8], but many surface properties are similar between the two methods[9].

1.5 Other fluids

The instantaneous interface can be applied to fluids other than water, although parameters must be adjusted accordingly.

The Dumbbell Fluid interface

To represent a dipole fluid, Cox and Geissler[10] proposed a simple dipolar dumbbell-shaped (DB) fluid composed of molecules of two equal and opposite point charges separated by a small distance and protected by Lennard-Jones parameters matching SPC/E water. The strength of the molecular dipole can be easily adjusted by modifying the charge and distance between the two theoretical atoms.

Figure 1.6 offers rough examples of a density profile for the dumbbell fluid with a Gaussian density kernel and a coarse-graining length of $\xi_{DB} = 4.0$, in LJ units. Four values of ρ^* are shown, from $\rho^* = 0.1$ to $\rho^* = 0.3$ according to the units shown on the y-axis, and demonstrate generally what substantial differences in density cutoff will do to a density profile. Figure 1.7 visualizes the same set of ρ^* as shown in Figure 1.6 and illustrates how changing the density cutoff shifts both the depth and roughness of the interface.

The dumbbell fluid forms a very sparse interfacial region that is particularly challenging to capture using a continuous interface, which is apparent from the broad density profiles in Figure 1.6.

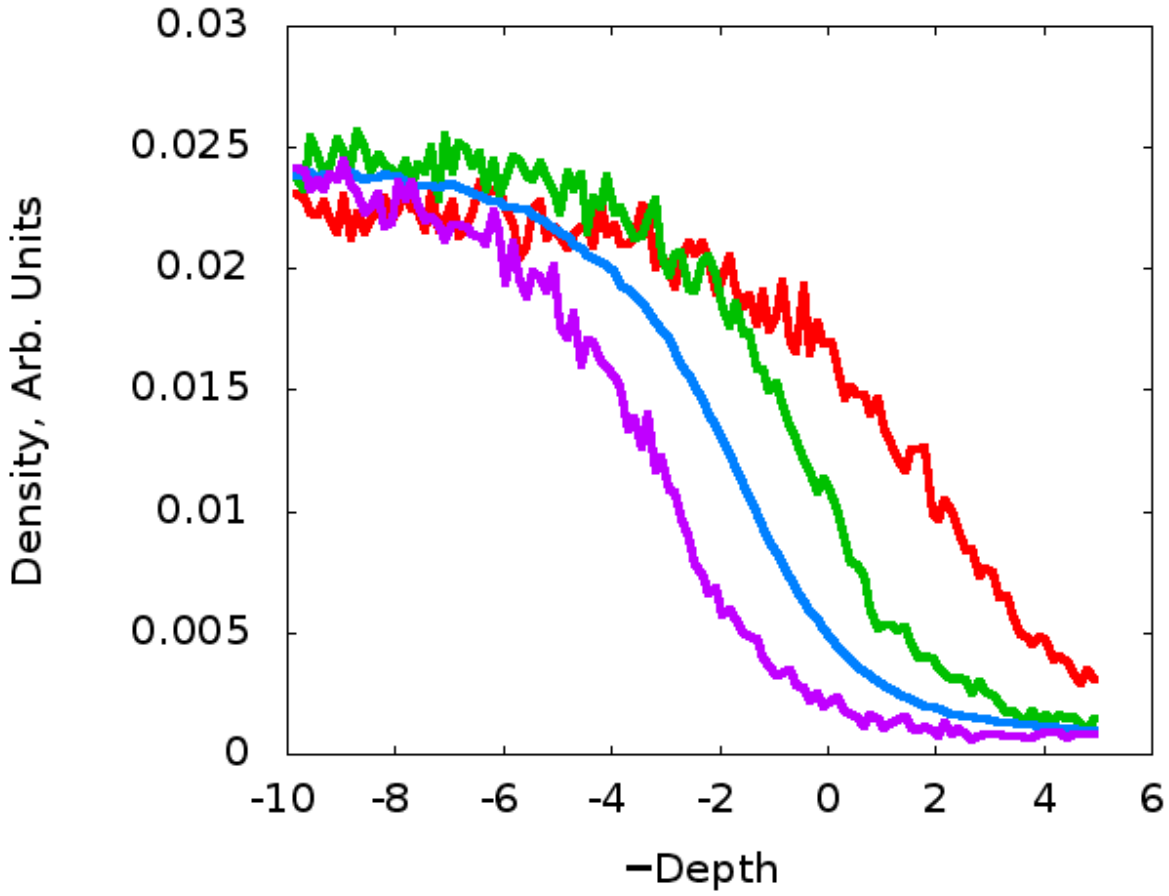

Figure 1.6: Four density profiles generated from the instantaneous interface over a liquid-vapor coexistence of the dumbbell fluid. Coarse-graining length $\xi_{DB} = 4.0$ and density cutoffs $\rho^* = 0.03$(red), 0.02(green), 0.015(blue), and 0.01(purple). The blue line was computed with extended data sampling to demonstrate that no features are hidden in the noise. Distance is measured in the DB model units, roughly equivalent to 1 Å.

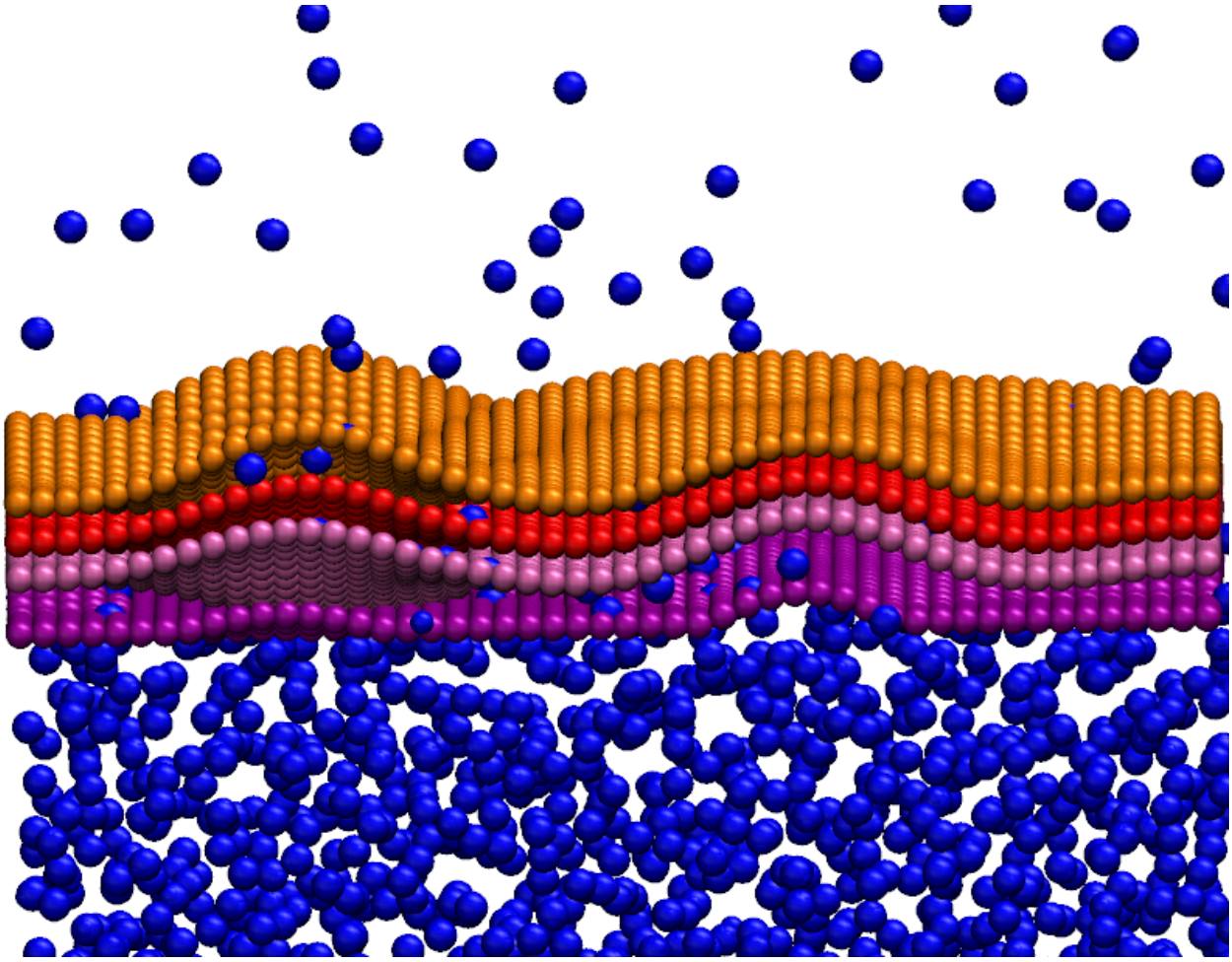

Figure 1.7: Examples of instantaneous interfaces of the DB fluid with various cutoff density ρ^* to match the profiles in Figure 1.6. Smaller density cutoff leads to a more distant and often flatter interface.

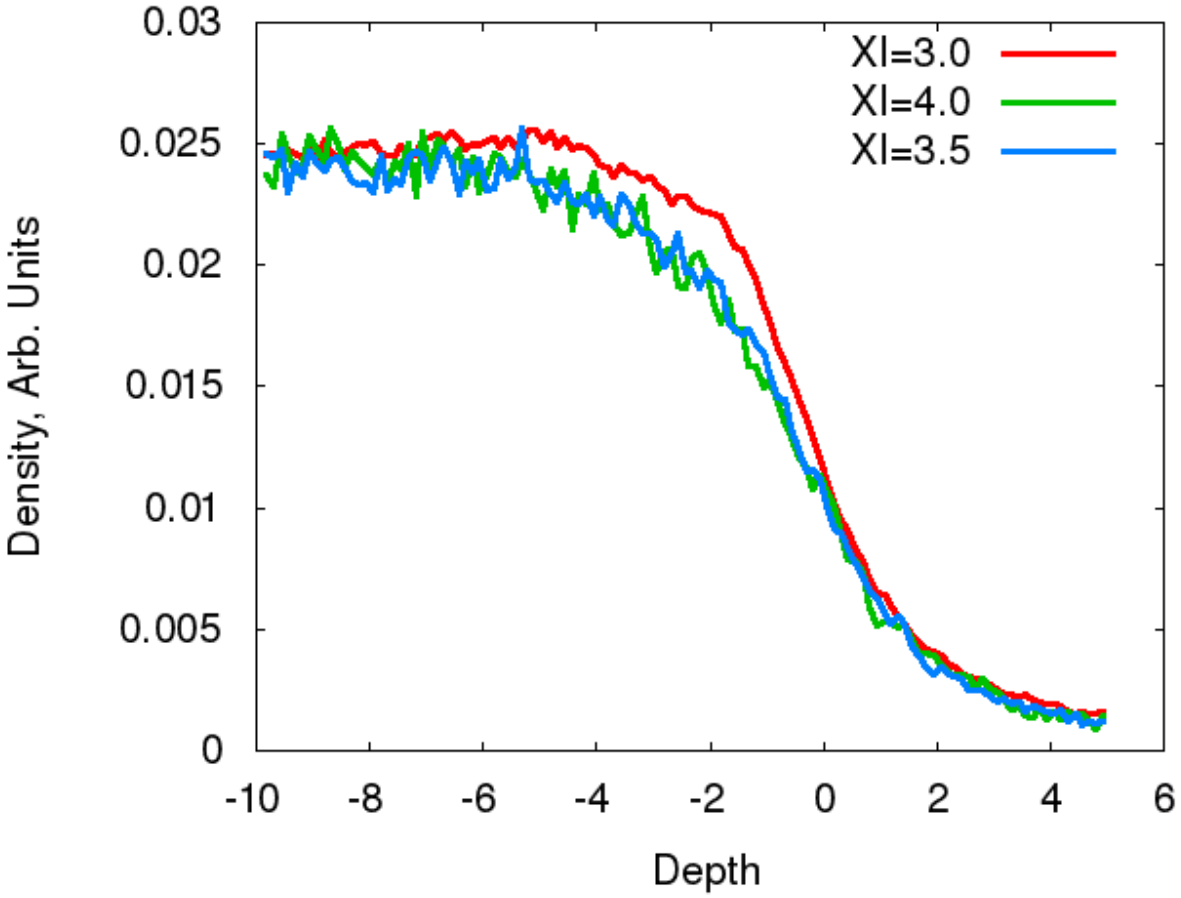

Figure 1.8: Interfacial density profiles for the DB fluid with constant $\rho^* = 0.020$ and with ξ_{DB} as indicated in the legend.

Figure 1.8 and its companion Figure 1.9 illustrate how changes to coarse-graining length ξ can be perilous. Although the density profiles in Figure 1.8 do not appear problematic, the visualizations in Figure 1.9 show that the smaller ξ leads to mistakes classifying droplets in vapor, sharp changes in depth, and bubbles.

The dumbbell fluid does not show strong density layering near the interface, contrary to tetrahedral water models like SPC/E and TIP5P. The unique interfacial patterns in water, and the potential causes of those patterns, will be examined more closely in subsequent chapters.

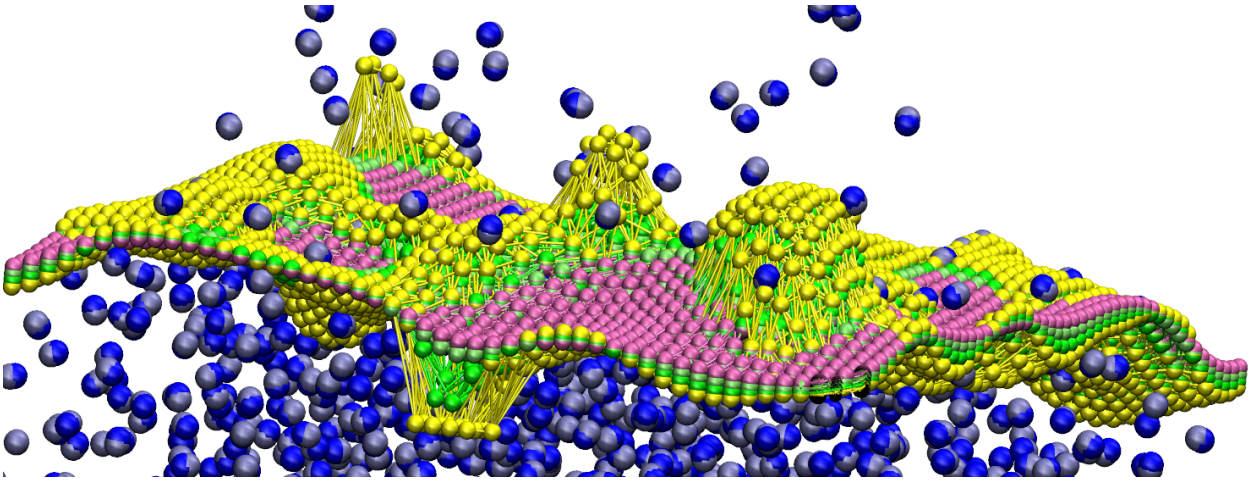

Figure 1.9: Snapshot of DB fluid interfaces generated with $\xi_{DB} = 2.5$(yellow), 3.0(lime), 3.5(green), and 4.0(pink). Smaller ξ_{DB} leads to droplets, abrupt jumps, and bubbles in the interface calculation.

The Mercedes-Benz Water interface

The Mercedes-Benz model (MB water), a water-like liquid in two dimensions, highlights two shortcomings with defining an instantaneous interface: liquid sparsity and surface roughness. Sometimes called BN2D as a 2D water model first proposed by Ben Naim in 1971[11], this model can be roughly understood as a projection of the 3-dimensional hydrogen-bonding behavior of water into two dimensions[12]. This model is explained in more depth in Chapter 3, mainly Section 2.3.

MB water forms extended hexagonal domains that resemble the hexagonal (basal) plane of frozen water. In liquid conditions, these domains are present but transiently reforming, often passing through rings of five to potentially dozens of members. A 6-membered ring is large enough to accommodate an interstitial water particle in the center of the hexagon, but this particle has no hydrogen bonds. See Figure 1.10 for a representative snapshot of a MB water trajectory at liquid conditions.

The algorithm for computing a WC instantaneous interface does not change significantly for MB water except that a 2-dimensional Gaussian kernel is used. Moving into two dimensions immediately poses one challenge to the instantaneous interface algorithm: that of sparseness. In 3-dimensions, a shorter coarse-graining length ξ can be chosen, meaning more detail is preserved. A tagged water molecule is likely to have around 3.5 hydrogen bond partners[13] and a few other neighbors close enough to generate a smooth, liquid-like density. MB water, on the other hand, will on average have fewer than 2.5 hydrogen bonds [12] and, by extension, the neighbors will also have fewer bonding partners than when compared to 3-dimensional liquid water. The lower dimensionality leads to sparseness, which in turn requires a relatively longer coarse-graining length.

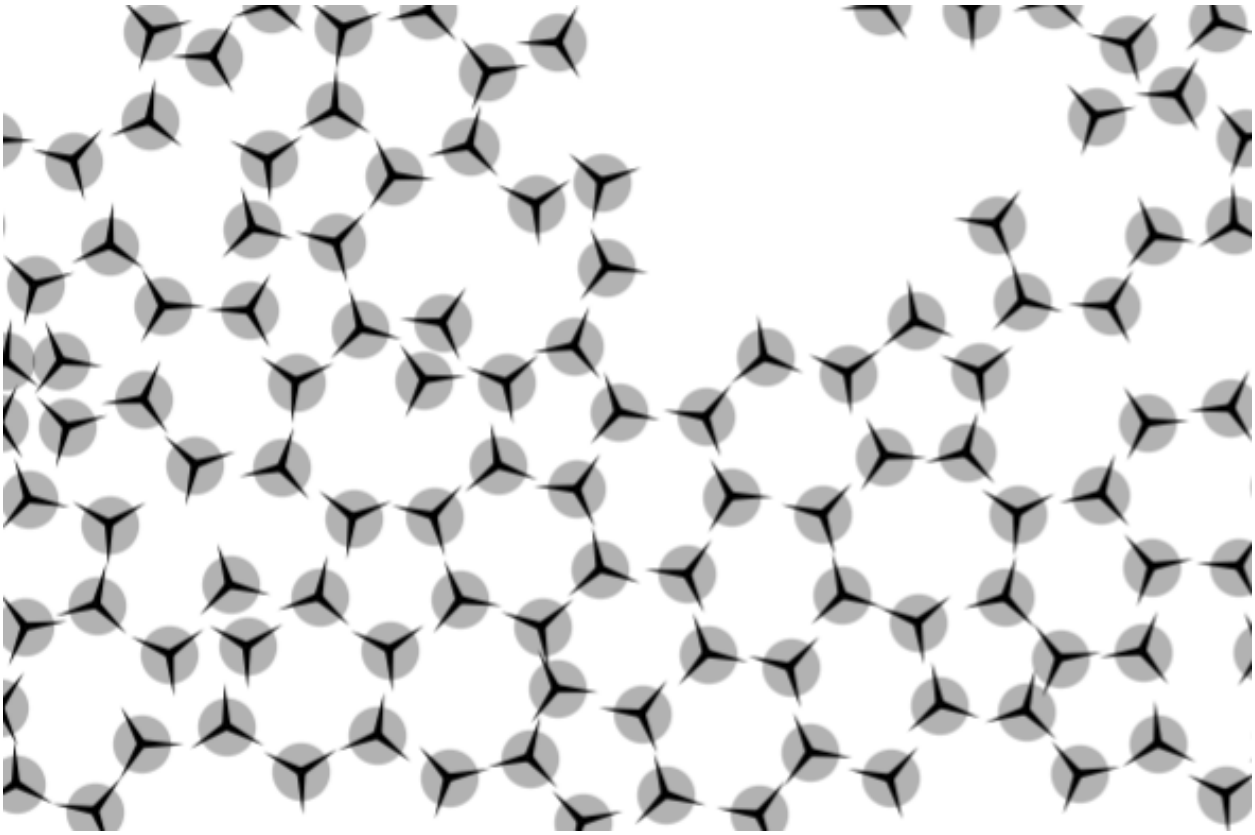

Figure 1.10: Example snapshot of the Mercedes-Benz model of water. Each particle has three spokes corresponding to the three directions of potential hydrogen bonding. The Lennard-Jones exclusion is roughly the size of the gray circle and is less than the ideal hydrogen bond length. Ideal hexagonal domains are widespread, but are transient in time and often interrupted by rings of smaller or larger size, or by interstitial particles that might not hydrogen bond.

A rough surface is, predictably, also challenging for the WC algorithm, which seeks to compute a smooth surface. Sudden or extreme changes in density, perhaps by a string of loosely hydrogen bonded particles that have nearly evaporated, can make the density gradient non-monotonic as a function of the surface-normal coordinate. In other words, following a straight line from bulk liquid to vapor there may be more than one altitude that matches an otherwise reasonable cutoff density value. There is no physical requirement that a liquid phase form a continuous, convex interface. However, much of the analytical benefit of the instantaneous interface comes from its smoothness. A disjointed interface confuses the surface curvature and molecular depth. The separation between liquid phase and vapor phase becomes difficult to define and therefore difficult to measure. For liquids that are strongly structured geometrically and yet have low surface tension, like the MB water model, interfaces are more likely to be rough and therefore become problematic to

analyze.

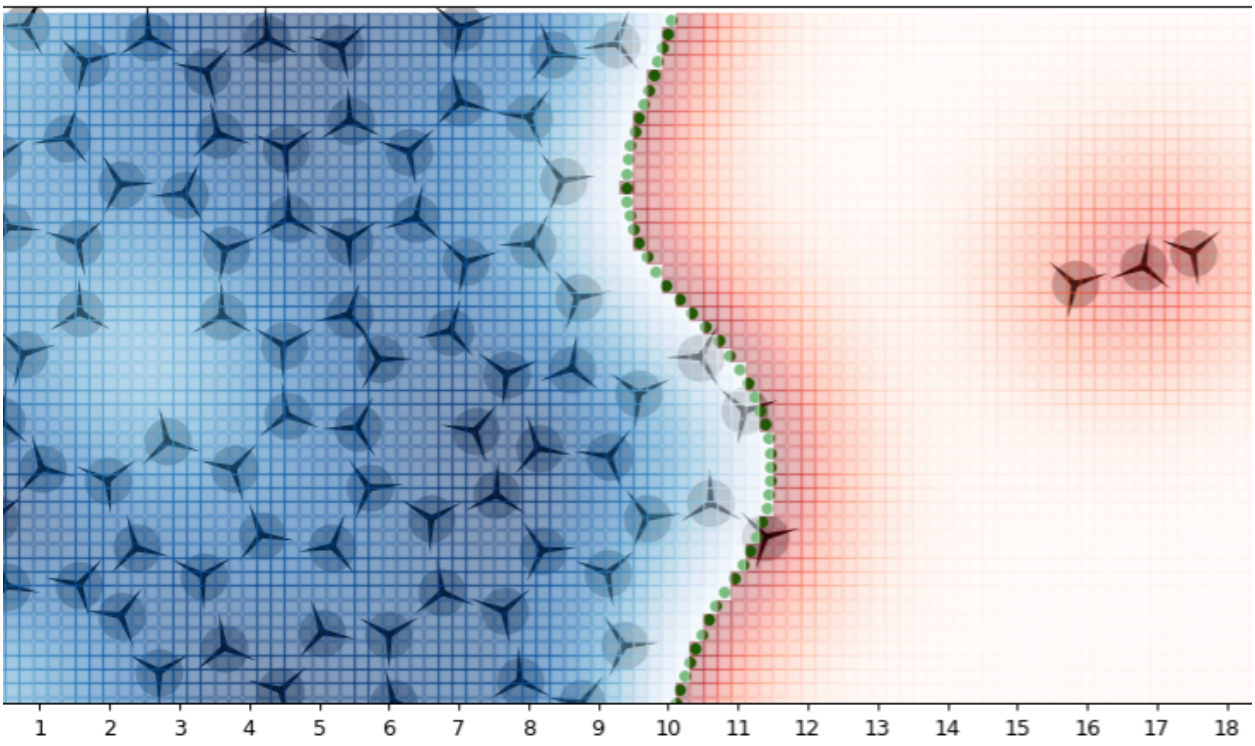

Figure 1.11: Density map and interface of a snapshot of MB water. MB water particles are represented as in 1.10. Green circles demarcate a Willard-Chandler style instantaneous interface. Blue squares correspond to the coarse-grained density $\rho_{WC}(\vec{r}, t)$ that is greater than the cutoff ρ^*, while red squares indicate a density below the critical cutoff. For both red and blue, a darker hue indicates a higher density value. The densities here were generated using a 2-dimensional Gaussian kernel with coarse-graining length $\xi_{MBW} = 1.3\ r_{HB}$ and cutoff density $\rho^*_{MBW} = 0.35\ r_{HB}^{-2}$, where r_{HB} is the reference length of a hydrogen bond in the MB water model.

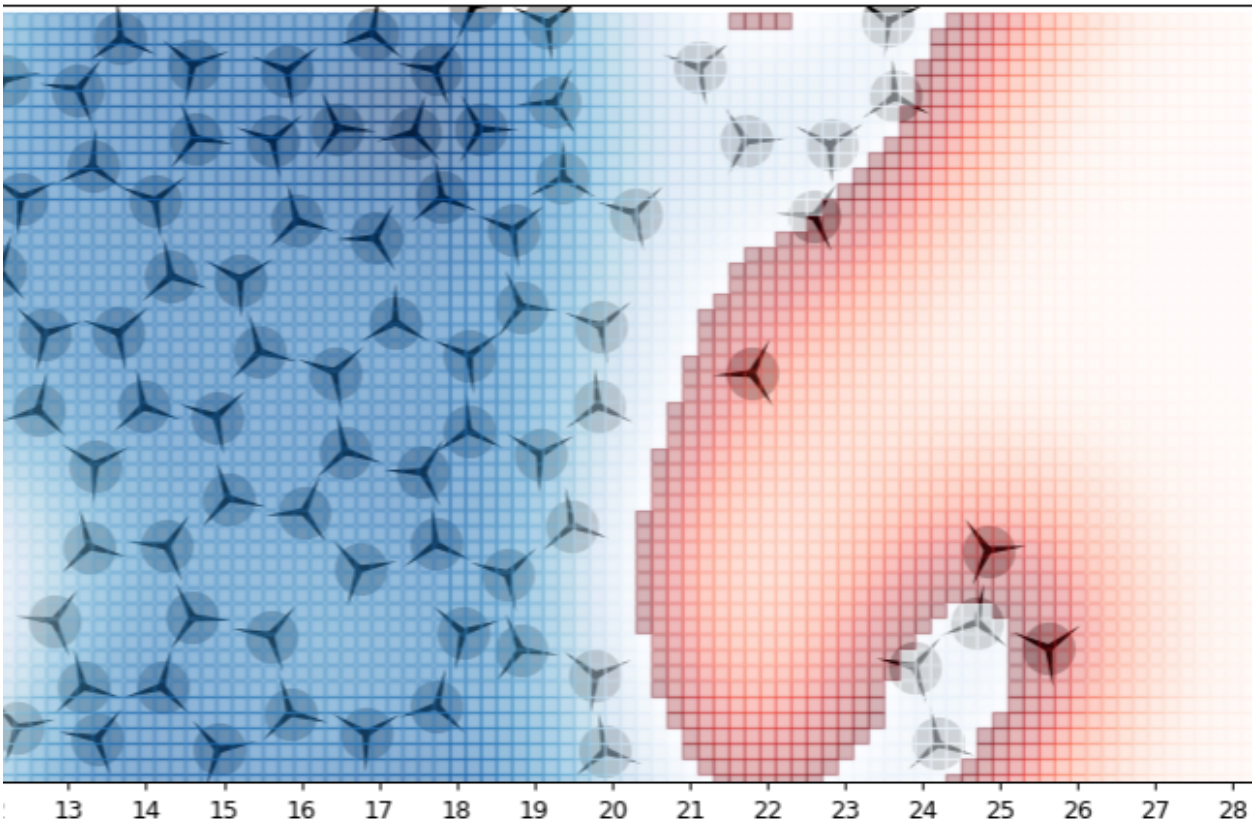

Figure 1.12: An example of one particularly problematic snapshot where a loosely-held string of MB particles extends partway into the vapor phase. A straight line from the bulk liquid to vapor will not necessarily have one clear point of delineation between phases. This sort of geometry is common to the MB water model. The roughness makes it difficult to gather statistics with reference to a single characteristic depth, which is one of the primary uses of an instantaneous interface.

1.6 Summary

The study of interfaces in molecular simulation is made easier through the use of the WC instantaneous interface, which offers arbitrary spatial and temporal resolution. Generating an interface by computing density with a kernel is robust, flexible, and can be extended to most liquids, especially those with a high surface tension. The remainder of this document will investigate some of the surface-related phenomena that become accessible computationally because of this interfacial scheme.

Chapter 2

The Surface of Water

Overview

All liquids make weak, short-lived, non-covalent inter-molecular bonds that are quickly broken and just as quickly reformed. Water as a liquid forms a tetrahedral network of hydrogen bonds, each of which persists for only a few picoseconds[14]. Despite the short bonding lifetime, each water molecule has 3 to 4 hydrogen bonds on average, in a mostly tetrahedral arrangement. The 3-dimensional nature of water's hydrogen bonding network is strongly disrupted by a flat, 2-dimensional interface.

Many water theories rely on treating water as an infinite "bulk" liquid, either with no edges or with no dependence on the water boundary. This is especially true in molecular dynamics simulations of liquid water, which predominantly use periodic boundary conditions to approximate bulk liquid. These assumptions are useful and the associated models are usually still accurate, but the interfaces that water forms do occasionally matter. Since all water inevitably forms interfaces–with solutes, with ice or vapor, with another liquid, or with a solid surface–it is essential to also build an understanding of water's response to a surface.

Of all the interfaces that water can form, the flat liquid-vapor boundary is one of the simplest. A finite droplet of water will form a sphere with a curved surface, but as the droplet grows its surface becomes nearly flat at the molecular length scale. While a flat liquid surface does not constitute a constraint on water, the strongly geometric hydrogen bonding network rearranges to minimize energy loss from the inability to form any favorable bonds outside the liquid. Hydrogen bonds lost at the liquid's edge are the main contribution to water's high surface tension and the preservation of near-surface hydrogen bonding is the main reason for short-lived structure at water's surface.

With a detailed depth profile, other properties of water's surface can be measured with better resolution, including density, orientational patterns, and hydrogen bonding information. Much about the surface patterns of liquid water can be accounted for by relating the water surface to the basal plane of ice.

2.1 The depth matters

How deep is the "surface" of water? What properties are sensitive to the presence of a surface, and how do these properties need to be modified? Many properties (dielectric constant, rates of motion, macroscopic thermodynamic properties) correlate strongly with the density of water, which is not constant at water's surface. Density correlations dissipate after only about 10 Å, so most liquid properties appear "bulk-like" after only a few layers of water molecules.

Figure 1.1 is a density profile that averages through two spatial dimensions, x and y, and in time, t. The mean interface is defined as a flat line that, on average, delineates liquid and vapor at one-half of liquid density. Water molecules are constantly fluctuating above and below that line, which limits the resolution of this interface. As discussed in Sections 1.2 and 1.3, using a simple flat line to represent the interface will not capture detailed density fluctuations in space and time.

Depth is a useful measurement that is especially sensitive to the definition of the interface. Water molecules respond to an imposed interface by realigning to the new geometric requirement. Each layer of water molecules will see a slightly different environment, and so each layer will have slightly different behavior. A highly averaged flat-line interface will be too broadly defined to distinguish one molecular layer from another. An instantaneous interface, on the other hand, is better able to categorize waters by defining a finer depth.

Depth from a flat, averaged interface is the perpendicular distance between the center of a water molecule and the nearest point on the interfacial line. In simulation this usually means that a slab of liquid is initialized with the liquid vapor interface set perpendicular to the lab-frame $\hat{z}$. A liquid with high enough surface tension will remain stable over simulation timescales and so the interface can be assumed to always span the xy-plane. If these conditions are met, then the interface is a plane composed of the set of points (x, y, z_{int}) where the average density over an arbitrary time t and space x, y is half of water density,

$$\langle \rho \rangle_{x,y,t} (z_{int}) = \rho_{bulk}/2 \tag{2.1}$$

and the altitude of a water molecule at $\boldsymbol{r} = (x, y, z)$ is

$$a(\boldsymbol{r}) = z - z_{int} \tag{2.2}$$

with depth

$$d(\boldsymbol{r}) = -a(\boldsymbol{r}). \tag{2.3}$$

The instantaneous interface does not form a flat plane along xy and will almost always have considerable curvature. The depth of a water molecule at $\boldsymbol{r}$ is defined in this case as the distance between the water center and the nearest interfacial point $\boldsymbol{s}$, projected onto the surface normal at $\boldsymbol{s}$, called $\hat{n}_s$,

$$d(\boldsymbol{r}) = (\boldsymbol{s} - \boldsymbol{r}) \cdot \hat{n}_s. \tag{2.4}$$

Computationally, this is easiest to do by first initializing a slab along x and y as before. Every coordinate on a grid along x and y will have at least one (or two if including the bottom interface in a periodic system) values of z that match the critical density within a numerical tolerance. The surface normal is found by computing the vectors between adjacent interfacial points in x and y and crossing them

$$v_x = s_{x+1} - s_{x-1} \tag{2.5}$$
$$v_y = s_{y+1} - s_{y-1} \tag{2.6}$$
$$\hat{n}_s = v_x \times v_y. \tag{2.7}$$

When depths are assigned using Equation 2.4, water molecules at similar depths exhibit similar, depth-dependent behavior. In Figure 2.1, the histogram of depths from an instantaneous interface shows layering, which is an important aspect of water's interface.

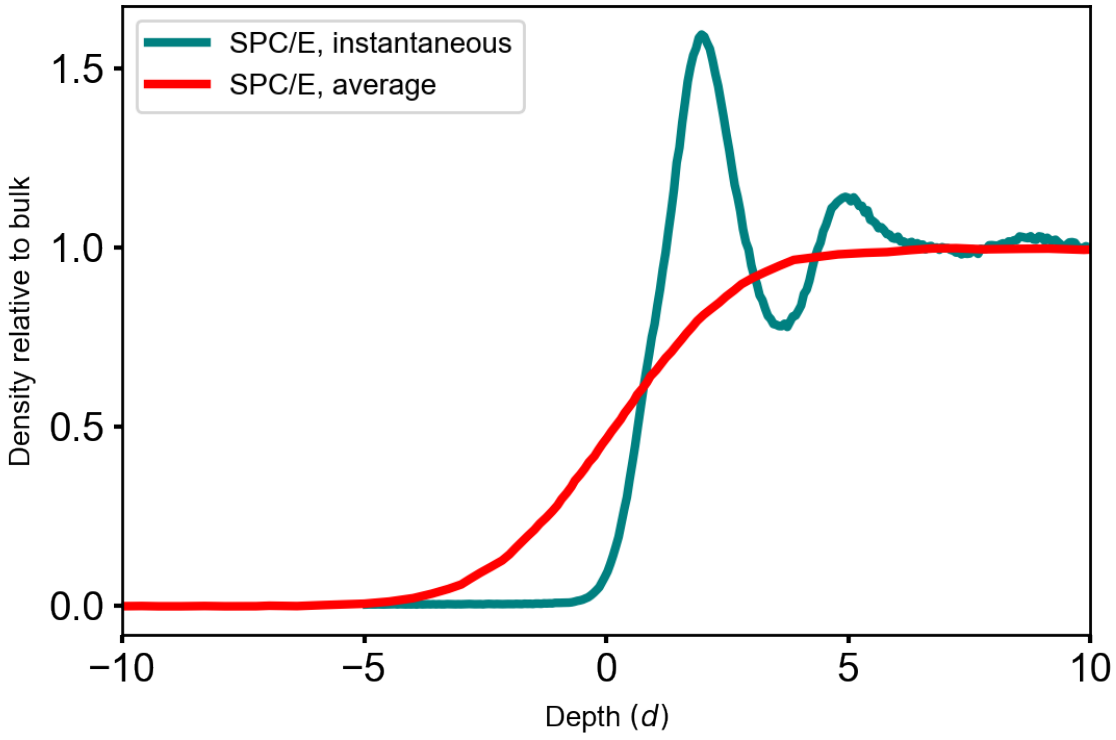

Figure 2.1: Relative density of water molecules according to depth for SPC/E, when depth is computed using either the lab-frame distance from a flat, average dividing surface, or the instantaneous interface. The instantaneous interface is better able to assign water molecules a depth that corresponds to a characteristic behavior.

2.2 Layering

The density profile of Figure 2.1 shows how water can potentially be divided into layers with depth. The series of maxima and minima are reminiscent of the shell-like structure of the radial distribution function of a liquid. Figure 2.2 compares the shells of the oxygen-oxygen radial distribution function $g_{OO}(r)$ of liquid SPC/E with the layers of density measured at the interface. The surface is linear while the radial distribution function measures over spherical shells, but the layering is similar in the two cases and is illustrative of the shape and scale of liquid water-water correlations

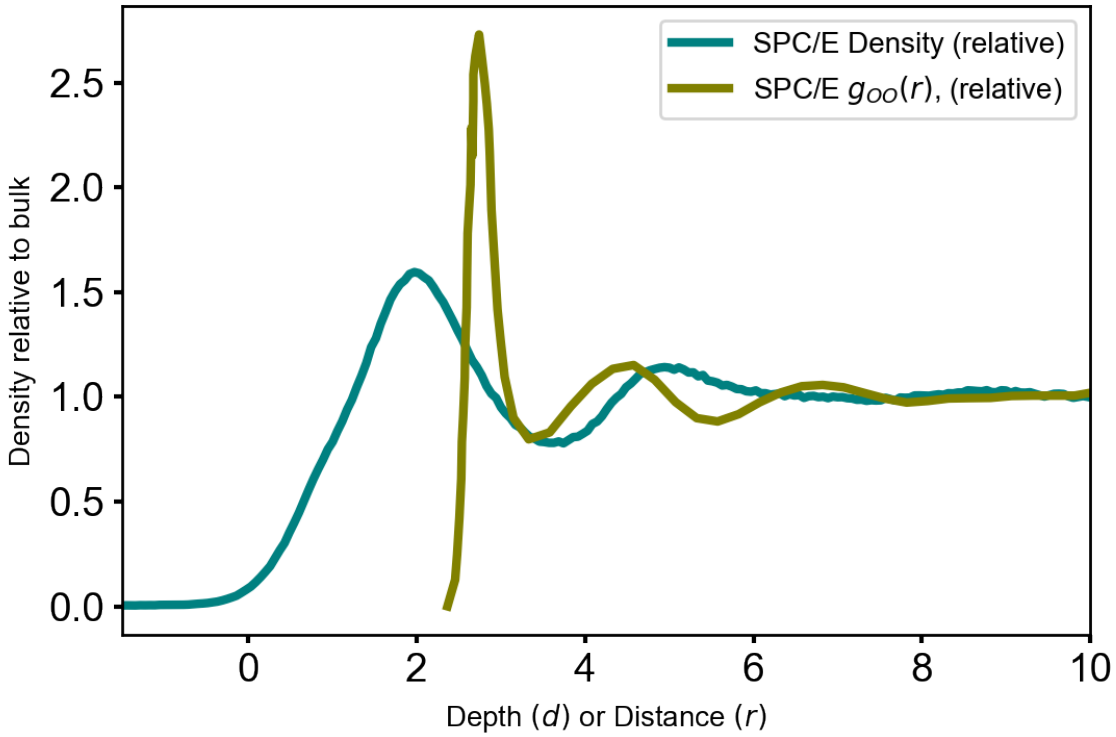

Figure 2.2: Density of liquid SPC/E according to depth from the instantaneous interface compared to the oxygen-oxygen radial distribution function from[15]. The shells of the radial distribution function are conceptually similar to the layers of the density profile.

More relevant than the radial distribution function is the layering similarity to ice, Figure 2.3. The basal plane of ice has a characteristic bilayer structure, Figure 2.17, that closely matches the layering of the surface of liquid water. It will be shown that the surface structure of water loosely follows the crystal structure of the ice basal plane.

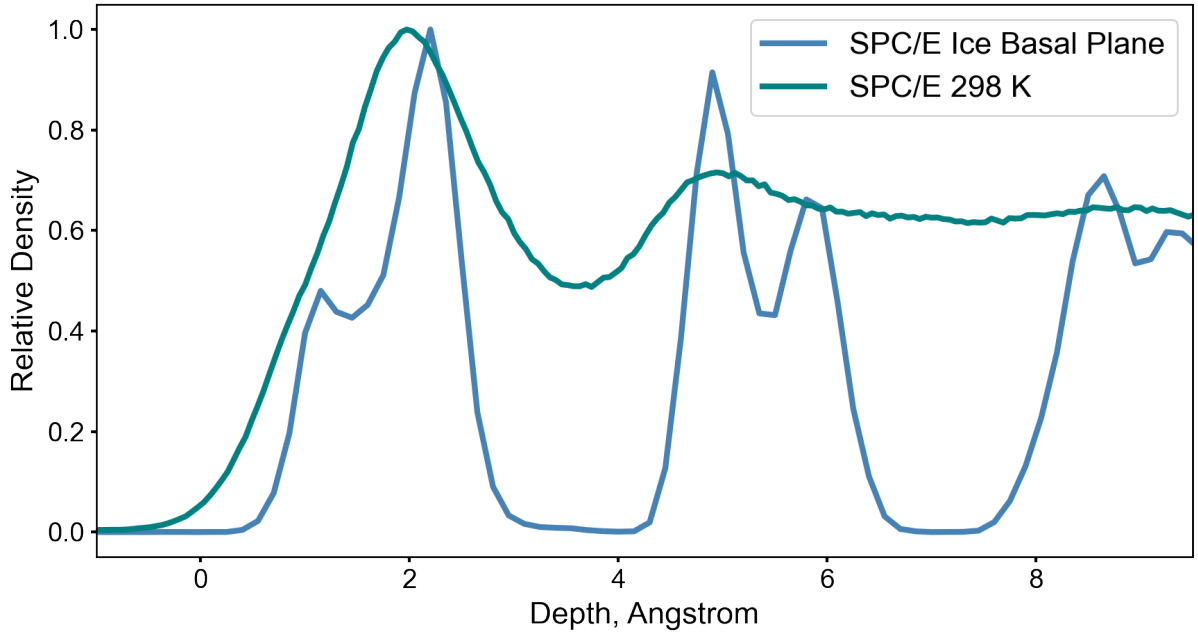

Figure 2.3: The surface of water compared at the surface and at depth of the basal plane of SPC/E ice. The layering of the liquid surface closely matches the bilayer pattern of the basal plane of ice.

At lower temperatures, most liquids should also exhibit layering as transient, solid-like behavior is characteristic of the cohesive forces that stabilize a liquid interface. Increasing the temperature of a liquid will continue to broaden the peaks in surface density until eventually the peaks flatten into a sigmoid, or a phase change occurs. A liquid with high surface tension will generally form a sharper, less diffuse interface. Any change to the surface tension will also lead to a change in surface density profile.

Categorizing water's surface

Kessler et al. [16] proposed dividing water's surface into distinct layers, termed L0, L1, L2, and L3, as in Figure 2.4. Each layer is 3 Å wide, which roughly corresponds to the diameter of one water molecule. The separation of vapor-like L0 and the sharp surface L1 at a depth of 0.5 Å is an arbitrary value that approximately separates layers at each minima. Nonetheless, each layer has characteristic patterns.

Although not stated in their work, the differentiation of characteristic layers by Kessler et al.[16], and the extension to a "2D H-Bond Network Skin" and "$L1_{\parallel}$" of ref[17], is really a measurement of the liquid quasi-bilayers that evolve from the basal plane of ice. In liquid water, the two bilayer peaks blur and merge into a single density layer. Shen and Ostro-

verkhov[18] speculated from spectroscopic measurements that the surface of water resembles the basal plane of ice, and that there is a smooth transition between an ice-like and a liquid-like surface as temperature increases. The connection to ice suggests that $3.5 - 3.6$ Å may be a better separation distance between layers.

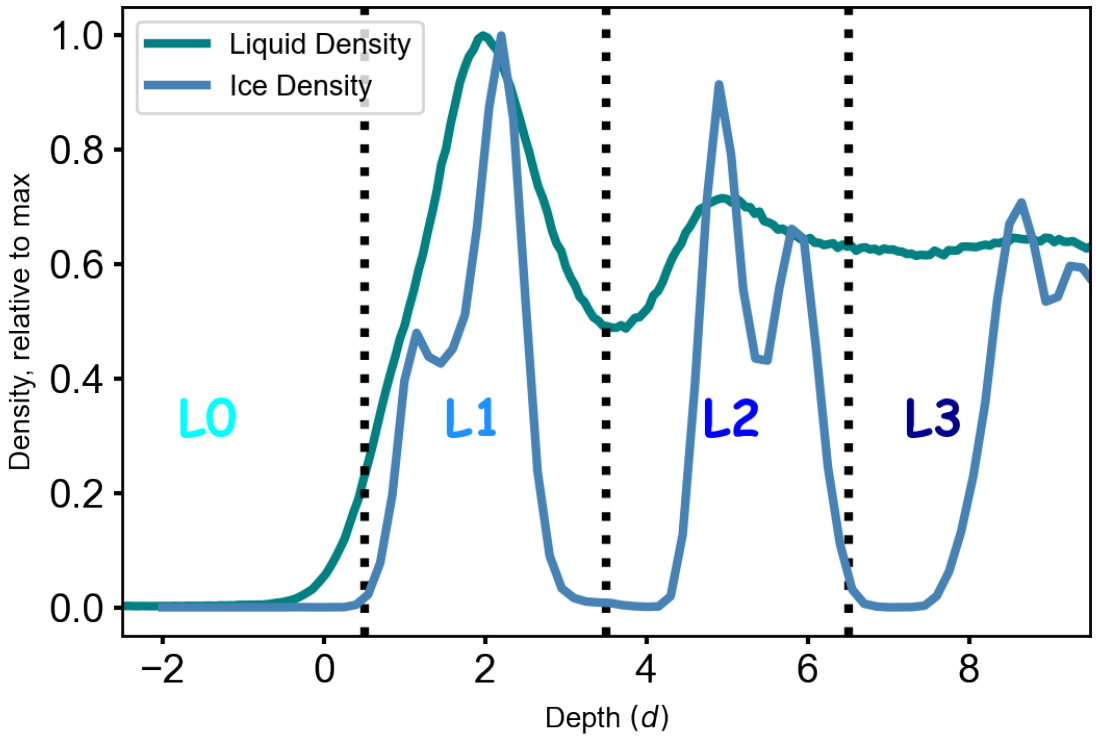

Figure 2.4: Classification of water surface layers, as proposed in ref [16]. Each layer of water has distinct, measurable behavior. The density profile of the ice basal plane explains why each layer has distinct character.

Refs [16] and [17] report on distinctions between each layer.

L0 is a mixture of low-density vapor and the outer extremes of the water surface. Due to low density, L0 has around one-tenth the number of water molecules as L1. These molecules are geometrically unable to form a third or fourth hydrogen bond because the density in this region is low, and in general L0 waters either donate one bond and accept one bond [16]. These molecules contribute strongly to the "dangling water" that produces a strong spectroscopic signal through the second harmonic generation technique and was first measured in 1988[19].

L1 is the outermost layer of nearly fully bonded water molecules, with only about 0.5 fewer hydrogen bonds on average than a bulk water molecule[16]. These molecules show

strong orientational[5], bonding[16], and network[17] patterns that are distinct from bulk. A sub-region in the top half of L1 is composed of water molecules which tend to point one or both OH-vectors parallel to the surface, and will be referenced throughout this work as either the parallel region or $L1_{\parallel}$. The $L1_{\parallel}$ region captures the upper half of the top bilayer of the basal plane of ice, the shoulder peak at $d = 1$.

L2 and L3 can mostly be understood as reverberations originating from the strong water structure in L1 in response to the sharp interface. Alternatively, L2 and L3 can be understood as less constrained liquid quasi-bilayers of ice. The density, and with it most other surface-sensitive properties, becomes indistinguishable from bulk water after about 10 Å.

While the details about each layer will change with choice of water model and slight differences can be found between classical and DFT simulations[16, 17], the general patterns are strong enough to be measured in every water model. This is likely because the layering is fundamentally related to the geometry of ice.

Orientations in each layer

When Willard and Chandler[5] first reported the layered interfacial density, they also observed a preference for orientation that changed with depth. Many orientational patterns were already known, but in less detail[20].

Because the position of water with respect to the interface is 3-dimensional, some cuts and averages must be made to digest the data. Unless noted, measurements are of SPC/E at STP. The Figures in this section will focus on the angle formed between the surface normal and one of two characteristic vectors within the water molecule. These are either the vector connecting the center oxygen and one of the bonded hydrogen atoms, or the dipole vector which points from the center oxygen and passes exactly in between the two hydrogen atoms. Alone, one of these vectors is not enough to uniquely describe the orientation of a water molecule, but either choice is sensitive enough to detect patterns at the surface. Because it is reasonably descriptive, this section will mostly refer to the angle formed between the surface normal and either of the oxygen-hydrogen separation vectors, which are indistinguishable but can form a joint probability distribution.

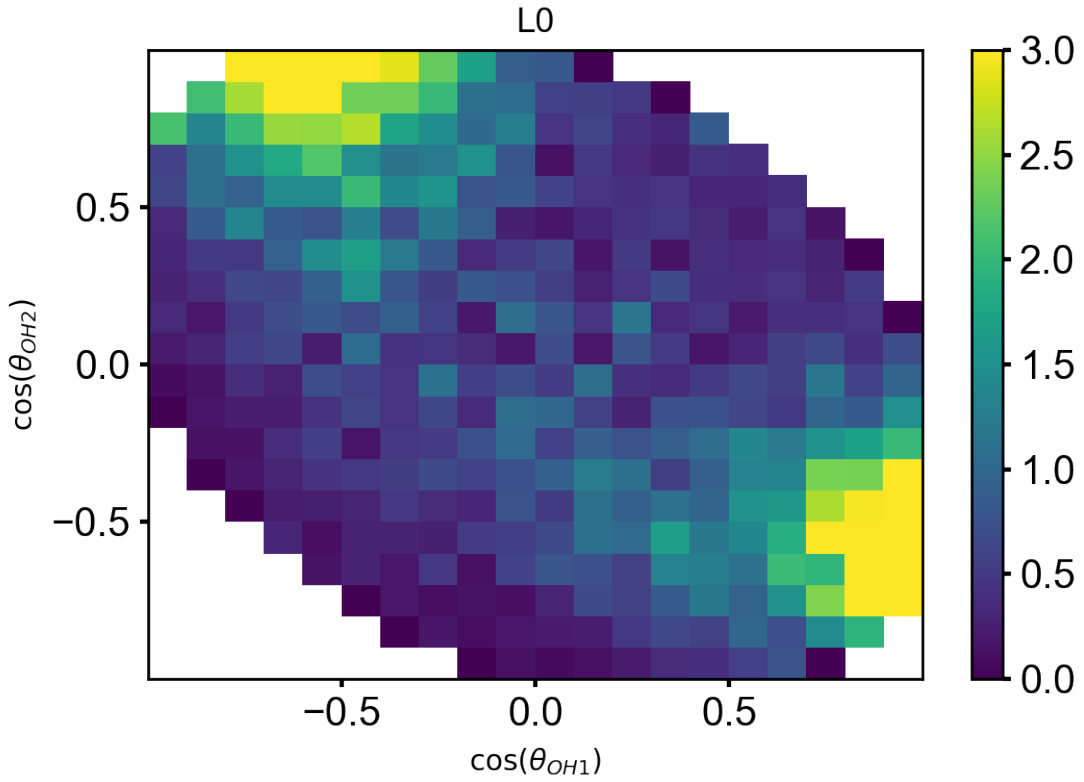

Figure 2.5: Probability of the angle made between the surface normal and either OH-vector in the region L0, relative to the probability in bulk. The OH-vectors show a strong tendency to point along the surface normal toward vapor. This type of plot is symmetric along the diagonal, but low sampling will appear as asymmetries.

Noted earlier, the L0 region exhibits a strong and characteristics "dangling water" feature, which derives from the surface-most water molecule tendency to orient one hydrogen directly out of the surface. These waters, on average, sacrifice one hydrogen bond to vapor, but as a result are able to fully satisfy two or three hydrogen bonds that connect downward toward bulk. See Figures 2.5 and 2.10.

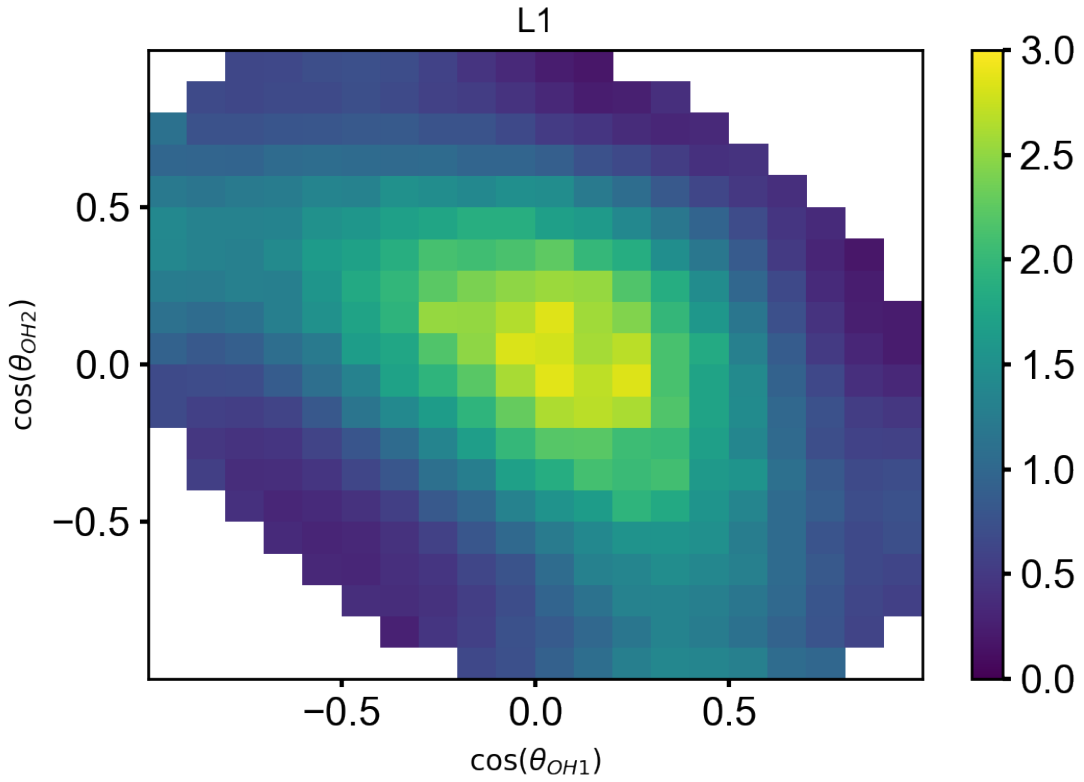

Figure 2.6: As in Figure 2.5, for waters in the L1 region. Waters in L1 region are more likely to orient with both OH-vectors parallel to the surface (perpendicular to the surface normal). L1 can be broken into two characteristic regions.

The L1 region is composed of two sub-regions, one where both OH-vectors point parallel to the surface, and one where at least one OH-vector points toward bulk. Figure 2.6 can be evenly split at depth $d = 2.0$ to produce Figures 2.7 and 2.8. The strongly parallel pattern is clear in the $L1_{\parallel}$ region, which is the half of L1 nearest to vapor. Although weaker, the trends in orientation within L0 and L1 are apparent when measured using electronic structure calculations[16].

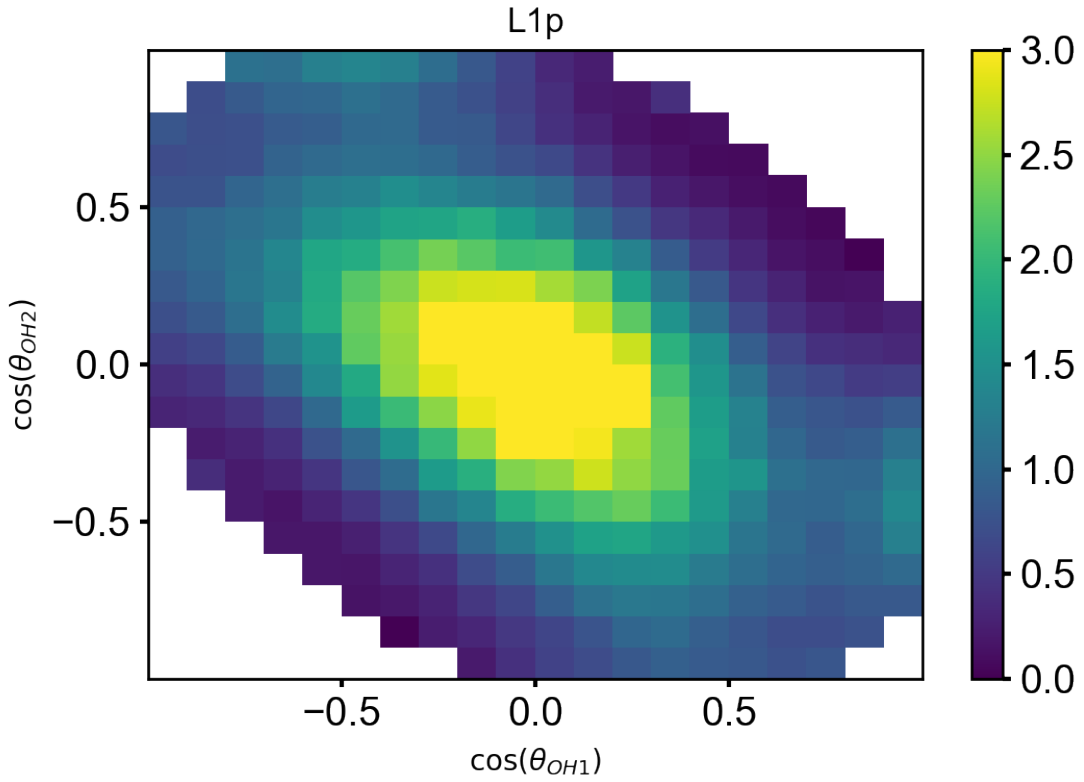

Figure 2.7: As in Figure 2.5, for waters in the $L1_{\parallel}$ region (depth 0.5 to 2). The first layer of water molecules are very likely to point parallel to the surface, i.e. $\cos(\theta_{OH}) \approx 0$. The "lone pairs", which are not explicitly included in SPC/E simulation, would therefore tend to point slightly upward and slightly downward, as a tetrahedron.

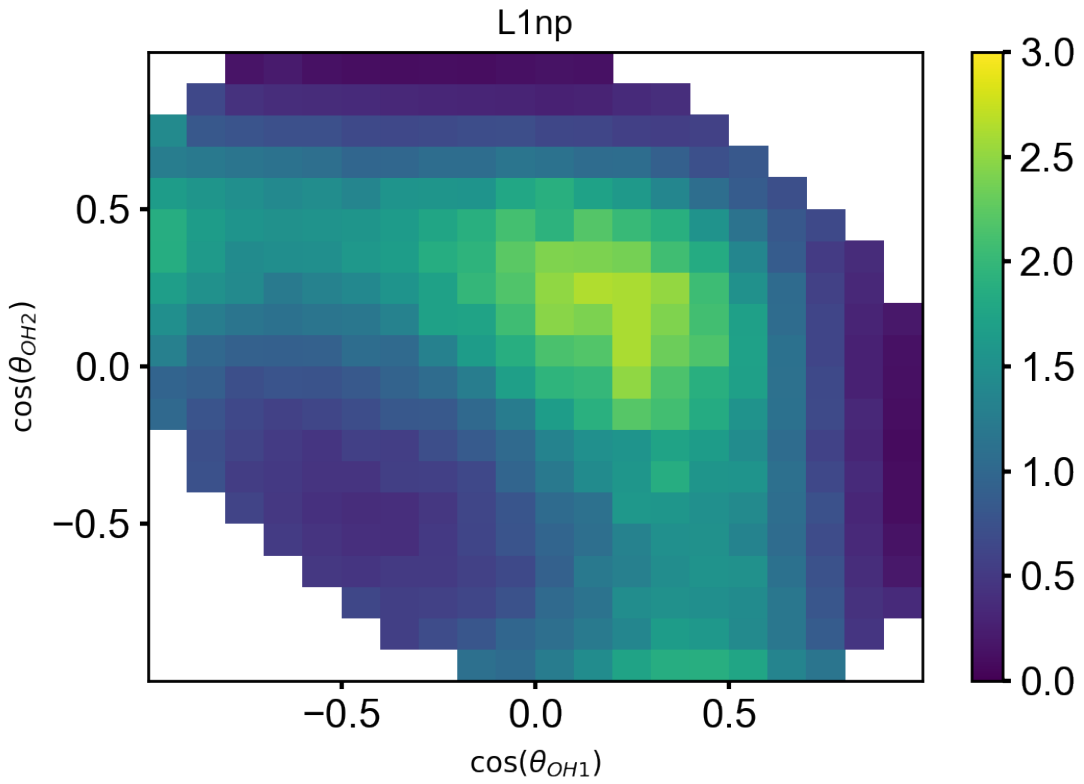

Figure 2.8: As in Figure 2.5, for waters in the lower half of the L1 region (depth 2 to 3.5). These waters are more likely to point one hydrogen downward, with the second hydrogen pointing slightly out of the interface, i.e. $\cos(\theta_{OH}) > 0$.

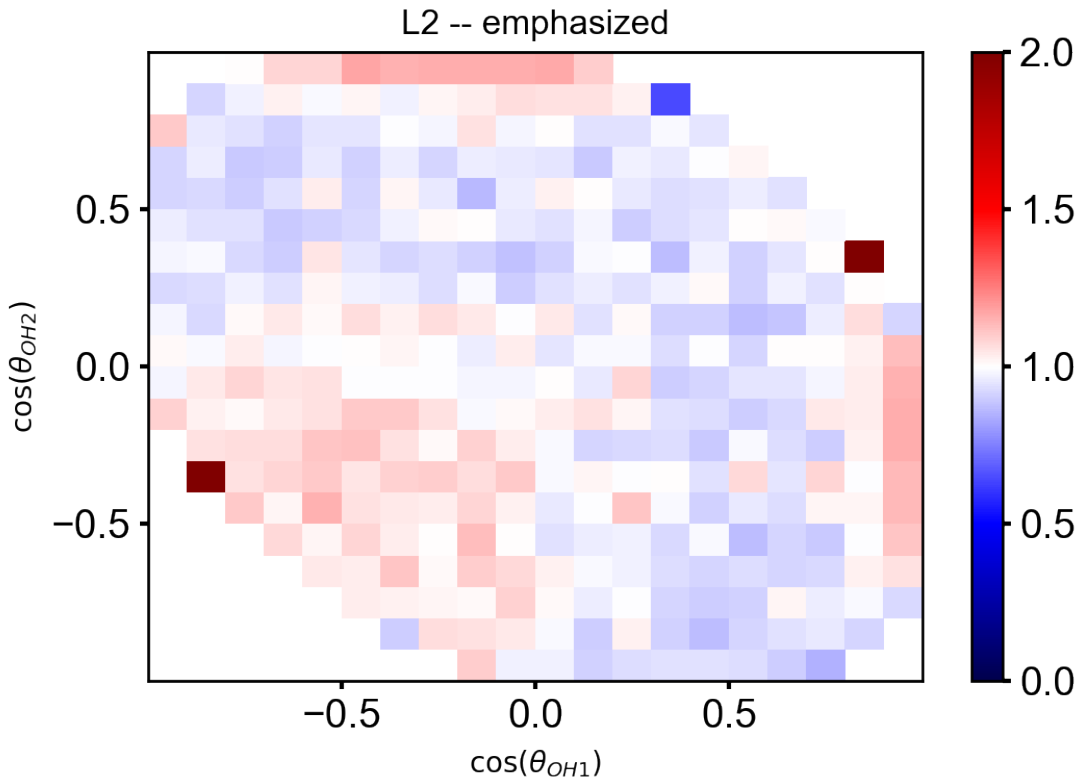

Figure 2.9: As in Figure 2.5, but with colors rescaled to emphasize the otherwise flat data in L2. Waters in this region have a small tendency to point one hydrogen toward vapor, likely in response to the downward-facing lone pairs in L1.

Although the density in L2 still has not settled to bulk values, the patterns in orientation rapidly disappear. Whether any signal is apparent in L2 depends on the choice of water model. SPC/E, which only loosely enforces tetrahedrality, does not show strong patterns beyond L1. A faint pattern can still be measured in L2, which is mostly an echo of L1, see Figure 2.9.

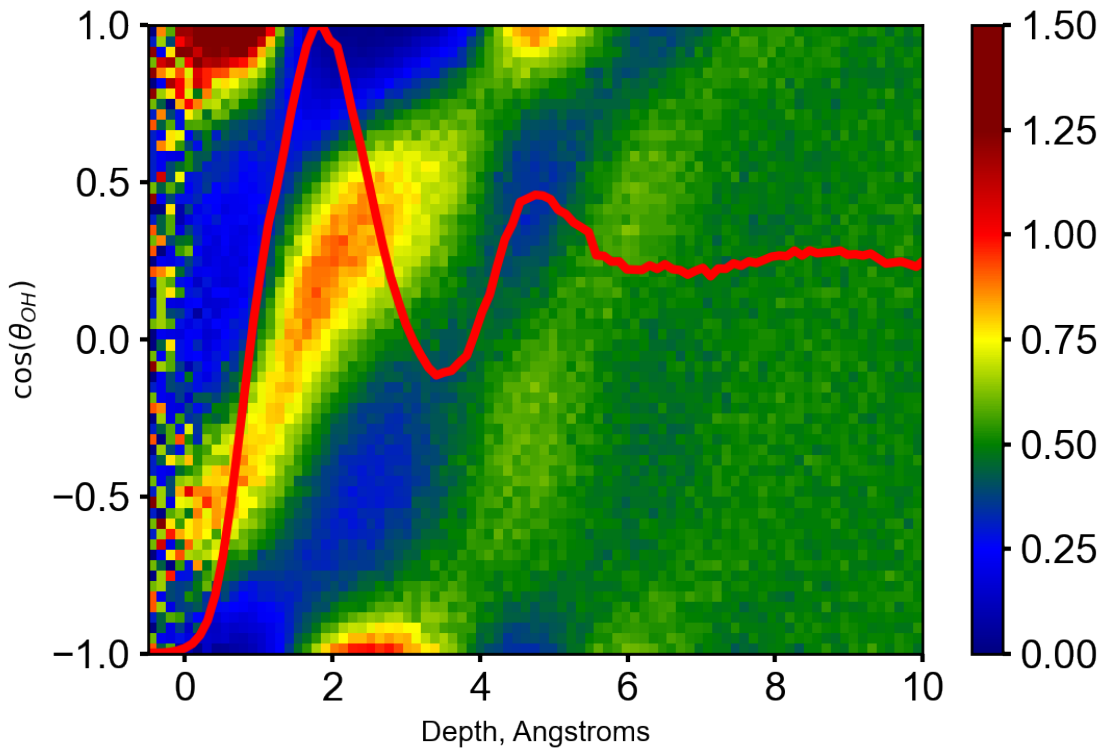

Figure 2.10: Distribution of water OH-vector angle with surface normal as a function of depth, $P(\cos(\theta_{OH})|d)$. The surface density profile $\rho(d)$ is superimposed in red with arbitrary units to indicate the connection between layering and orientation. Data is for SPC/E; other water models have similar trends, but with recognizable differences. The color map describes the probability density of a given θ_{OH}, which is normalized separately for each depth, $\int d\cos(\theta) P(\cos(\theta)|d) = 1$.

Figure 2.10 makes a single distribution of $\cos(\theta_{OH})$, combining both of the OH-vectors, to produce a heat map of angular probability at each depth. Near $d = 0$, there is a very strong tendency for one OH-vector to point directly upward, $\cos(\theta_{OH}) = 1$, and a corresponding probability at $\cos(\theta_{OH}) \approx 0.5$, which is necessarily where the second OH-vector points, at about 109.5 degrees. The parallel layer near $d = 1$ is subtle in this figure but becomes more apparent with the density correction in Figure 2.11. Subsequent oscillations in orientation are apparent descending from the interface and attenuate to bulk at depths near 10 Å.

The form of Figure 2.10 is effective at summarizing orientational information and so will be used frequently in later Chapters. However, there are some shortcomings to this representation.

The $\cos(\theta_{OH})$ axis is not uniform in angle, but it is uniform in surface area of the sphere

that the OH-vector can explore. At a glance, the representation of Figure 2.10 emphasizes the behavior when the OH-vector points toward the poles and obscures the data near the equator.

Furthermore, because the probability is normalized by column, regions of low density appear as important as regions of higher density. The dangling hydrogen feature at depth $d = 0$ is actually a small, if strong, feature by number of total molecules. A much larger number of waters actually point parallel or downward in the L1 region, so these two orientations should be considered the dominant behavior of surface water. Figure 2.11 preserves the relative number of water molecules by keeping the sum of each column proportional to the density at that depth.

Finally, there is a rotation about the OH-vector that is averaged to produce these plots. If $\cos(\theta_{OH_1}) = 1$, the second OH-vector has a constant $\theta_{OH_2} = 109.5°$ and taking its average is no issue. However, if $\cos(\theta_{OH_1}) = 0$ then a large range of θ_{OH2} is averaged together, obscuring the actual orientation of the water molecule. This ambiguity can be reduced by plotting a joint $\cos(\theta_{OH1}) - \cos(\theta_{OH2})$ distribution, as above, or by considering a similar heat map with the dipole, $\cos(\theta_\mu)$ in Figure 2.12.

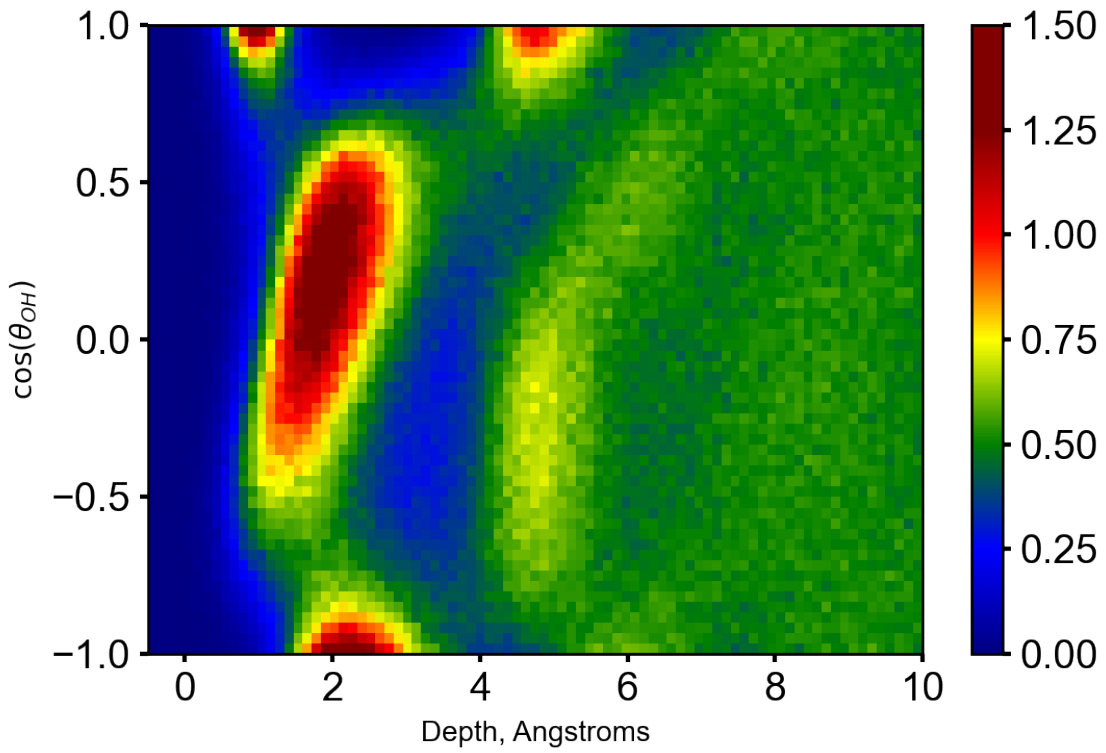

Figure 2.11: Measurements as in Figure 2.10, but with the sum of each column kept proportional to density. This presentation makes it more clear how common it is to find a water molecule at a particular depth and orientation. Waters are rare above $d = 0.5$, so almost no patterns are discernible in that region. The parallel layer is more apparent with this presentation, in the range $0.5 < d < 2$, but is still obscured because equatorial angles are underrepresented when referring to $\cos(\theta)$. The color map describes the product of probability density and molecular density of a given θ_{OH}, which is normalized separately for each depth, $\int d\cos(\theta) P(\cos(\theta)|d) = \rho(d)$.

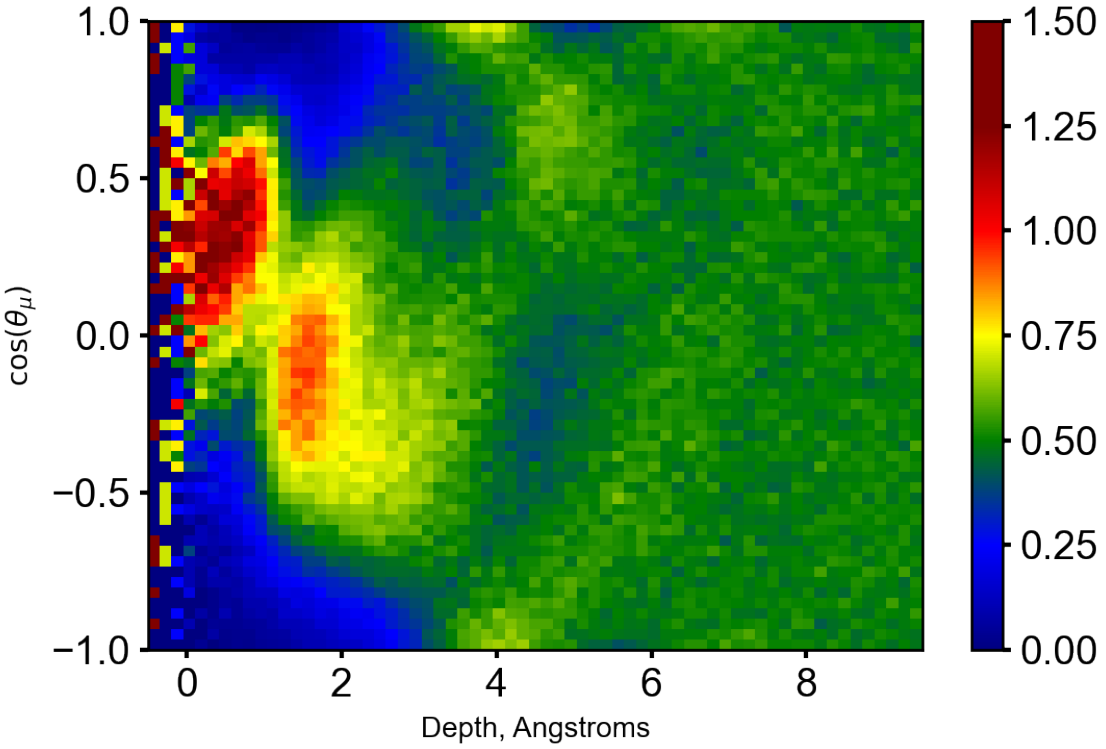

Figure 2.12: Heat map for the probability of an angle between the SPC/E dipole and the surface normal at each given depth. There is a tilt upward in L0 corresponding to the dangling hydrogen feature, then a slight tilt downward in L1.

The dipole of SPC/E also shows strong surface structure that seems directly related to the strong hydrogen bonding structure seen in Figure 2.10. The direction of surface dipoles in Figure 2.12 does not seem to contradict the direction of the OH bond vectors in Figure 2.10, and the two probability maps in combination provide a reasonably complete picture of water orientations near the water surface.

Strong geometric preferences for surface hydrogen bonding likely dictates the direction of the dipole, but it is interesting that the surface dipoles tend to point in parallel with the surface. It is possible that the average dipole creates a local field that attenuates water orientations, as has been widely suggested. Also, given that the dipole patterning largely disappears after about 5 Å, it is not unreasonable to think that the dielectric behavior of water (see Chapter 4) may be bulk-like at scales beyond 5 Å.

Comparison to TIP5P

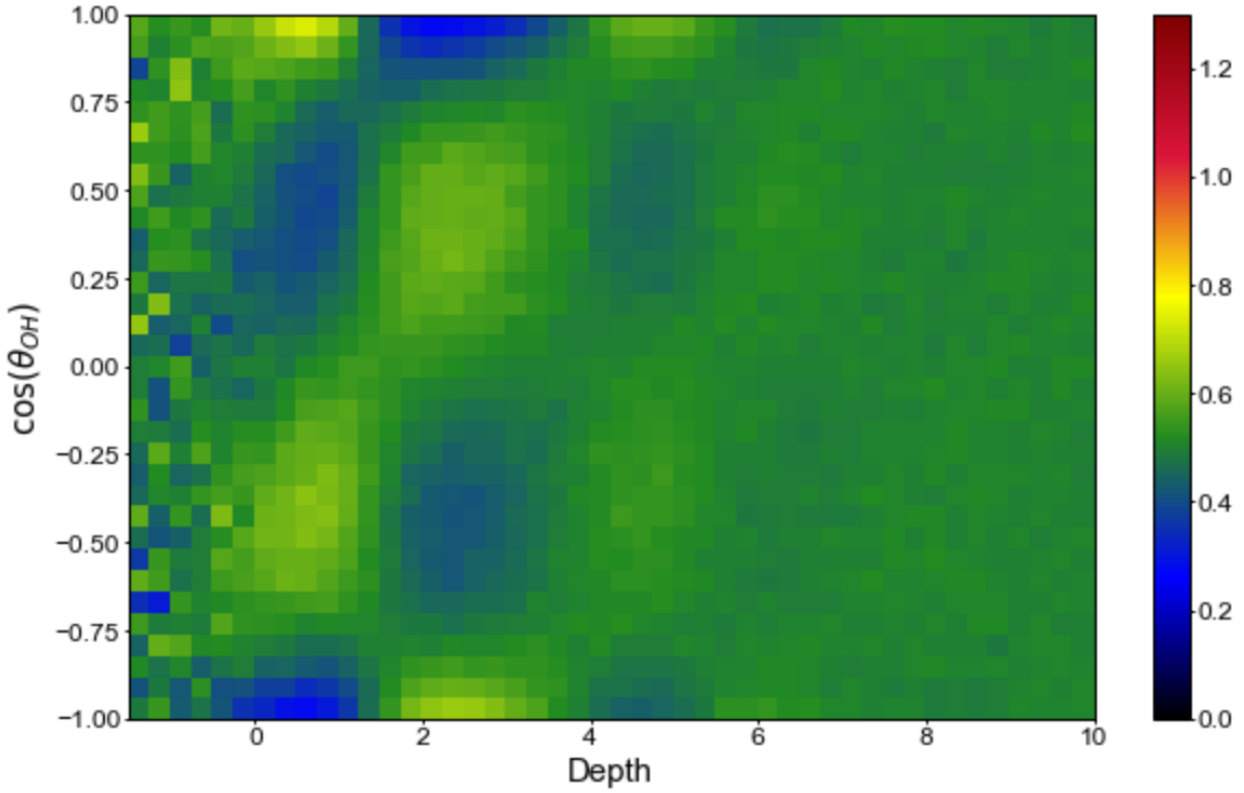

Figure 2.13: TIP5P OH-vector orientations heat map, compare Figure 2.10. Because TIP5P is more strongly tetrahedral, the angular patterns correspond more strongly to the ideal tetrahedral angles. In general, orientational patterning is weaker in TIP5P.

Different water models will show different angular preferences[16, 17, 21], but the strongest patterns are present in all water models. SPC/E does not explicitly enforce the two water electron lone-pairs, but because both hydrogen atoms in SPC/E form a nearly tetrahedral angle with the central oxygen, a tetrahedral network structure tends to form in SPC/E systems. As a result, when SPC/E is measured in the less-constrained interfacial region, a single hydrogen bond can be accepted evenly across a whole range of angles opposite from the hydrogen atoms. TIP5P, on the other hand, explicitly represents water lone-pairs with negative point charges. For this reason, TIP5P will exhibit strongly tetrahedral structure, even when near a sparse interface, see Figure 2.13. TIP5P, notably, does not show a strong parallel layer at the water-vapor boundary, Figure 2.14, while SPC/E and various DFT methods do[16]. The strong difference in signals suggests that the parallel layer is facilitated largely by the relaxed geometric requirements of a model like SPC/E.

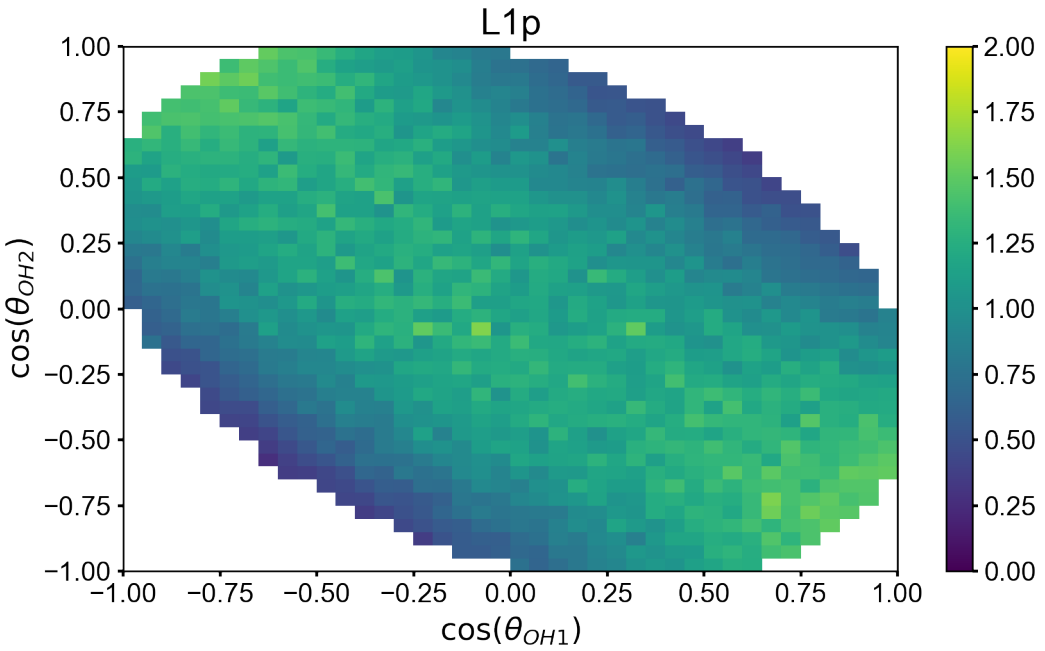

Figure 2.14: The parallel layer, $L1_{\parallel}$, is almost imperceptible in TIP5P (a change in scale is made from the previous SPC/E cos-cos heat maps).

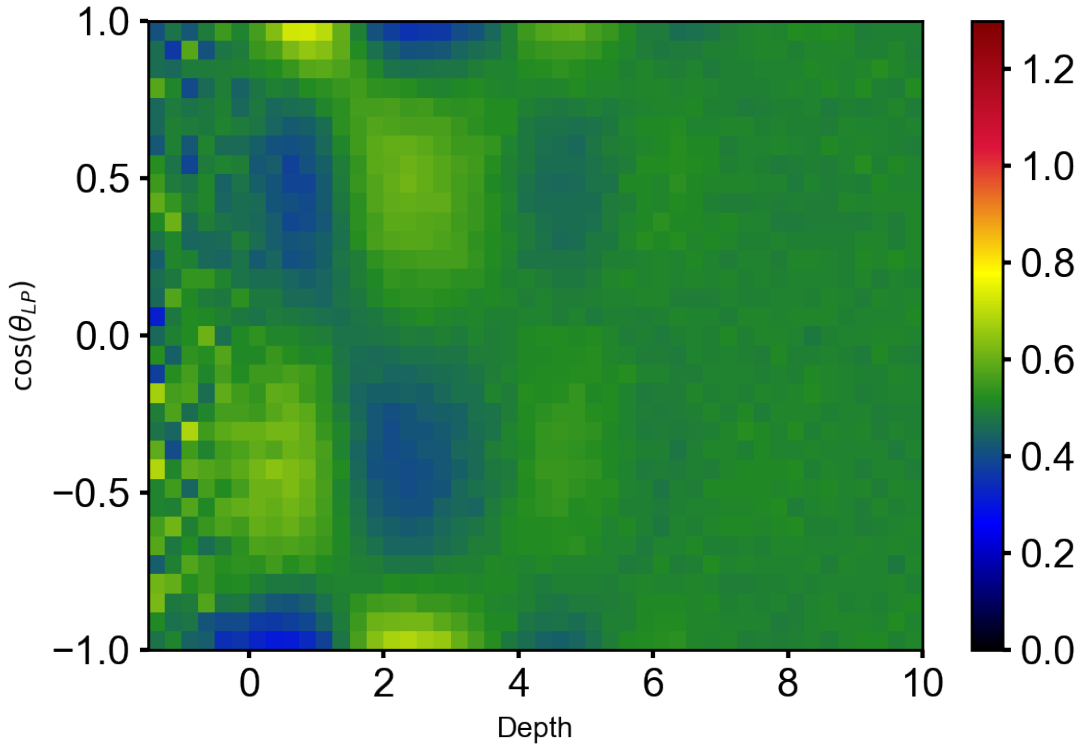

Figure 2.15: TIP5P lone-pair vector orientations heat map. The patterns in lone-pair angle are a close match to the patterns in the $\cos_{OH}$ distribution, which is expected for symmetric tetrahedral behavior.

Since TIP5P includes explicit charge sites representing water's electron lone-pairs, the distribution of Figure 2.13 can be repeated using the lone-pair angles. The result should be unsurprising given that TIP5P is a largely symmetric, tetrahedral model, see Figure 2.15.

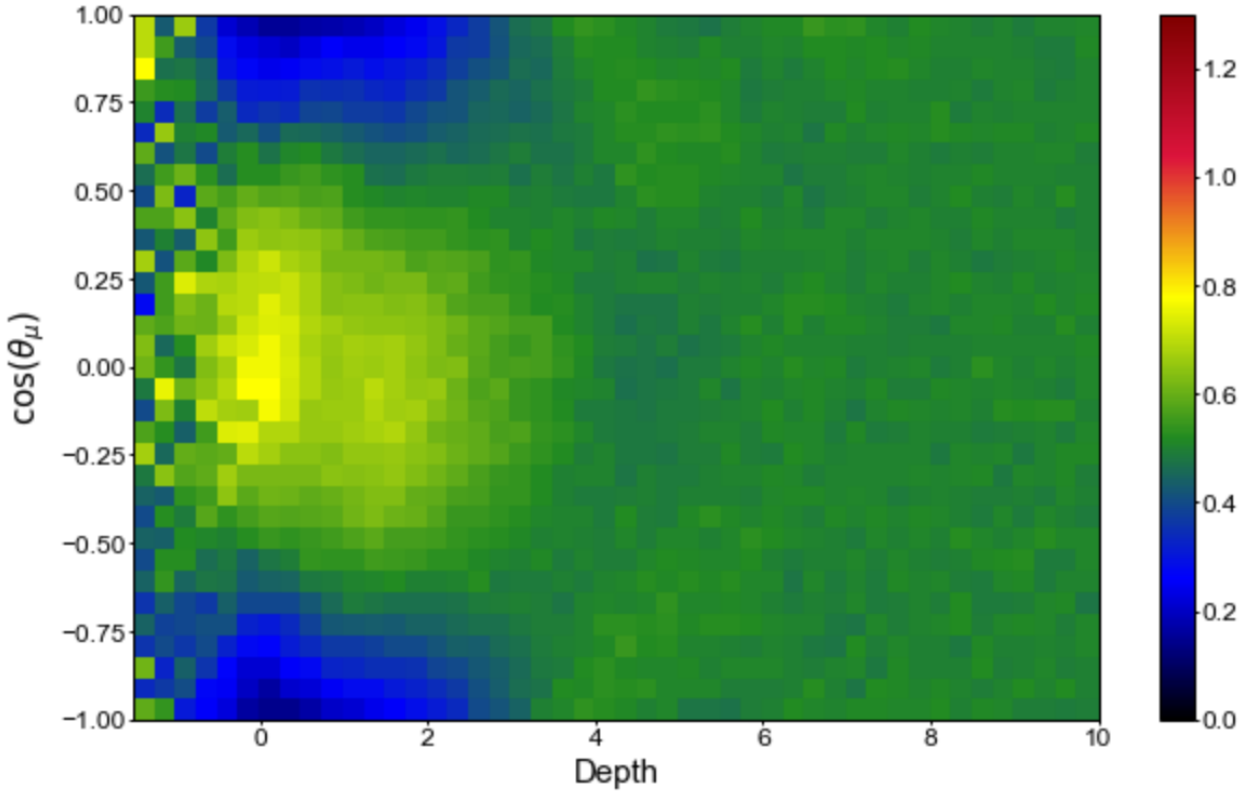

Figure 2.16: Dipole distribution of TIP5P.

The dipole distribution of TIP5P, however, largely matches with SPC/E, although the strength of the pattern in dipole is weaker here.

Orientational patterns are ice-like

Many of water's surface motifs match with the orientations of the basal plane of ice. Figure 2.17 is a representation of the first two bilayers of the basal plane of ice (surface toward the top of the page). Measurements of the simulated ice layer show strong similarities to liquid, Figures 2.18 and 2.19.

The ice crystal structure is based primarily off the fundamental building block of ice: the puckered hexagon. Water molecules form an extended lattice of hexagons, and each water oscillates sitting within either a higher or lower position. The high and low of the puckered hexagon is equivalent to the characteristic bilayer of the basal plane of ice. Each vertex of the hexagon forms three hydrogen bonds with three neighbors in the same bilayer, while a fourth hydrogen bond extends directly upward or downward and connects bilayers to one another. In effect, ice is composed of stacks of extended, puckered hexagons.

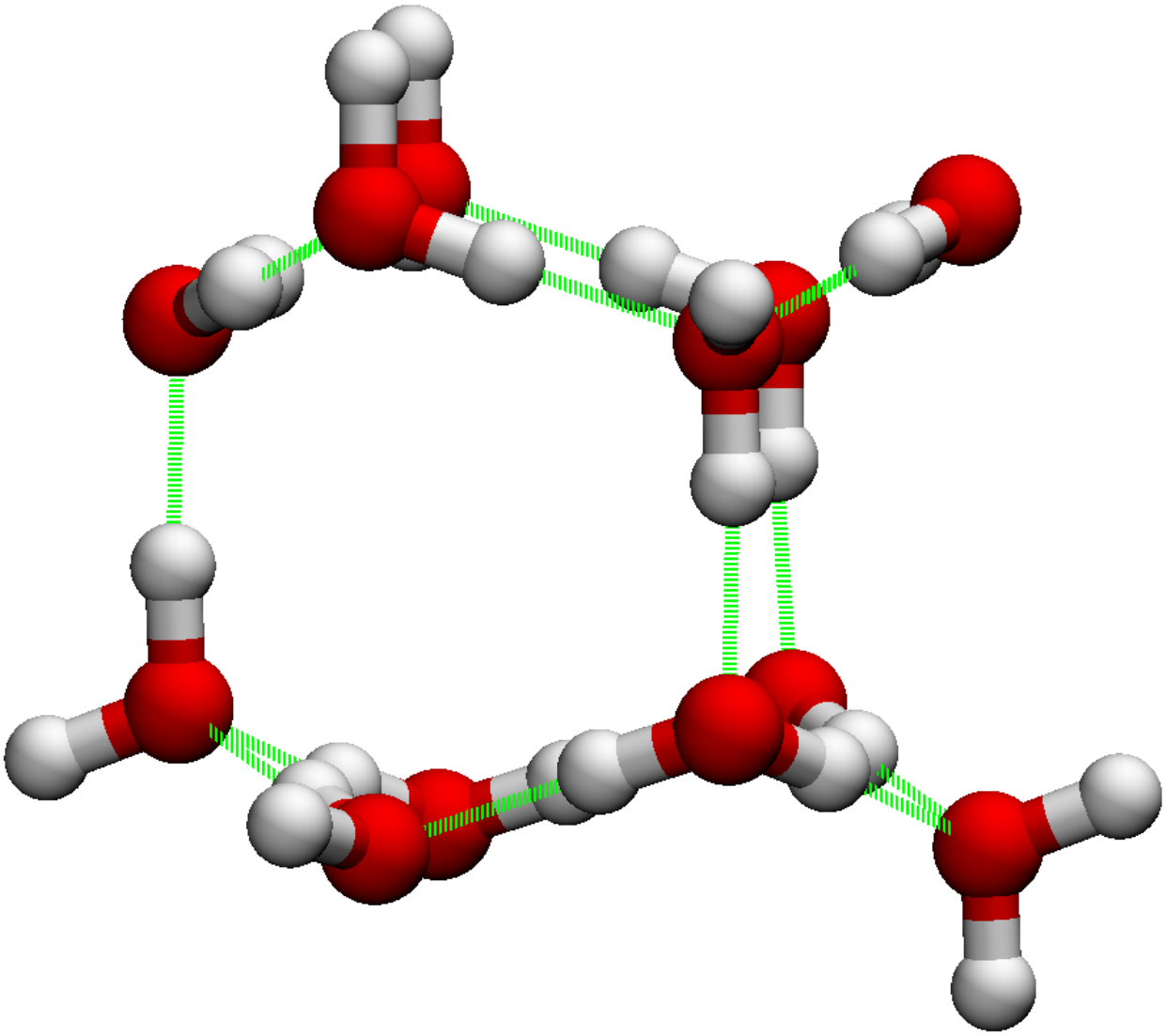

Figure 2.17: Visualization of two bilayers of SPC/E ice (oxygen-red, hydrogen-white, hydrogen bonds-green). The bilayer-bilayer separation of approximately 3.6 Å matches the separation of layers in liquid water. Two pictured molecules in the top bilayer point one hydrogen upward, while two other molecules in the same bilayer point both hydrogen atoms mainly parallel, in line with the interface, and the last two molecules in the top bilayer bond with the layer of water below them.

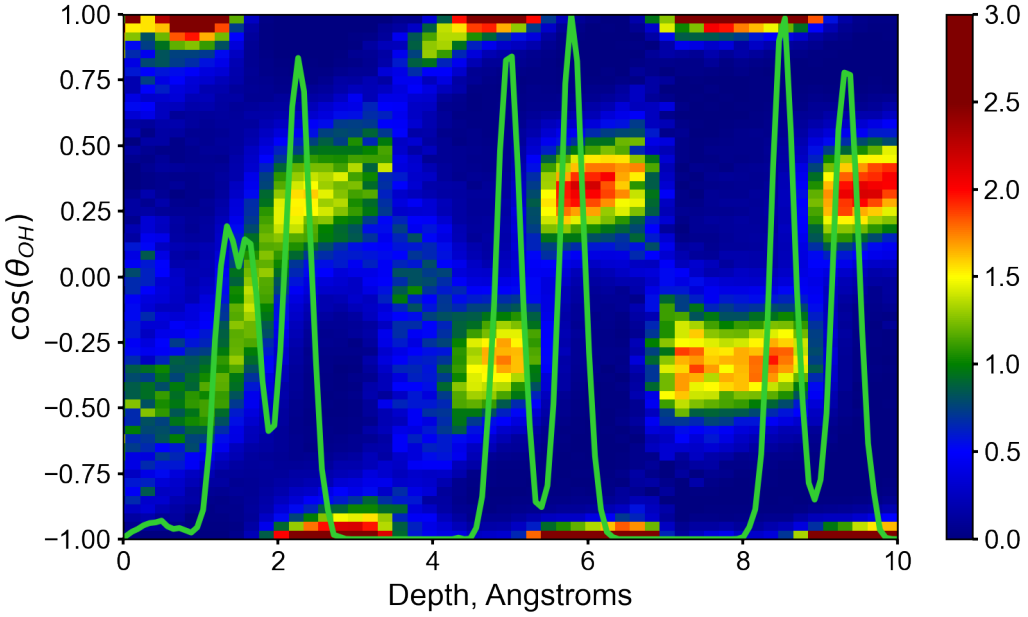

Figure 2.18: Orientation of the OH-bond vectors in SPC/E ice below freezing, compare Figures 2.10 and 2.13. The main structural motif at the liquid surface matches the repeated lattice patterns of ice.

The hydrogen bond acceptor/donor pattern throughout ice is not predetermined by the ice crystal except that each water molecule settles into a tetrahedral arrangement. Contrary to some illustrations, there is no requirement that the topmost water molecules point only a dangling hydrogen outward, while accepting two hydrogen bonds, since this would be entropically unlikely and, due to a buildup of dipole interactions, also energetically unfavorable.

Accordingly, there is no requirement that the top half of the surface bilayer of ice (or, by connection, liquid water) exhibit an excess of "dangling hydrogen bonds", as has been suggested[18], but such a feature may appear to grow in as all features near the surface become more exaggerated at low temperatures. Figure 2.19 suggests that the top bilayer of ice does show the dangling hydrogen feature, but that the parallel feature is also strong.

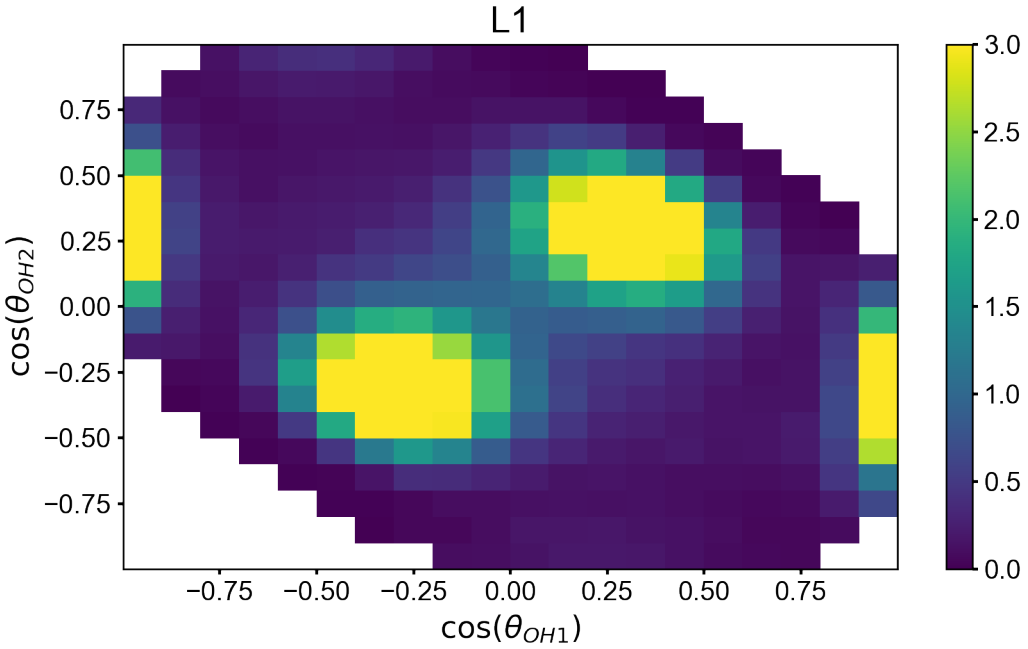

Figure 2.19: A parallel layer is also observable in the crystal structure of ice, compare Figure 2.7. Since OH1 and OH2 are indistinguishable, this heat map should be symmetric. Ice diffuses very slowly and is hard to sample, but the orientational probabilities match expectations from the ice crystal structure.

Bonding trends at the water interface

Water's high surface tension derives from the strength of hydrogen bonds in bulk. The strong patterning of water's surface is mainly a result of water aligning itself in a way that maximizes total hydrogen bonds, which is key for the surface-most waters. Water necessarily sacrifices bonds at an interface, but even at the extreme surface, water still tends to form two hydrogen bonds on average [16], see Figure 2.20.

Normally, a hydrogen bond is classified as such if the two water centers are less than 3.5 Å apart and the relevant OH bond vector forms an angle of less than 30° with the water-water separation vector. A strictly geometric definition of a hydrogen bond is necessary to classify hydrogen bonds in a classical simulation of point-charge molecules.

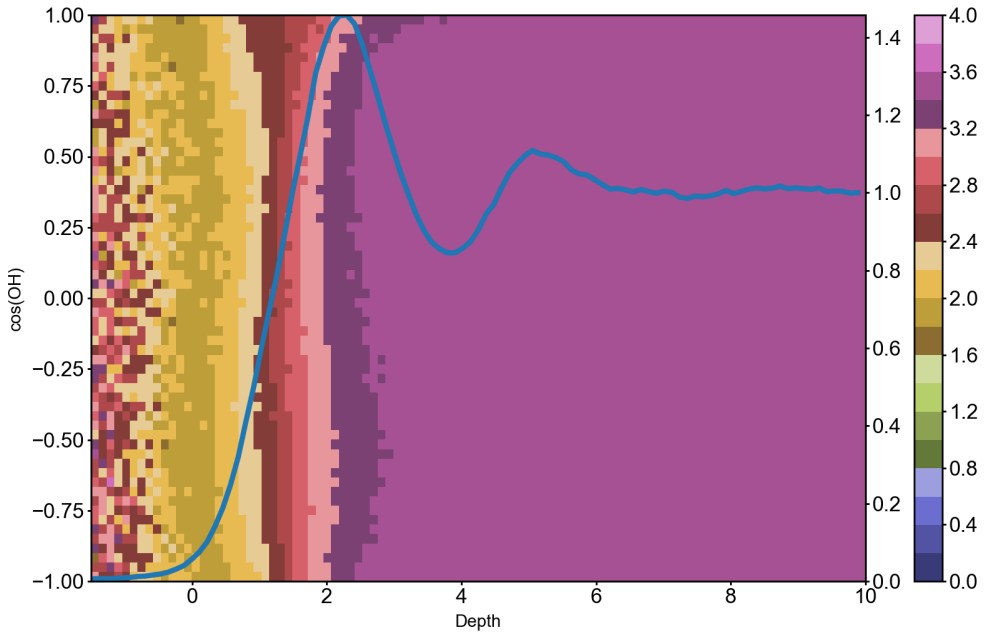

Figure 2.20: Number of hydrogen bonds, on average, at a given depth and orientation. Data beyond $d = -0.5$ becomes very sparse. The tendency to hydrogen bond is not strongly dependent on angle, though some dependence exists. The bonding behavior depends strongly on depth. Even at the interfacial boundary, water molecules tend to form two hydrogen bonds.

More cuts to the data could describe bonding patterns near the surface, and the trends will vary slightly depending on the water model. SPC/E and TIP5P were shown above to differ in orientation near the surface and the bonding patterns are likewise governed mainly by water's intrinsic tetrahedrality.

Figure 2.21 shows the angular bonding behavior of SPC/E near the interface. The angle is measured between the SPC/E dipole and the water-neighbor separation vector of hydrogen-bonded neighbors (see inset). The two vertical lines mark where a perfectly tetrahedral bonding arrangement would sit. Figure 2.21 shows mainly that bonding geometries do not differ substantially moving from bulk to the surface until the very surface-most waters in $L1_{\parallel}$ and L0. The distribution of neighbors for SPC/E is rather sharp relative to the hydrogen atoms near 55°, but becomes very loose on the tail end where the implied lone pairs sit. SPC/E represents the two water lone pairs in a very loose way, with only a negative charge on the central oxygen atom. The opposite extreme can be seen with a model like TIP5P that explicitly represents water lone pairs with negatively charged dummy atoms. Actual water,

and actual water bonding behavior near the interface, is probably somewhere between these two fully-classical extremes.

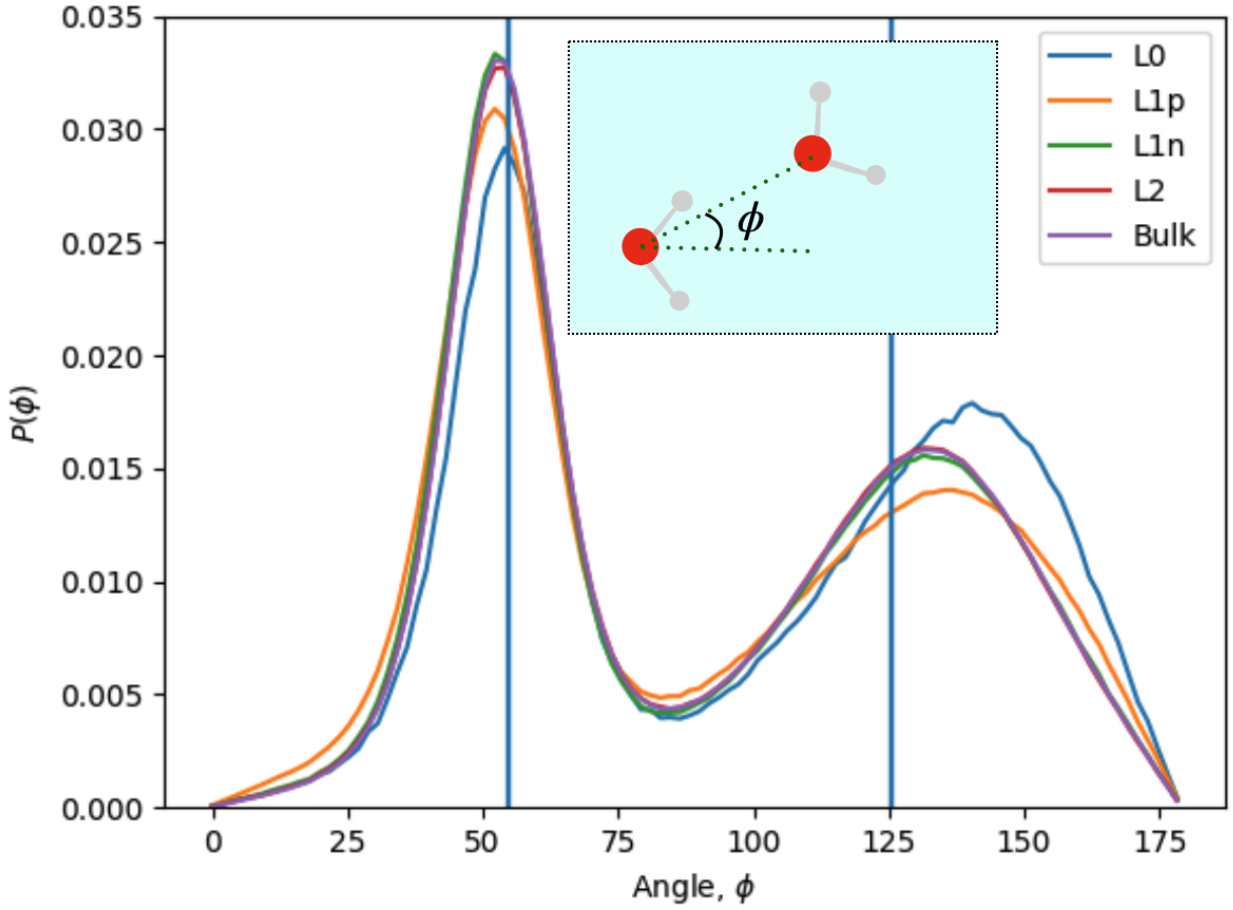

Figure 2.21: Distribution of hydrogen-bonding neighbors around a central SPC/E molecule. Comparison between bulk liquid and surface layers shows almost no change to neighbor geometries until the top half of L1 and in L0. Angle ϕ is measured relative to the central atom dipole and the two vertical blue lines indicate where four neighbors in a perfectly tetrahedral arrangement would sit, illustrated in the inset.

With fewer than 4 neighbors, the likely configuration of bonding changes, especially around the negative side of an SPC/E molecule, which does not explicitly include lone pairs. Figure 2.22 shows that neighbors tend to sit around the "southern hemisphere" of an SPC/E molecule, i.e. the half of the sphere where the lone pairs would sit, with the "south pole" defined as the negative of the SPC/E dipole, $-\mu$. An SPC/E molecule that only accepts one bond is more likely to accept that bond directly along $-\mu$, rather than at a tetrahedral angle. However, in $L1_{\parallel}$ and L0, the average single hydrogen bond donating neighbor moves away from the negative dipole.

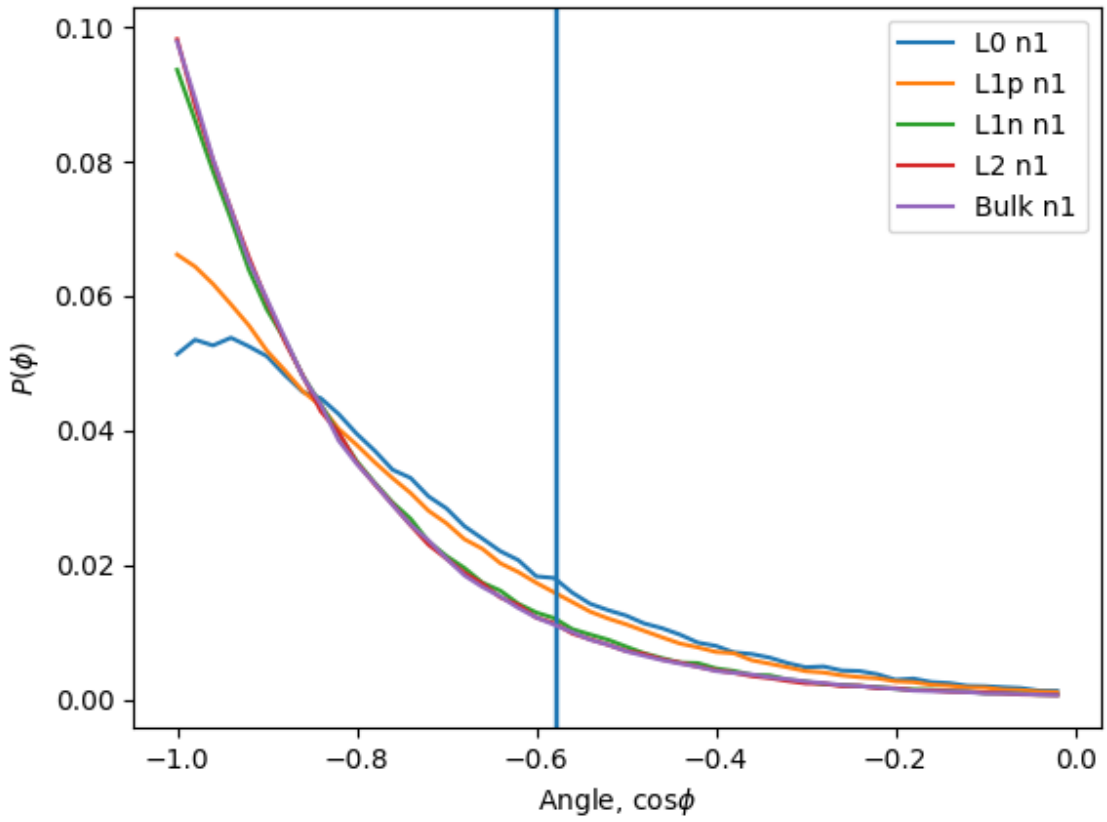

Figure 2.22: When only one SPC/E neighbor donates a hydrogen bond, the neighbor tends to sit more along the negative dipole, i.e. "at the south pole". At the very surface, however, that trend reverses, likely because sitting along the dipole at the surface would be too exposed to be favorable.

While the details in this section about bonding may have implications for water's surface, SPC/E is not a perfect representation of water. Many of the bonding and orientational

patterns of SPC/E near the water surface should be interpreted skeptically. Since the SPC/E model chooses not to explicitly represent lone pairs and includes only one partial negative charge on the central oxygen, it is not surprising that bonding behavior toward the "southern hemisphere" of SPC/E is flexible.

Surface dipole behavior

Much attention has been paid to the electrostatic structure at water's surface[21–29]. A classical simulation of point charges, like SPC/E, is conceptually easy to measure at water's surface but may not represent the reality of physical water, which is not fully classical in nature[30]. Regardless, simple point charge models are often schematically true and are much easier to treat theoretically, and so are worth studying. Figures 2.23 and 2.24 show measurements for SPC/E.

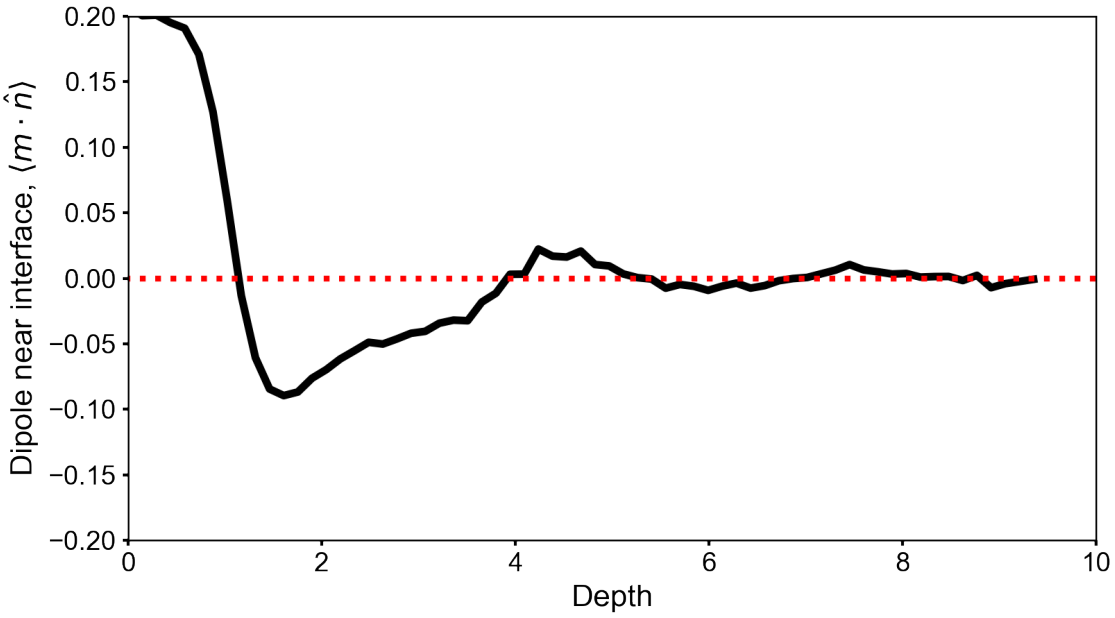

Figure 2.23: Expectation of SPC/E dipole component along surface normal as a function of depth, $\langle m \cdot \hat{n} \rangle (d)$, in units of SPC/E dipole.

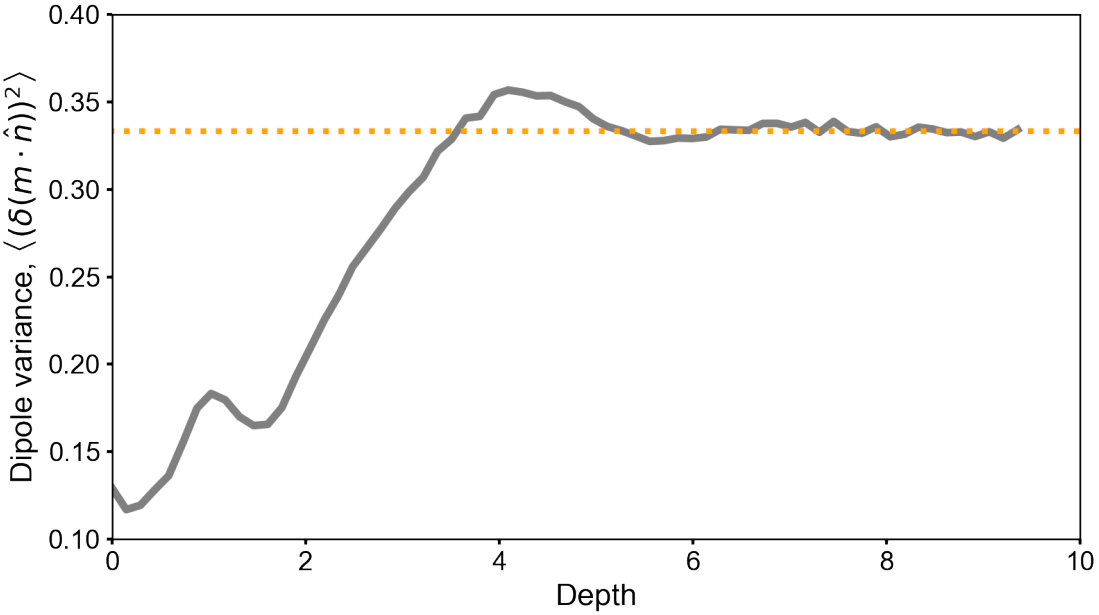

Figure 2.24: Variance of SPC/E surface-normal dipole component as a function of depth, $\langle(\delta(m\cdot\hat{n}))^2\rangle\,(d)$.

Similar dipole measurements for the the basal plane of SPC/E ice, alternatively, show significantly different dipole trends, despite the similarities in OH-vector orientations. Figures 2.25 and 2.26 display the dipole trends in for the first bilayer of ice, which is approximately repeated and made sharper toward the center of the ice crystal. It should be noted that the average dipole measured at the basal plane of ice can vary drastically for slightly different crystal arrangements, but dipole rearrangements for a frozen crystal are likely to be extremely slow.

While many orientational patterns match between the basal plane of ice and the surface of liquid SPC/E, the dipole trends differ substantially. This is another indication that the immediate neighbors of an SPC/E molecule at the liquid surface do not determine the dipole behavior and that other effects, such as multi-body configurations or local dipole fields, must be influential.

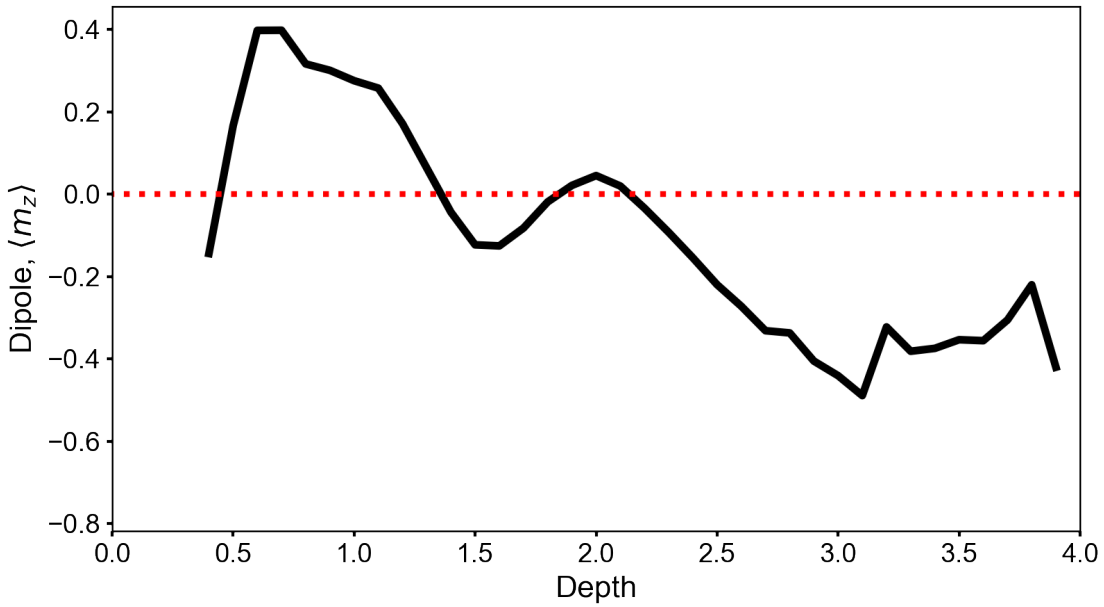

Figure 2.25: Average angle made between the dipole and surface normal in the first bilayer of SPC/E ice. These trends do not match the mean surface dipole of liquid SPC/E. Dipole statistics for ice surface converge slowly and can change drastically with initial configuration, so while the dipole average here is not converged and may not be representative, there is still no match with liquid SPC/E.

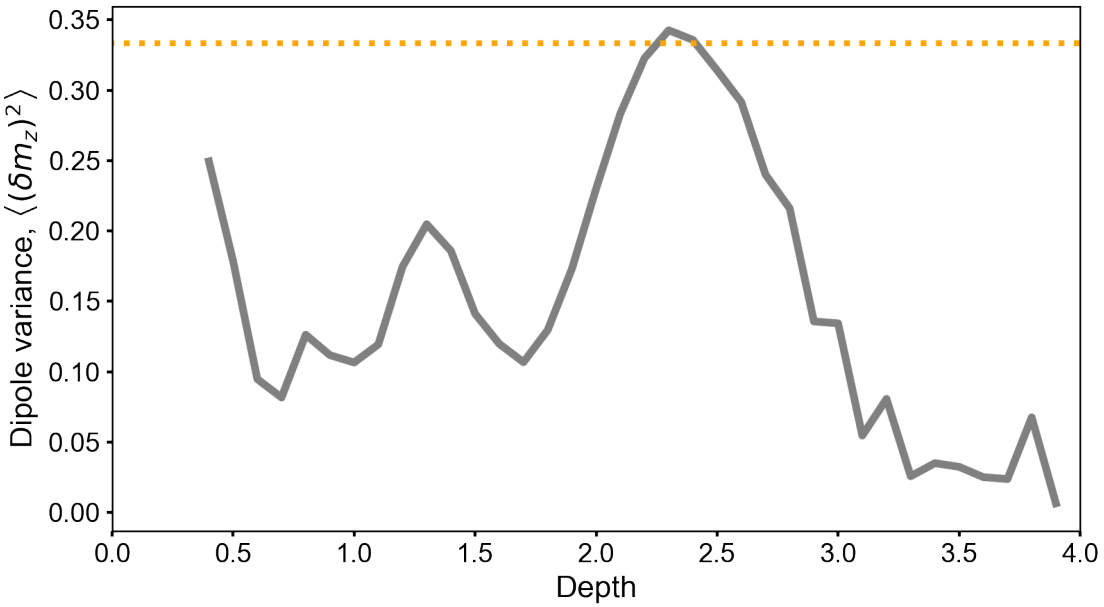

Figure 2.26: Variance of the angle made between the dipole and surface normal in the first bilayer of SPC/E ice. These trends do not match the mean surface dipole of liquid SPC/E.

2.3 Similarities in the Mercedes-Benz water model

The MB water model

The Mercedes-Benz water model (also called MB or BN2D) is a 2-dimensional simplification of water that enforces strong bonding structure, is inspired by the hydrogen bonding structure of liquid water, and has been studied in some depth[11, 12, 31–33]. Through its simplicity, MB water is an analytically approachable model of water that offers a 2-dimensional validation case for a theoretical understanding of 3-dimensional water.

MB molecules add a 2-dimensional Lennard-Jones energy E_{LJ} to three equidistant bonding arms that extend radially outward and interact with neighboring molecules using a special bonding energy E_{HB}. The three arms resemble the three spokes of the Mercedes-Benz logo (Figure 1.10), leading to the name Mercedes-Benz water.

The position and orientation of an MB particle i can be described with a 2-dimensional coordinate and the minimum angle that an arm makes with $\hat{x}$, as (x_i, y_i, θ_i), where $\hat{x}$ and $\hat{y}$ span the dimensions of the simulation box and θ_i is in the range $[0°, 120°)$.

Two MB molecules i and j are separated by a vector $\boldsymbol{r}_{ij}$ which is defined

$$\boldsymbol{r}_{ij} = \boldsymbol{r}_j - \boldsymbol{r}_i = r_{ij}\hat{\boldsymbol{u}}_{ij}. \tag{2.8}$$

The particle-particle interaction energy for those MB molecules is

$$E_{ij} = E_{LJ}(r_{ij}) + E_{HB}(r_{ij}, \theta_i, \theta_j), \tag{2.9}$$

$$E_{LJ}(r_{ij}) = 4\epsilon \left[\left(\frac{\sigma_{LJ}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{LJ}}{r_{ij}} \right)^{6} \right], \tag{2.10}$$

$$E_{HB}(r_{ij}, \theta_i, \theta_j) = -\epsilon_{HB} G(r_{ij} - r_{HB}) G(\boldsymbol{b}_i \cdot \hat{\boldsymbol{u}}_{ij} - 1) G(\boldsymbol{b}_j \cdot \hat{\boldsymbol{u}}_{ij} + 1), \tag{2.11}$$

with $\boldsymbol{b}_i$ and $\boldsymbol{b}_j$ defined as the unit vectors along the respective particle's arm that forms the smallest angle with $\hat{\boldsymbol{u}}$, as

$$\hat{\boldsymbol{b}_i} = \arg\min_{b_k; k=1,2,3} \left[-\boldsymbol{b}_k \cdot \hat{\boldsymbol{u}}_{ij} \right] \tag{2.12}$$

$$\hat{\boldsymbol{b}_j} = \arg\min_{b_k; k=1,2,3} \left[\boldsymbol{b}_k \cdot \hat{\boldsymbol{u}}_{ij} \right], \tag{2.13}$$

and k indexes over each of the particle's three arms.

An unnormalized Gaussian function (having a maximum value of 1) is used to compute the hydrogen bond energy,

$$G(x) = \exp \left[-\frac{x^2}{2\sigma_{HB}^2} \right]. \tag{2.14}$$

Table 2.1 summarizes the parameters and the values that are generally used to compute the energy of a system of MB particles[12, 31–33].

Table 2.1: Summary of MB water parameters.

Parameter	Value
r_{HB}	1
σ_{LJ}	0.7 r_{HB}
ϵ_{HB}	1
ϵ_{LJ}	0.1 ϵ_{HB}
σ_{HB}	0.085 r_{HB}

Surface trends in MB water

Interfacial structure of the MB water model can be characterized in close analogy to the treatment of 3-dimensional water models in Section 2.2. As discussed in Section 1.5, the low surface tension and relative sparsity of the MB water surface does not always allow for a smooth, continuous instantaneous interface at the liquid-vapor boundary. A careful coordinate projection or an artificial increase in surface tension may lead to a more flat interface, but would also significantly complicate the analysis. Alternatively, a hard wall mimics the loss of interactions to a vapor interface while also suppressing the problematic

surface fluctuations. The following data are measured using hard walls at x_{max} and x_{min} with periodic boundary conditions in y.

The density profile of MB water shows layering much like other water models, and there is persistent geometric structure at the interface. Figure 2.27 combines the density profile (blue) with the orientational preference of MB molecules (gray-scale). As with water, the pattern in orientations is explained by the tendency of an MB water near the interface to sacrifice one bond along an arm that is pointed into vapor while forming two bonds with neighbors just below it. The remaining patterns are mainly in response to the strong top-layer pattern. Compare to SPC/E in Figure 2.10.

The three-fold symmetry in Figure 2.27 arises from the definition of the angle θ. Because the three arms are indistinguishable, the minimum angle formed between a bonding arm and the surface normal cannot exceed 60 degrees,

$$\theta = \cos^{-1}\left(\frac{b_k \cdot \hat{n}}{|b_k|}\right), \qquad \theta \in [0, \pi/3]. \tag{2.15}$$

Data is reflected and repeated to cover the full range $[0, 2\pi)$ as a representation of one arm exploring the full rotation of a MB particle.

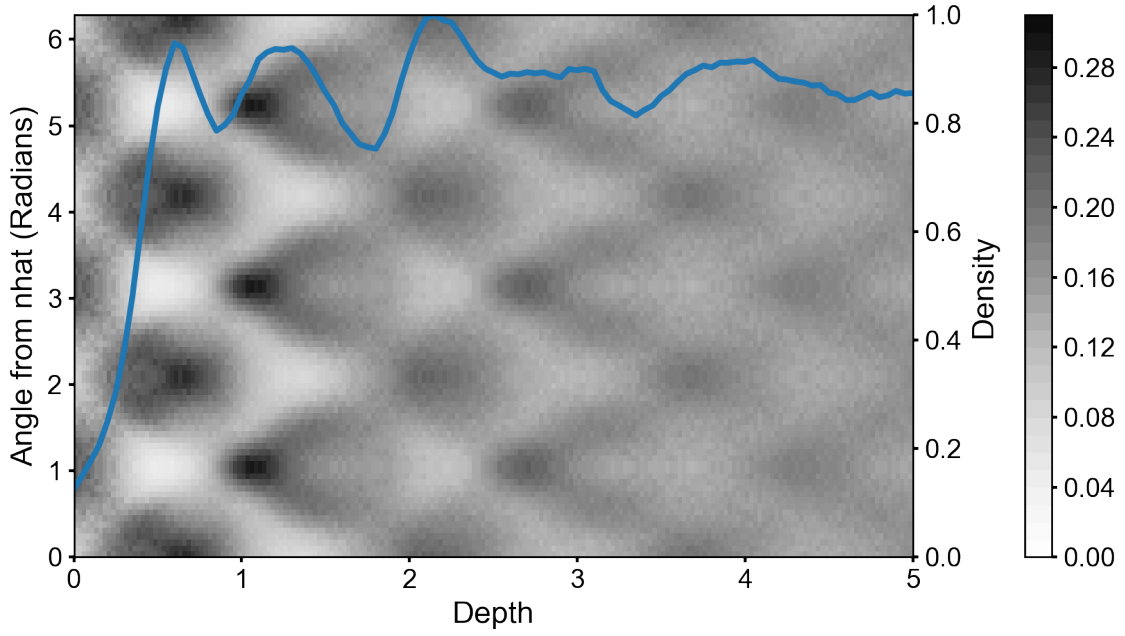

Figure 2.27: Density profile (blue) and probability of orientations (gray-scale) of the MB water model. Surface patterns are similar to the patterns observed in 3-dimensional water models. Data gathered from 260 particles at temperature $T^* = 0.16$, mean density 0.8, with hard walls in the positive and negative x-direction and a periodic y-coordinate 12 units long.

It is likely that the MB water density layering measured from Monte Carlo simulation averages over multiple structural motifs. Whereas SPC/E water exhibits diminishing oscillatory peaks, the density of MB water is not so regular. It is possible that a handful of long-lived and non-overlapping surface structures compete at the surface, like different flat cuts across a hexagon. Despite the irregular density profile, the angular patterns remain strong.

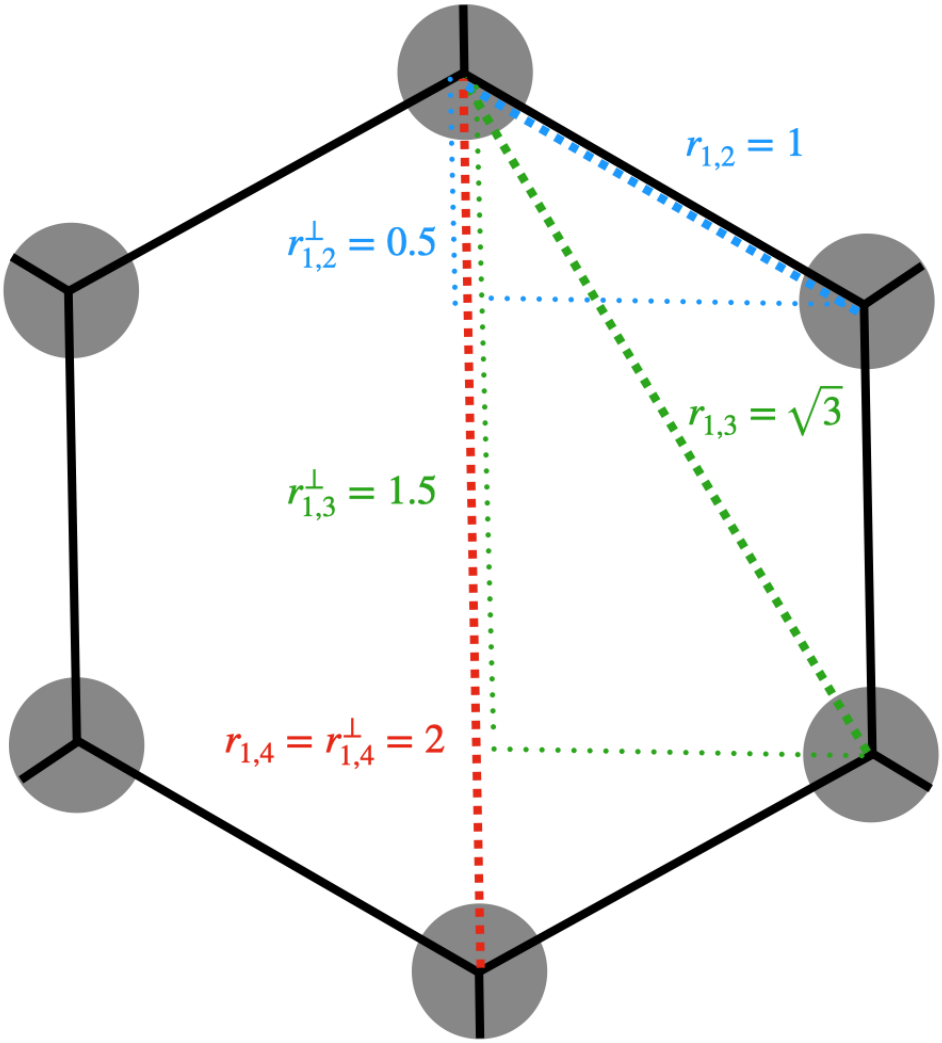

Figure 2.28: Frozen MB particles form an extended crystal lattice of perfect hexagons. Relevant lattice distances are shown in color and are in units of the MB model, r_{HB}.

Figure 2.28 depicts the ice geometry of MB water, which forms neat hexagons when perfectly frozen. The first density peaks in Figure 2.27 are at approximately the following depths $d_i = \{0.62, 1.25, 2.18, 2.83\}$ with deltas $d_i - d_0 = \{0.63, 1.56, 2.21\}$, which roughly

match the expected spacing from the hexagonal crystal structure of MB water. Moreover, in a perfect crystal the second and third peaks would each account for twice as many molecules as the first and fourth peaks, possibly explaining the relatively weak first peak in liquid MB water.

Summary

The abruptness of a liquid-vapor interface leads to a strong average geometric response in water characterized by a combination of layering, orientational, and bonding patterns. The interfacial trends are interrelated and stem mainly from water maximizing hydrogen bonding in response to the loss of bonding partners in the transition to the vapor phase. The outermost layer of water responds strongly to the loss of bonding opportunities, while subsequent layers react to the waters directly around them. Water behavior mostly decorrelates from the interface after about 1 nm. The structures at the surface of a liquid are closely related to the structure of the corresponding ice crystal, implying that molecules in both ice and liquid surfaces are oriented in a way that maximizes bonding.

Chapter 3

A Mean Field Explanation for Water's Surface

Overview

A method for predicting water surface orientation by treating density layers as a mean field was proposed by Shin and Willard in 2018[21], which was called the rigid tetrahedral model. A water molecule and the four neighbors it forms hydrogen bonds with can be represented as a rigid tetrahedron, and when near water's surface the vertices of that tetrahedron (the bonded neighbors) are more likely to be found in regions of higher relative water density.

This chapter provides a more thorough statistical mechanical foundation for the rigid tetrahedral approach than was originally offered. The derivation herein suggests improvements to the original theory in the form of self-consistency requirements that enhance the model's ability to capture collective stabilization of orientational patterns. Any strongly geometric liquid surface can be reasonably approached with this focus on the density-bonding relationship. such as the 2-dimensional Mercedes-Benz water model.

The rigid tetrahedral model is easily modified to imitate different behaviors of water or the water surface, such as flexibility in the tetrahedral geometry of a fully bonded water or by adding texture to the relevant water interface. By proposing specific changes to the mean field surface model, some characteristics of water are shown to be unimportant to surface geometry, like small changes to bond strength and flexibility, while others are essential, like the general tetrahedral shape and trends in surface dipole.

3.1 Theoretical basis

The mean field surface model is based on a central assumption that the layered density, a function of depth d from the instantaneous interface (see Chapter 2, especially Figure 2.4),

dictates the likelihood of forming a successful hydrogen bond at that depth,

$$P_{HB}(\boldsymbol{r}) \propto \rho(d). \tag{3.1}$$

The goal of the mean field water surface model is to describe orientational statistics of water molecules at the surface of water using only the geometry of water and a measured density field as inputs. The relevant probability distribution is

$$P(\kappa|r) \tag{3.2}$$

where r is the position of a water molecule and κ fully describes the molecule's orientation. In the case of MB water, a single angle θ is enough to uniquely describe orientation, but κ for 3-dimensional water is a more complex quantity.

Any given water molecule may donate and accept multiple hydrogen bonds. The hydrogen-bonding state α describes which of the bonding arms has a corresponding neighbor at an appropriate distance and orientation to form a strong bond. For example, for MB water the bonding state denoted $\alpha = [0, 1, 1]$ would specify that only the second and third arms participate in a hydrogen bond, while the first arm either does not point to a nearby neighbor, or the neighbor is not aligned so as to form a bond. For MB water, there are $2^3 = 8$ unique α. The exact molecular distance and angle that constitute a successful bond do not need to be defined at this point and it is advantageous to keep the definition flexible.

The notations $r^N = \left\{r^{(1)}, r^{(2)}, ..., r^{(N)}\right\}$ and $\kappa^N = \left\{\kappa^{(1)}, \kappa^{(2)}, ..., \kappa^{(N)}\right\}$ describe the positions r and orientations κ for one configuration of a collection of N molecules.

For a particular water molecule at position r, which will be referred to here as the "tagged water", the joint probability distribution of orientation κ and bonding state α can be written

$$P(\kappa, \alpha|r) = \frac{1}{P(r)} \int dr^N \int d\kappa^N \exp\left[-\beta U_N(r^N, \kappa^N)\right] \delta(r - r^{(1)})\delta(\kappa - \kappa^{(1)}) h_\alpha(r^{(N-1)}, \kappa^{(N-1)}; r^{(1)}, \kappa^{(1)}). \tag{3.3}$$

Equation 3.3 integrates over all possible configurations of the N particles and uses two delta-functions to isolate those configurations where the first particle has the specified position r and orientation κ. A Boltzmann factor weights the probability of the total configuration with the energy function $U_N(r^N, \kappa^N)$. The last term, h_α, is an indicator function that returns value 1 for configurations where the tagged water has neighbors satisfying the bonding state α and returns 0 otherwise.

The distribution of κ alone is the sum over the bonding state probabilities,

$$P(\kappa|r) = \sum_\alpha P(\kappa, \alpha|r). \tag{3.4}$$

Isolating the tagged water in Equation 3.3 leads to,

$$P(\kappa, \alpha|r) = \frac{1}{P(r)} \int dr^{N-1} \int d\kappa^{N-1} \exp\left[-\beta U_{N-1}(r^{N-1}, \kappa^{N-1}) - \beta V(r^{N-1}, \kappa^{N-1}; r, \kappa)\right] h_\alpha(r^{(N-1)}, \kappa^{(N-1)}; r, \kappa), \tag{3.5}$$

where the energy has been decomposed into U_{N-1}, which is the energy of the configuration without considering the tagged water, and V, which is the energy contributed by adding the tagged water.

The energy of the tagged water V is decomposed further into an isotropic piece V_{iso} and a non-isotropic piece ΔV,

$$V(r^{N-1}, \kappa^{N-1}; r, \kappa) = V_{iso}(r^{N-1}, \kappa^{N-1}; r) + \Delta V(r^{N-1}, \kappa^{N-1}; r, \kappa) \tag{3.6}$$

to be used to define an isotropic energy term,

$$U_0(r^N, \kappa^{N-1}) = U_{N-1}(r^{N-1}, \kappa^{N-1}) + V_{iso}(r^{N-1}, \kappa^{N-1}; r). \tag{3.7}$$

The energy U_0 does not change as the orientation κ of the tagged water is varied.

The probability of Equation 3.5 expressed in terms of the fully isotropic energy

$$\begin{aligned} P(\kappa, \alpha|r) = \frac{1}{P(r)} \int dr^{N-1} \int d\kappa^{N-1} \\ \exp\left[-\beta U_0(r^N, \kappa^{N-1}) - \beta \Delta V(r^{N-1}, \kappa^{N-1}; r, \kappa)\right] h_\alpha(r^{(N-1)}, \kappa^{(N-1)}; r, \kappa) \end{aligned} \tag{3.8}$$

can be abbreviated further by referencing the isotropic partition function,

$$Q_0 = \int dr^{N-1} \int d\kappa^{N-1} e^{-\beta U_0}, \tag{3.9}$$

$$P(\kappa, \alpha|r) = \frac{Q_0}{P(r)} \left\langle h_\alpha e^{-\beta \Delta V} \right\rangle_0. \tag{3.10}$$

With the property

$$\langle f \rangle_{0,\alpha} = \frac{1}{Q_{0,\alpha}} \int dr^{N-1} \int d\kappa^{N-1} e^{-\beta U_0} h_\alpha f \tag{3.11}$$

$$= \frac{Q_0 \langle h_\alpha f \rangle_0}{Q_0 \langle h_\alpha \rangle_0}, \tag{3.12}$$

the probability is succinctly written

$$P(\kappa, \alpha|r) = \frac{Q_0}{P(r)} \langle h_\alpha \rangle_0 \left\langle e^{-\beta \Delta V} \right\rangle_{0,\alpha}. \tag{3.13}$$

To summarize, the individual distributions can be written, without any assumptions, as follows

$$Z(r) = \int d\kappa \sum_\alpha \langle h_\alpha \rangle_0 \left\langle e^{-\beta \Delta V} \right\rangle_{0,\alpha} \tag{3.14}$$

$$P(\alpha|r) = \int d\kappa P(\kappa, \alpha|r) = \frac{1}{Z(r)} \int d\kappa \langle h_\alpha \rangle_0 \left\langle e^{-\beta \Delta V} \right\rangle_{0,\alpha} \tag{3.15}$$

$$P(\kappa|r) = \sum_\alpha P(\kappa, \alpha|r) = \frac{1}{Z(r)} \sum_\alpha \langle h_\alpha \rangle_0 \left\langle e^{-\beta \Delta V} \right\rangle_{0,\alpha}. \tag{3.16}$$

Necessary assumptions for the mean field model

In order to use the formulation in Equation 3.16, some assumptions need to be made about h_α and ΔV.

The expectation of the indicator function $\langle h_\alpha \rangle$ is the probability of finding a configuration with multiple neighbors aligning to match hydrogen bonding state α. The multiple neighbors must meet two conditions: (A) be in the specified location, and (B) point in the right direction.

Condition (A) is addressed by assuming that the existence of a neighbor very near a site of potential hydrogen bonding is proportional to the average density $\rho(d)$ at the relevant depth d, i.e. Equation 3.1.

Condition (B) references the probability of a neighbor orientation $P(\kappa|d)$, but this is exactly the quantity that the mean field surface model seeks to predict. Some form of $P(\kappa|r)$ must be assumed, measured from simulation, or approximated in another way. The final mean field model described in this chapter will first assume a uniform $P(\kappa|r)$, and then iteratively self-reference to solve for a final, self-consistent $P(\kappa|r)$.

These two assumptions are combined in an ansatz for the expectation of the hydrogen bonding state indicator function,

$$\langle h_\alpha \rangle_0 (r) \approx \left[\prod_{i=1}^{m_\alpha} a_i \rho(s_i) \int d\kappa_i P(\kappa_i|s_i) e^{-\beta E(\kappa_i|s_i)} \right] \times \left[\prod_{j=m_\alpha+1}^{M} 1 - a_j \rho(s_j) \int d\kappa_j P(\kappa_j|s_j) e^{-\beta E(\kappa_j|s_j)} \right]. \tag{3.17}$$

For a tagged molecule at position r, the probability of bonding state α is the product of the probabilities of m_α successful bonds and $M - m_\alpha$ potential but unfulfilled bonds. Realistically, water can form up to $M = 5$ hydrogen bonds and MB water can form $M = 3$, although a larger M could index other possible bonding locations. Each potential bonding site is referenced by position s_i and orientation κ_i. Equation 3.17 asserts that the probability of a successful bond at site s_i is proportional to both the local density $\rho(s_i)$ and probability of a valid orientation $P(\kappa_i|s_i)$. The probability of a neighbor orientation is modified by an energy function $E(\kappa|s)$, which accounts for imperfect angle alignments.

The distances between the tagged water and its bonding neighbors are not allowed to vary in this formulation, and all $r - s_i = b$, the bonding length. In reality, the length of a hydrogen bond may vary, but within a relatively small window. The length of a hydrogen bond in water is well estimated by the first peak in the water $g_{OO}(r)$, see Figure 2.2, which spans from a maximum near 2.8 Å to a minimum around 3.5 Å[14, 34]. So long as density and orientational probability do not vary sharply on the 0.5 Å length scale, an set distance of $b = 3$ should only add small errors to functions of s_i.

A constant of proportionality, a_i, normalizes the density $\rho(s_i)$ and the orientation term $P(\kappa_i|s_i)$. The constant a_i also potentially encapsulates all other information about the average of configurations $\{(r^N, \kappa^N)\}$, and could be different for every bonding site i. However,

since $\rho(s_i)$ and $P(\kappa|s_i)$ were chosen specifically to capture the most relevant factors driving bonding behavior, it is reasonable to set $a_i = a$ as a constant fit and normalization parameter.

An additional assumption is made for the non-isotropic energy function ΔV. In its exact form, ΔV is different for every configuration (r^N, κ^N). However, ΔV can be closely approximated by only considering the molecules near enough to form a bond, i.e. the M neighbors specified by bonding state α,

$$\Delta V_\alpha = \epsilon_\alpha, \tag{3.18}$$

$$\left\langle e^{-\beta \Delta V} \right\rangle_{0,\alpha} = e^{-\beta \epsilon_\alpha}. \tag{3.19}$$

The Shin-Willard model assumes a uniform energy for every hydrogen bond formed, suggesting

$$\Delta V = m_\alpha \epsilon_W, \tag{3.20}$$

where ϵ_W is a parameter fit to a measured probability distribution.

With the added assumptions, Equation 3.16 becomes a self-referencing mean field expression for predicting the orientation of a water molecule at an arbitrary position,

$$P(\kappa|r) = \frac{1}{Z(r)} \sum_\alpha e^{-\beta \epsilon_\alpha} \left[\prod_{i=1}^{m_\alpha} a\rho(s_i) \int d\kappa_i P(\kappa_i|s_i) e^{-\beta E(\kappa_i|s_i)} \right] \left[\prod_{j=m_\alpha+1}^{M} 1 - a\rho(s_j) \int d\kappa_j P(\kappa_j|s_j) e^{-\beta E(\kappa_j|s_j)} \right] \tag{3.21}$$

This form of $P(\kappa|r)$ can be solved numerically using the density mean field $\rho(d)$ measured from MD simulation. The energies ϵ_α can also be inferred from simulation, or the Shin-Willard assumption of a constant energy-per-bond can be used. The neighbor orientations $P(\kappa|r)$ are first assumed to be uniform, then are solved through iterative self-reference. The specific form of Equation 3.21 will be elaborated upon through the rest of this chapter.

Validity in bulk

Although designed for modeling the interface, Equation 3.21 makes predictions for bulk liquid as well. Given an isotropic bulk, dependence on r and κ should disappear,

$$P(\kappa|r) = \frac{1}{\Omega} \tag{3.22}$$

$$e^{-\beta E(\kappa|r)} = \delta(\kappa - \kappa') \tag{3.23}$$

$$\rho(r) = \bar{\rho}, \tag{3.24}$$

where the constant Ω normalizes a uniform distribution of κ.

Defining a constant,

$$c = \frac{a\bar{\rho}}{\Omega}. \tag{3.25}$$

and noting that Z is independent of r in bulk, the probability distribution simplifies to

$$P_{bulk}(\kappa, \alpha|r) = \frac{1}{Z} e^{-\beta\epsilon_\alpha} c^{m_\alpha} (1-c)^{M-m_\alpha} \tag{3.26}$$

$$= \frac{1}{Z} (1-c)^M e^{-\beta\epsilon_\alpha} \left(\frac{c}{1-c}\right)^{m_\alpha} \tag{3.27}$$

$$= CL(\alpha) \tag{3.28}$$

where $C = \frac{(1-c)^M}{Z}$.

The behavior of the neighborhood is encapsulated in the function

$$L(\alpha) \equiv e^{-\beta f_\alpha}, \tag{3.29}$$

where f_α has the form of a free energy,

$$f_\alpha = \epsilon_\alpha - m_\alpha k_B T \ln\left(\frac{c}{1-c}\right). \tag{3.30}$$

The bulk probability distribution is now simply normalization constants and a Boltzmann factor that changes with total bonding state α

$$P_{bulk}(\kappa, \alpha) = Ce^{-\beta f_\alpha}. \tag{3.31}$$

The form of f_α suggests strongly that the bin-widths within c define the scale of entropy. The constant a, which contributes to c, combines both a normalization for $\rho(r)$ and an effective confinement for a neighbor at s_i.

Equivalence to the rigid tetrahedral model

The rigid tetrahedral model proposed by Shin and Willard[21] effectively makes two assumptions to the mean field surface formulation:

1. The free energy ΔV is linear with the number of hydrogen bonds formed, m_α, and each hydrogen bond contributes an energy ϵ_W, as $\epsilon_\alpha = \sum_{i=1}^{m_\alpha} \epsilon_W$.

2. All neighbor orientations are considered equally likely, $P(\kappa|s) = \frac{1}{\Omega}$

   Because there is no reference to neighbor orientations, $e^{-\beta E(\kappa|s)} = \delta(\kappa - \kappa')$

With these assumptions,

$$P(\kappa|r) = \frac{1}{Z(r)} \sum_{\alpha} \left[\prod_{i=1}^{m_\alpha} \frac{a}{\Omega} \rho(s_i) \right] \left[\prod_{j=m_\alpha+1}^{M} 1 - \frac{a}{\Omega} \rho(s_j) \right] e^{-\beta m_\alpha \epsilon_W}. \tag{3.32}$$

Additionally, an orientation κ is defined by the set of four bonding vectors $\{b_1, b_2, b_3, b_4\}$ that point to the vertices of a tetrahedron centered on the tagged water at r. A bonding state $\{\alpha\}$ is described explicitly with an indicator function $h_b(s_i)$, where $s_i = r + b_i$.

$$P(\kappa|r) = \frac{1}{Z(r)} \sum_{\alpha} \left[\prod_{i=1}^{M=4} h_b(s_i) e^{-\beta \epsilon_W} \frac{a}{\Omega} \rho(s_i) \right] \left[\prod_{i=1}^{M=4} (1 - h_b(s_i))(1 - \frac{a}{\Omega} \rho(s_i)) \right] \tag{3.33}$$

$$= \frac{1}{Z(r)} \sum_{\alpha} \prod_{i=1}^{M=4} \left[h_b(s_i) e^{-\beta \epsilon_W} \frac{a}{\Omega} \rho(s_i) + (1 - h_b(s_i))(1 - \frac{a}{\Omega} \rho(s_i)) \right] \tag{3.34}$$

$$= \frac{1}{Z(r)} \sum_{h_b(b_1)=0,1} \sum_{h_b(b_2)=0,1} \sum_{h_b(b_3)=0,1} \sum_{h_b(b_4)=0,1} \prod_{i=1}^{M=4} \left[h_b(s_i) e^{-\beta \epsilon_W} \frac{a}{\Omega} \rho(s_i) + (1 - h_b(s_i))(1 - \frac{a}{\Omega} \rho(s_i)) \right]. \tag{3.35}$$

Evaluation through the sums over bonding indicators h_b makes clear the match to the Shin-Willard rigid tetrahedral form, but with more explicit constants of proportionality,

$$P(\kappa|r) = \frac{1}{Z(r)} \prod_{i=1}^{M=4} \left[1 + \frac{a}{\Omega} \rho(s_i)(e^{-\beta \epsilon_W} - 1) \right]. \tag{3.36}$$

3.2 Computational approach to the mean field surface

The theoretical formulation in Equation 3.21 is concise, but the computational implementation needs a more explicit description of each term.

To make the mean field surface functional form independent of units and binning, the density and neighbor orientation are expressed as probability functions relative to a bulk liquid reference. Using $\bar{\rho}$ as the mean density of the bulk liquid, define

$$\phi(r) = \frac{\rho(r)}{\bar{\rho}}. \tag{3.37}$$

For orientations,

$$W(\kappa|r) = \int d\kappa \frac{P(\kappa|r)}{P_{bulk}(\kappa)} e^{-\beta E(\kappa|r)} \tag{3.38}$$

The form of the energy $E(\kappa|r)$ remains flexible and can include an angular tolerance or a bonding potential. Options for reasonable energy functions to model bonding are discussed in Section 3.2.

As in Section 3.1, define also

$$C(r) = \frac{(1-c)^M}{Z(r)}. \tag{3.39}$$

The probability function can be expressed in terms of f_α as

$$P(\kappa|r) = C(r) \sum_\alpha e^{-\beta f_\alpha} \prod_{i=1}^{m_\alpha} \phi(s_i) W(\kappa_i|s_i) \prod_{j=m_\alpha+1}^{M} \frac{(1 - c\phi(s_j) W(\kappa_j|s_j))}{1-c}, \tag{3.40}$$

or in terms of ϵ_α,

$$P(\kappa|r) = C(r) \left(\frac{c}{1-c}\right)^M \sum_\alpha e^{-\beta \epsilon_\alpha} \prod_{i=1}^{m_\alpha} \phi(s_i) W(\kappa_i|s_i) \prod_{j=m_\alpha+1}^{M} \frac{1}{c} - \phi(s_j) W(\kappa_j|s_j). \tag{3.41}$$

Defining an orientation, κ

Up to this point, κ has remained an abstract specification of a molecular orientation. For MB water, $\kappa = \theta$ is enough to uniquely describe a molecular orientation, where θ is the angle one arm makes with a reference direction. Defining the orientation of a tetrahedron in 3-dimensions requires at least three pieces of information, and a few conventions for this already exist. Euler angles, notated here as $\kappa = [\theta, \eta, \psi]$, define three successive angular rotations of a vector. Take $\kappa = [0, 0, 0]$ as indicating that the water OH_1 bond vector points along the interface surface normal, which is assumed here to be along the lab frame $\hat{n} = \hat{z}$ for clarity. OH_2 sits in the xz-plane at a tetrahedral angle (109.5°) to OH_1, and $OH_2 \cdot \hat{x} > 0$.

The declination angle θ defines the angle made between OH_1 and $\hat{z}$, where $\theta \in [0, \pi]$. The angle η sets the rotation of OH_1 around $\hat{z}$ as the angle made by $\hat{x}$ and a projection of OH_1 onto the xy-plane, OH_1^{xy}. The combination of θ and η fully describes the position of the bond vector OH_1, but does not uniquely define the other three corners of the tetrahedron. The final angle ψ does so by rotating the tetrahedron about the OH_1 vector, expressed as an angle between $OH_1^\perp$ (which points along $\hat{x}$ when $\kappa = [0, 0, 0]$) and $OH_2^\perp$ (which is projected to the plane normal to OH_1). The angles are computed

$$\theta = \cos^{-1}\left(\frac{OH_1 \cdot \hat{z}}{|OH_1|}\right) \qquad \theta \in [0, \pi] \tag{3.42}$$

$$\eta = \cos^{-1}\left(\frac{OH_1^{xy} \cdot \hat{x}}{|OH_1^{xy}|}\right) \qquad \eta \in [0, 2\pi) \tag{3.43}$$

$$\psi = \cos^{-1}\left(\frac{OH_1^\perp \cdot OH_2^\perp}{|OH_1^\perp||OH_2^\perp|}\right) \qquad \psi \in [0, 2\pi) \tag{3.44}$$

The Euler angles are a conventional and convenient way of describing the orientation of a body in three dimensions. Unfortunately, a histogram of the rotations of a tetrahedron in θ, η, ψ does not produce a uniform bin density. All four bond vectors are the same length and will each point to a different location on the surface of a sphere, but the surface area defined by a step in θ, η, ψ will not be uniform for all θ, η, ψ. Moreover, not every combination of the three angles describes a unique state. When $\theta = 0$, a rotation in η and a rotation in ϕ become indistinguishable, which is a problem known as gimbal lock. The degeneracy of rotations and angle-dependent bin density is remedied by exploiting symmetries and making a geometric transformation.

While the height of the instantaneous interface fluctuates substantially in the lab-frame z, the roughness can be projected out into the depth coordinate, d. Since the interface tends to vary slowly on the scale of a water molecule, the local interface can be modeled as a flat plane. The angle η describes a rotation in the xy plane and so does not change the depth of any of the neighbor sites s_i. Since only changes to depth matter in Equation 3.41, all η are equivalent in terms of depth for the mean field surface formulation and η can be held constant by symmetry.

$$P(\theta, \psi|d) = \frac{1}{V_\eta} \int d\eta P(\kappa|d) \tag{3.45}$$

The surface area swept out by the OH_1 vector on a sphere is uniform when θ is transformed to $\cos(\theta)$ without any need to change ψ. Although the other three bonding vectors are not uniform in $(\cos(\theta), \psi)$, a meaningful probability can average through the rotation ψ, leading to the simplified probability function

$$P(\cos(\theta)|d) = \int d\psi P(\cos(\theta), \psi|d), \tag{3.46}$$

which still carries significant information about a given water's orientation.

Options for bonding tolerance

The probability of forming a bond with a neighbor at s_i in the mean field surface model is proportional to the mean density at s_i times the probability of a bond-like neighbor orientation,

$$P_{HB}(\kappa_i|s_i) \propto a\rho(s_i) \int d\kappa_i P(\kappa_i|s_i) e^{-\beta E(\kappa_i|s_i)}. \tag{3.47}$$

The match with a neighbor's orientation can be expressed in a number of reasonable ways. The rules for "matching" a neighbor orientation are encapsulated in the function $E(\kappa|s)$. It will be shown that the details of E are largely unimportant to the numerical results.

Direct match

The simplest function E makes a direct orientational match: given a tagged water at r with bond vector pointing to s_i, there is a neighbor centered at s_i in one of a few possible orientations κ_r with a matching bond vector pointed directly back at r.

$$E(\kappa_i|s_i) = \delta(\kappa' - \kappa_r) \tag{3.48}$$

$$W_D(\kappa_i|s_i) = a_D \int d\kappa' \frac{P(\kappa'|s_i)}{P_{bulk}(\kappa')} \delta(\kappa' - \kappa_r). \tag{3.49}$$

Matched window

Alternatively, W can sum matches within a window of reasonable angles, where the tagged bond vector b_r is nearly aligned within a tolerance γ_{max} to a valid neighbor bond vector b_s,

$$W_w(\kappa_i|s_i) = a_w \int d\kappa' \frac{P(\kappa'|s_i)}{P_{bulk}(\kappa')} h_\theta(b_r, b_s). \tag{3.50}$$

$$\gamma = \cos^{-1}\left(\frac{-b_r \cdot b_s}{|b_r||b_s|}\right) \tag{3.51}$$

$$h_\theta(b_r, b_s) = \begin{cases} 1 & 0 \le \gamma \le \gamma_{max} \\ 0 & o.w. \end{cases}. \tag{3.52}$$

If $P(\kappa|s_i)$ is constant over the window of γ, then

$$\frac{W_D}{W_w} = \frac{a_w}{a_D}, \tag{3.53}$$

and so the direct match and matched window methods are equivalent for small windows or regions of nearly flat $P(\kappa|s_i)$.

Harmonic bond angle mismatch

It may be more physically correct to model the bonding with an energy that is harmonic in γ, leading to an additional Boltzmann factor,

$$E_H(\gamma) = -\frac{1}{2} k_H \gamma^2 \tag{3.54}$$

$$W_H(\kappa_i|s_i) = a_H \int d\kappa' \frac{P(\kappa'|s_i)}{P_{bulk}(\kappa')} e^{-\beta E_H(\gamma)}. \tag{3.55}$$

The bonding tolerance k_H can be made asymmetric for hydrogen bond donors and hydrogen bond acceptors.

The numerical γ problem

The calculation for γ is simple in the MB water model which, for bond vector at angle θ to $\hat{x}$, forms a bond angle γ with a neighbor at s_i

$$\gamma_{MB} = \theta - (2\pi - \theta_i). \tag{3.56}$$

Conversely, in three dimensions, angles are binned according to $\cos(\theta)$, which unfortunately makes sampling very poor near the poles, $\cos(\theta) = \pm 1$. This problem arises because for constant spacing in $\cos(\theta)$, the stride in θ becomes disproportionately small at the equator and large near the poles. Neighbor angles are well sampled near the equator but unavoidably large bin widths lead to rounding errors near the poles. Expressed in discrete terms,

$$W_H(\kappa|s_i) = a_H \sum_{\ell} w(\theta_\ell) \frac{P(\cos(\theta_\ell|s_i))}{P_{bulk}(\cos(\theta_\ell))} e^{-\beta E_H(\gamma_\ell)} \tag{3.57}$$

where ℓ iterates through the relevant bins near the matching angle and $w(\theta)$ is the discretized bin width in terms of θ rather than $\cos(\theta)$. In practice, the transformation back to θ makes bin widths that differ by a factor of 30 or more. Increasing the number of bins increases sampling, but the number of bins added near the equator grows much faster than the number of bins near the poles, so the computational scaling is very poor.

To address this, γ needs to be defined in a more numerically friendly way. Take two water molecules, as in Figure 3.1, whose bond vectors $\vec{b}$ (blue) are not in perfect alignment, and form an angle γ. Computationally, the tagged molecule may be manipulated, but only the neighbor declination angle θ is accessible through $P(\cos(\theta)|r)$.

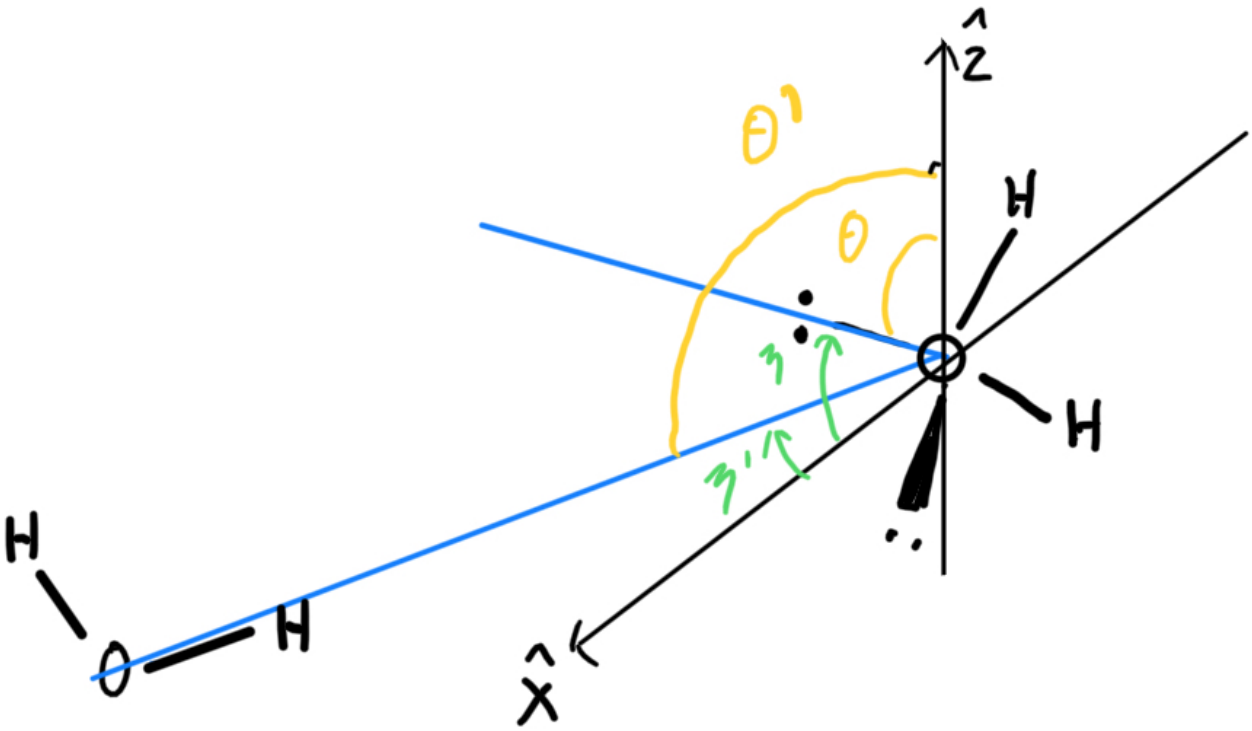

Figure 3.1: Schematic of the angle γ defined between two water molecules.

To compute γ, define (as in Figure 3.1) the declination angles $\theta,\ \theta'$ and the rotations about $\hat{z}$, $\eta,\ \eta'$, where $'$ denotes quantities relating to the neighbor water.

$$\cos(\gamma) = \frac{\vec{b}\cdot\vec{b'}}{|\vec{b}||\vec{b'}|} \tag{3.58}$$

$$= \frac{1}{|\vec{b}||\vec{b'}|}\left[\sin(\theta)\sin(\eta)\sin(\theta')\sin(\eta') + \sin(\theta)\cos(\eta)\sin(\theta')\cos(\eta') + \cos(\theta)\cos(\theta')\right] \tag{3.59}$$

Without loss of generality, define the axes along the flat interface so that $\eta = 0$ and set the bond vectors $\vec{b}$ to unit magnitude,

$$\cos(\gamma) = (0) + \sin(\theta)\sin(\theta')\cos(\eta') + \cos(\theta)\cos(\theta'). \tag{3.60}$$

A convenient energy function resembles the harmonic energy $E_H(\gamma)$ with a shift for numerical stability,

$$E_\gamma(\cos(\gamma)) = k_\gamma(1 + \cos(\gamma)) \tag{3.61}$$

The new expression for the immediate probability of an orientation is then

$$W_\gamma(\cos(\theta)|s_i) = \frac{a_\gamma}{Z_\gamma}\int_{-1}^{1} d\cos(\theta')\int_0^{2\pi} d\eta' \frac{P(\cos(\theta')|s_i)}{P_{bulk}(\cos(\theta'))} e^{-\beta k_\gamma} e^{-\beta k_\gamma[\cos(\theta)\cos(\theta')+\sin(\theta)\sin(\theta')\cos(\eta')]} \tag{3.62}$$

$$Z_\gamma = \int_{-1}^{1} d\cos(\theta')\int_0^{2\pi} d\eta' e^{-\beta k_\gamma} e^{-\beta k_\gamma \cos(\gamma)}. \tag{3.63}$$

The integration over η' can be simplified by using the order 0 Bessel function of the first kind,

$$I_0(x) = \frac{1}{2\pi}\int_0^{2\pi} e^{x\cos(\eta)} d\eta, \tag{3.64}$$

$$W_\gamma(\cos(\theta)|s_i) = \frac{2\pi a_\gamma}{Z_\gamma}\int_{-1}^{1} d\cos(\theta')\frac{P(\cos(\theta')|s_i)}{P_{bulk}(\cos(\theta'))} e^{-\beta k_\gamma} e^{-\beta k_\gamma \cos(\theta)\cos(\theta')} I_0(-\beta k_\gamma \sin(\theta)\sin(\theta')), \tag{3.65}$$

$$\approx \frac{2\pi a_\gamma}{Z_\gamma}\sum_{\cos(\theta')=-1}^{1} \Delta\cos(\theta')\frac{P(\cos(\theta')|s_i)}{P_{bulk}(\cos(\theta'))} e^{-\beta k_\gamma} e^{-\beta k_\gamma \cos(\theta)\cos(\theta')} I_0(-\beta k_\gamma \sin(\theta)\sin(\theta')) \tag{3.66}$$

This expression can be implemented numerically and does not require the aggressive bin-width corrections that the previous harmonic energy required.

The role of each term in Equation 3.65 can be understood as follows:

- $P(\cos(\theta'))$ is the probability of finding a neighbor with declination angle $\cos(\theta)$
- $e^{-\beta k_\gamma \cos(\theta)\cos(\theta')}$ is the portion of the Boltzmann factor that measures the degree that θ and θ' match.
- $I_0(-\beta k_\gamma \sin(\theta)\sin(\theta'))$ summarizes all the Boltzmann factors deriving from η, i.e. the mismatch to the neighbor's rotation about $\hat{z}$. It is fortunate that $P(\eta')$ is constant by symmetry about a flat interface. The Bessel factor makes essentially a geometric correction for changing importance of η; the η mismatch could be severe when θ is equatorial, but is negligible when θ is polar.

For small k_γ, Equation 3.65 works well. For large k_γ, the chosen Bessel function is still not necessarily numerically stable. One more approximation is made where for large arguments x,

$$I_0(x) \approx \frac{e^x}{\sqrt{2\pi x}} \tag{3.67}$$

which substitutes in combination with the identity

$$\cos(a+b) = \cos(a)\cos(b) + \sin(a)\sin(b) \tag{3.68}$$

so that for large k_γ, W_γ is approximated by

$$W_\gamma(\cos(\theta)|s_i) \approx \frac{2\pi a_\gamma}{Z_\gamma} \int_{-1}^{1} d\cos(\theta') \frac{P(\cos(\theta')|s_i)}{P_{bulk}(\cos(\theta'))} \frac{e^{-\beta k_\gamma(\cos(\theta+\theta')+1)}}{(-2\pi\beta k_\gamma \sin(\theta)\sin(\theta'))^{\frac{1}{2}}}. \tag{3.69}$$

Inspection of the exponential shows that W_γ is greatest when the bond vectors are exactly anti-parallel on the declination coordinate θ, i.e. $\theta + \theta' = \pi$.

3.3 Mean field surface of MB water

For the 2-dimensional MB water model (see Section 2.3), orientations are simple to manipulate and express only in the angle θ, as

$$P(\cos(\theta)|d) = \frac{1}{Z(d)} \sum_{\alpha} \prod_{i=1}^{m_\alpha} e^{-\beta\epsilon_{MB}} a\phi(s_i) W_{MB}(\Delta\theta_i|s_i) \prod_{j=m_\alpha+1}^{M} 1 - a\phi(s_j) W_{MB}(\Delta\theta_j|s_j). \tag{3.70}$$

The bond-match function W_{MB} uses the Gaussian energy form of the MB water bonding, Equation 2.11. Because the mean field surface sets the depth and orientation of the tagged molecule, the MB energy can be abbreviated and shifted,

$$E_{MB}(\Delta\theta) = \epsilon_{HB}[1 - G_{MB}(\cos(\Delta\theta) - 1)], \tag{3.71}$$

$$W_{MB}(\Delta\theta_i|s_i) = \int d\cos(\theta) P(\cos(\theta_i)|d) e^{-\beta E_{MB}(\Delta\theta_i)}, \tag{3.72}$$

using $G_{MB}(x)$, the Gaussian function used for the MB water model bonding, written in Equation 2.14, and $\Delta\theta_i = \theta - \theta_i$.

Monte Carlo simulation of MB water results in an orientational probability heat map repeated here in Figure 3.3 from Figure 2.27. Using the measured density profile, Figure 3.2 is an impressive match with the simulation measurements. While not every feature is a perfect match and there is some distortion, the distances, patterns, and even relative heights are predicted by the self-consistent mean field model.

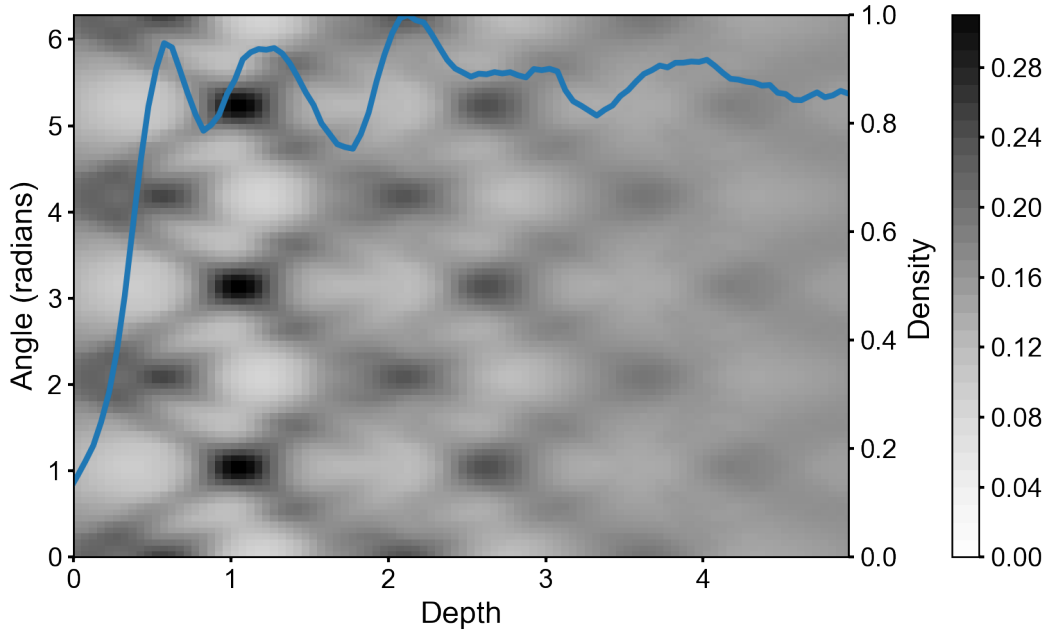

Figure 3.2: Mean field surface results for MB water with a density profile taken from a hard wall simulation. The predicted orientations largely match simulation measurements.

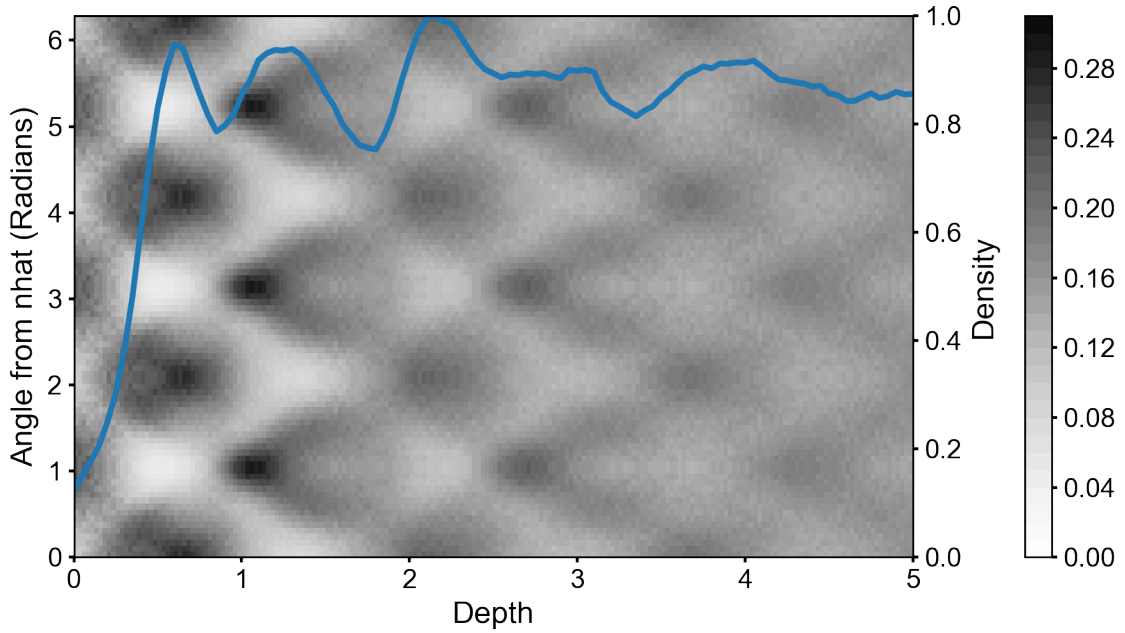

Figure 3.3: MB water orientational trends measured from Monte Carlo simulation.

About the fit parameters

Table 3.1 summarizes the parameters used in the various mean field surface computations of the MB water model. For every change in the form of ϕ or W, the parameters a and ϵ_W also must shift to preserve the scale of free energy. Although it would be reasonable to use a fitting algorithm to choose a and ϵ_W to match the Monte Carlo measurements, the goal of the following sections is to understand cause and effect, not to make quantitative predictions. Accordingly, the fitting algorithm is "by hand".

It would be convenient to set $\epsilon_W = \epsilon_{MB} = 1$ because the hydrogen bond of the MB water model is the unit of energy. However, ϵ_W is used to approximate $\langle \Delta V \rangle$, which contains the average non-isotropic energy contribution from all configurations. The non-isotropic energy is left generic and may be a combination of the hydrogen bond energy ϵ_{HB} and the Lennard-Jones energy ϵ_{LJ}, and would necessarily change with each bonding state α. As a result, while it is expected that ϵ_W resembles the MB water model ϵ_{HB}, there should be no expectation that the energies will match.

In practice, the energy ϵ_W and the parameter a overlap almost entirely, and are interchangeable when considering the m_α successful bonds. The importance of non-bonding neighbors, however, depends entirely on a. As a result, the scale of ϵ_W controls the first layer orientations, where the number of bonds can fluctuate wildly, while a controls the importance of neighbor alignment, which is most important toward bulk. In other words, the first layer peaks are mainly tied to ϵ_W while the peaks at greater depths will grow with a.

Table 3.1: Summary of Mean Field Surface MB water parameters and their useful ranges.

Parameter	Value
a	$0.0005 - 0.01$
ϵ_W	$0.5 - 1.5\ k_B$
Bonding sites, M	3
Temperature, T	0.14 to 0.18

Modifications to the MB mean field surface model

The mean field surface model can be adapted to probe how changing a given physical constraint leads to a change in surface structure.

Density modifications

Figure 3.4 is the mean field surface output when the measured density profile is replaced by a smooth sigmoidal function,

$$\rho_\Sigma(z) = \frac{1}{2}\left[\tanh\left(\frac{d-0.5}{\sigma_\Sigma}\right) + 1\right], \tag{3.73}$$

where $\sigma_\Sigma = 0.2$ roughly matches the thickness of the measured density profile (one molecular length). The surface patterns change, but also largely match with simulation measurements. While the details should depend on the fine density layering, a smooth vapor boundary of approximately one molecular length predicts the main surface patterns.

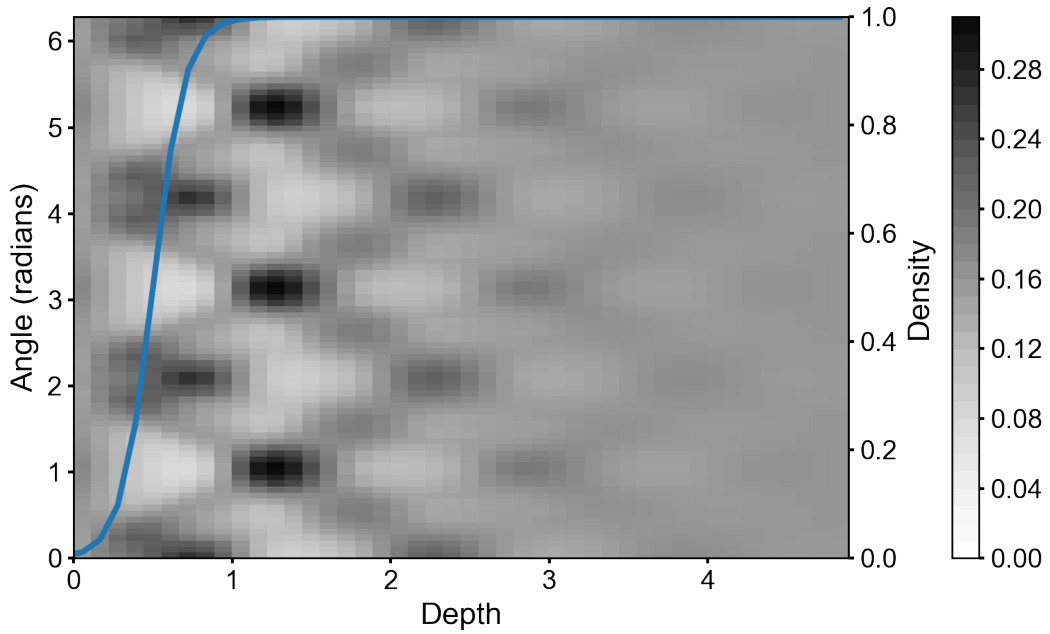

Figure 3.4: Mean field surface calculation for a general sigmoidal interface in two dimensions. The predictions largely match with Monte Carlo simulation measurements, so the density layering cannot be the underlying cause of water surface structure in this case. For MB water, trends in orientation are almost entirely due to distance-dependent correlation with the first molecular layer, which itself maximizes bonding by sacrificing one bond toward vapor.

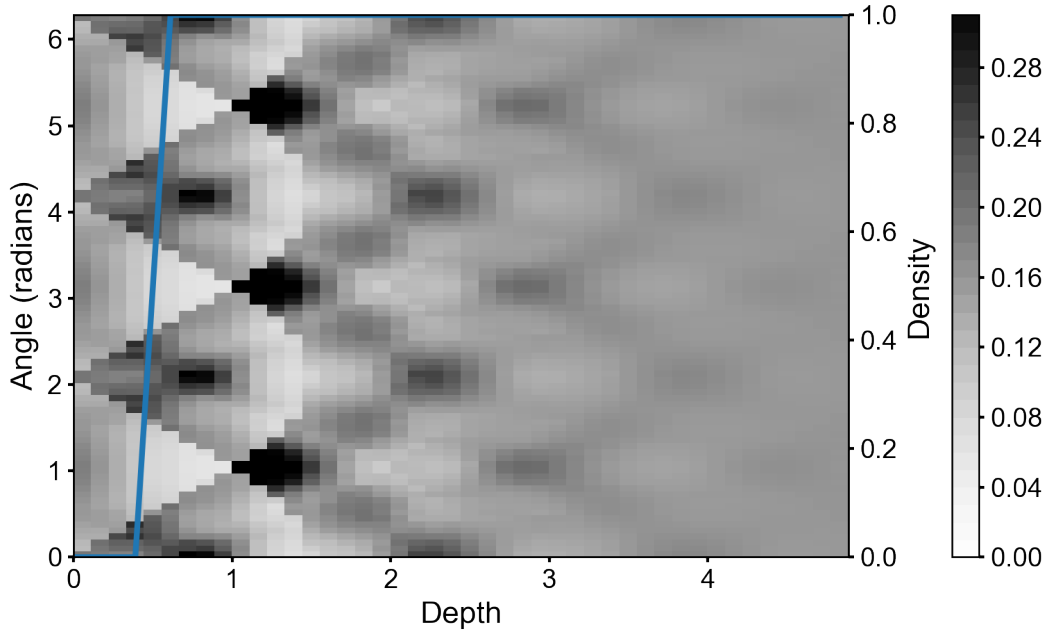

Figure 3.5: Mean field surface calculation for a very sharp interface. While the patterns schematically match those in Figure 3.3, the interfacial sharpness leads to unreasonably sudden changes to orientational probability.

Even a very sharp change in density brings about the same 3-fold symmetric directional oscillations, although the sudden change in $\rho(d)$ leads to exaggerated, sharp features. Figure 3.5 is generated with $\sigma_\Sigma = 0.01$. The density layering cannot be the main cause of orientations at any depth, though it is likely secondary. The most important orientational trends are echoes of the topmost molecular layer rotating to point as many bonds as possible toward the liquid phase.

Without neighbor orientations

Removing the dependence on neighbor orientations is equivalent to setting $W(\kappa_i|s_i) = \Omega^{-1}$. Doing so means only the density plays any role in predicting the neighbor orientations. Figure 3.6 removes W and reveals how the initial, top-layer interfacial patterning can form, but there is no mechanism to capture the diminishing oscillations into bulk. The difference is even more severe when a sigmoidal density is used, see Figure 3.7.

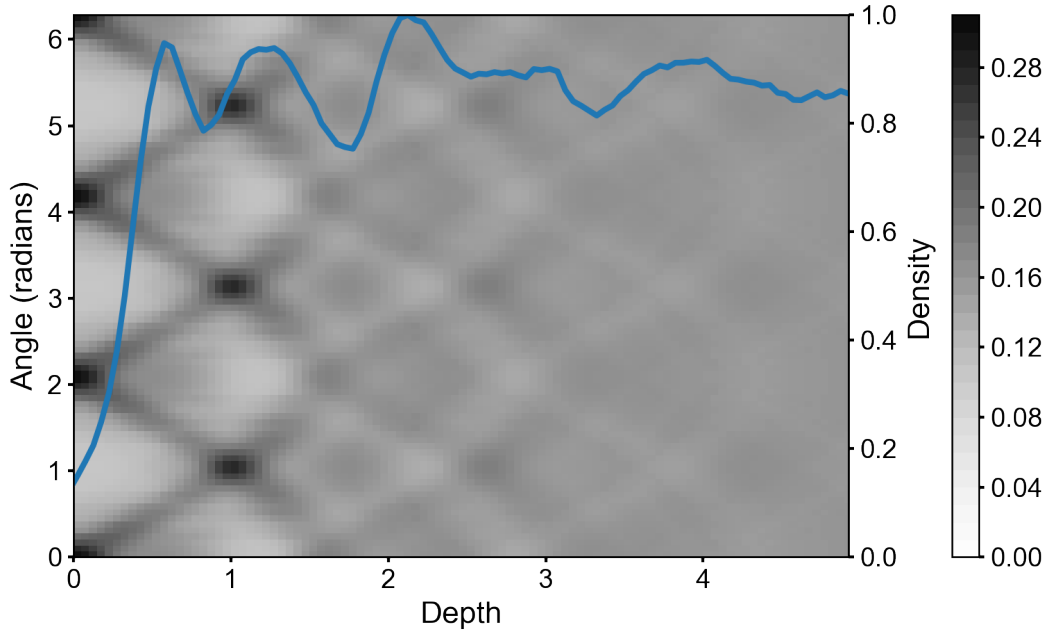

Figure 3.6: Mean field surface calculation when neighbor orientations are ignored. While the first interfacial layer agrees on the principle orientations, there is no opportunity within this model for orientational correlations to extend into bulk without explicitly referencing neighbor positioning.

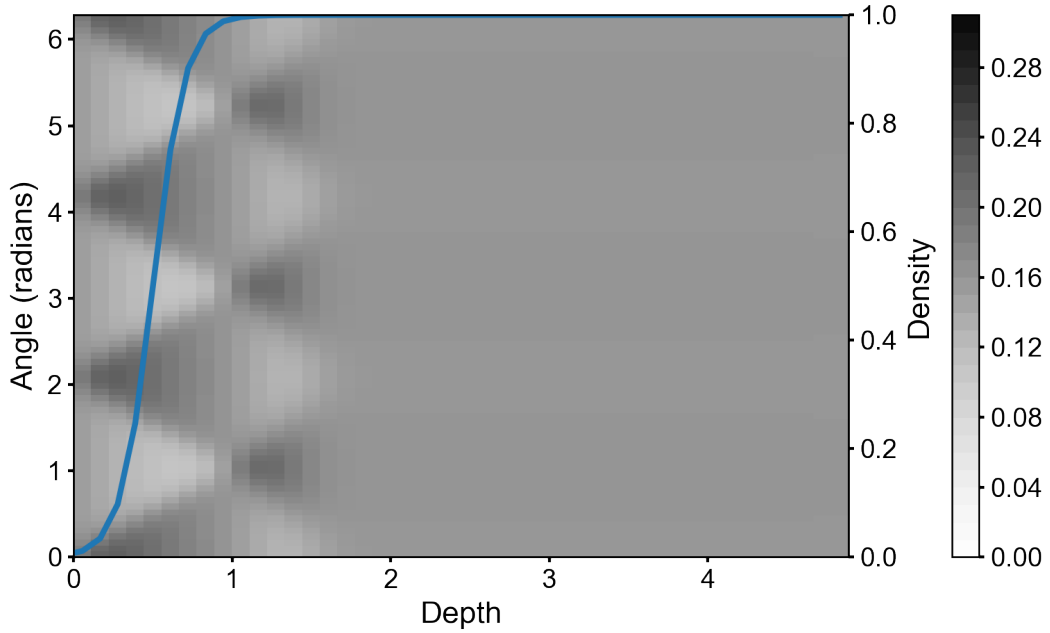

Figure 3.7: Mean field surface calculation when neighbor orientations are ignored and density layering is replaced with a smooth sigmoid. Only the surface-most orientations can be predicted without neighbor orientational information, and positioning here is due entirely to the first layer of molecules maximizing bonding partners.

Loose or strict bonding enforcement

The MB water model specifies a conventional bonding energy, Equation 2.11, which can be used directly in computing the mean field surface model, Equation 3.71. The effect of other bonding rules on the interfacial structure can hypothetically be tested by modifying the mean field surface bonding rules. Figure 3.8 shows the subtle changes to surface features if W_w (Equation 3.51) is used. Compare to Figures 3.3 and 3.2. The positions of the most important peaks do not move, but a narrow bond-angle window makes many features correspondingly too narrow. The tolerance for neighbor bonding angle is an important detail in modeling surface layers. However, once the width of bonding is roughly matched, the underlying potential of a neighbor bond angle is not critical.

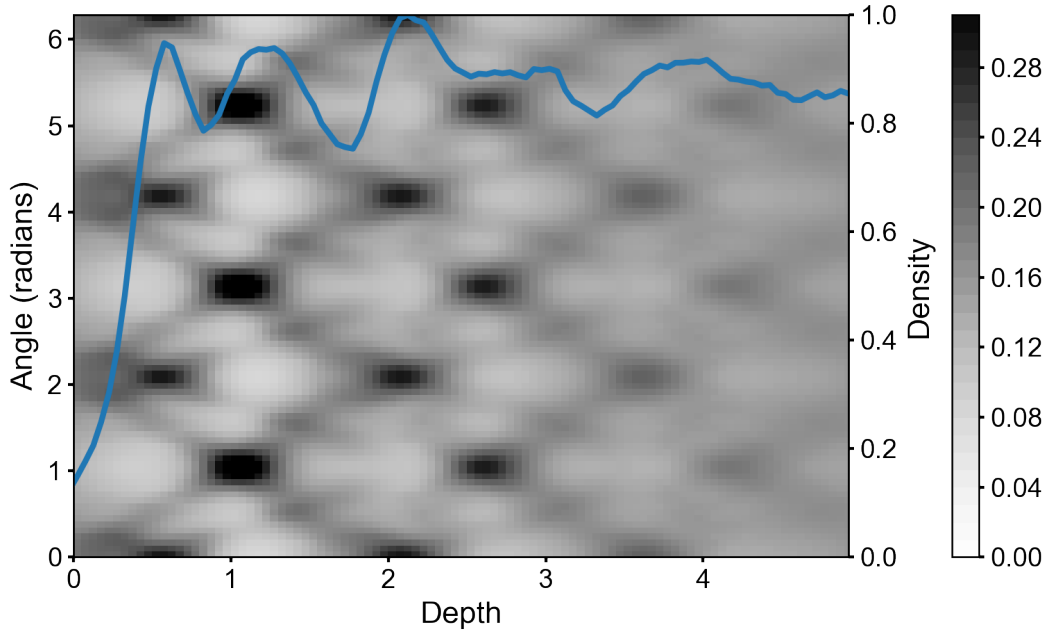

Figure 3.8: Mean field surface calculation when the flat-window W_w scheme is used to match bonding angle. When the window of allowed bonding angle is too narrow–as is the case here– the peaks are also correspondingly too narrow. A wider window would have the opposite problem, with peaks bleeding over one another.

3.4 The mean field surface of water

For the 3-dimensional mean field surface water model, the central equation for $P(\kappa|r)$ is either of the two equivalent forms in Equation 3.40 or 3.41. The following heat maps of probability summarize the output under various conditions. Compare the mean field surface results to the similar outputs for the mean field surface MB water, and for the simulation measurements for SPC/E and TIP5P in Figures 2.10 to 2.16.

Unlike MB water, the orientation of a 3-dimensional water molecule cannot be fully described with a single angle θ. The two OH bonds are indistinguishable and the lone pairs may or may not be explicitly represented in a water model. The water dipole, which points between the two OH bonds, is also important because the direction of the water dipole has implications for water's response to an electric charge or electric field. Figures 3.9, 3.10, and 3.11 show directional predictions for water OH, lone pair (LP), and dipole (mu), respectively.

Figures 3.9-3.11 are generated using a tetrahedron ($M = 4$), assume a unit energy per bond ϵ_W (Equation 3.41), enforce the bonding angle with W_γ (Equation 3.65), have a separate k_γ for donors and acceptors, and constrain bonds so that only OH+LP bonds are favorable.

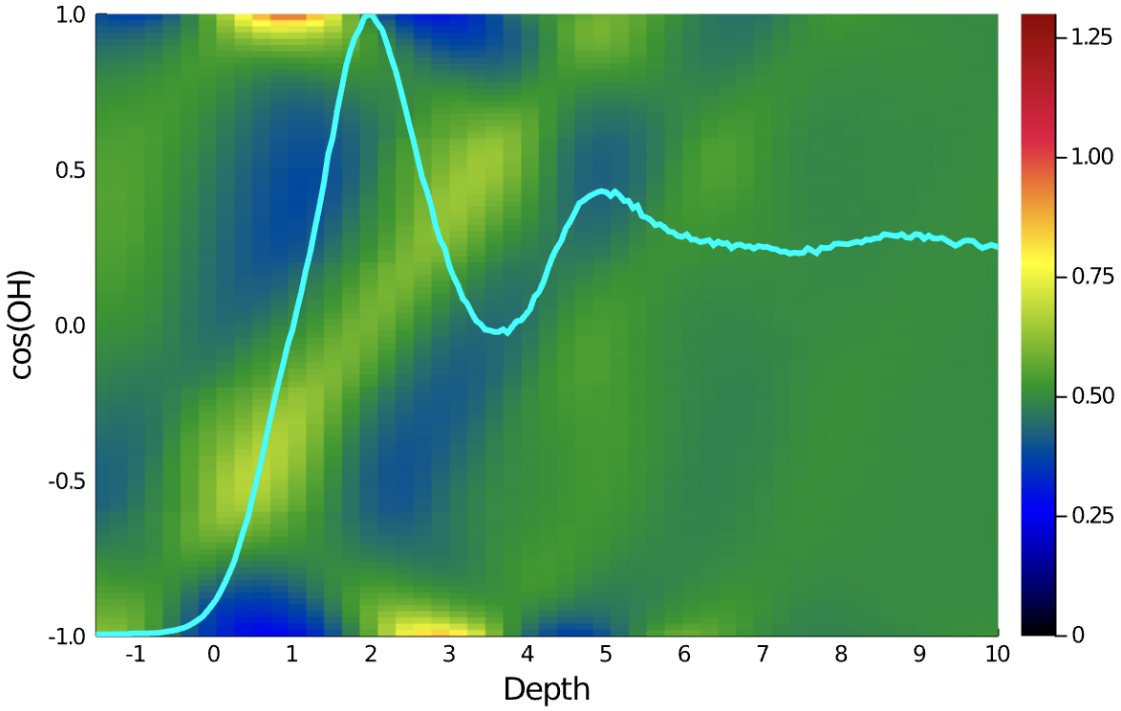

Figure 3.9: Mean field surface prediction for water OH bond direction.

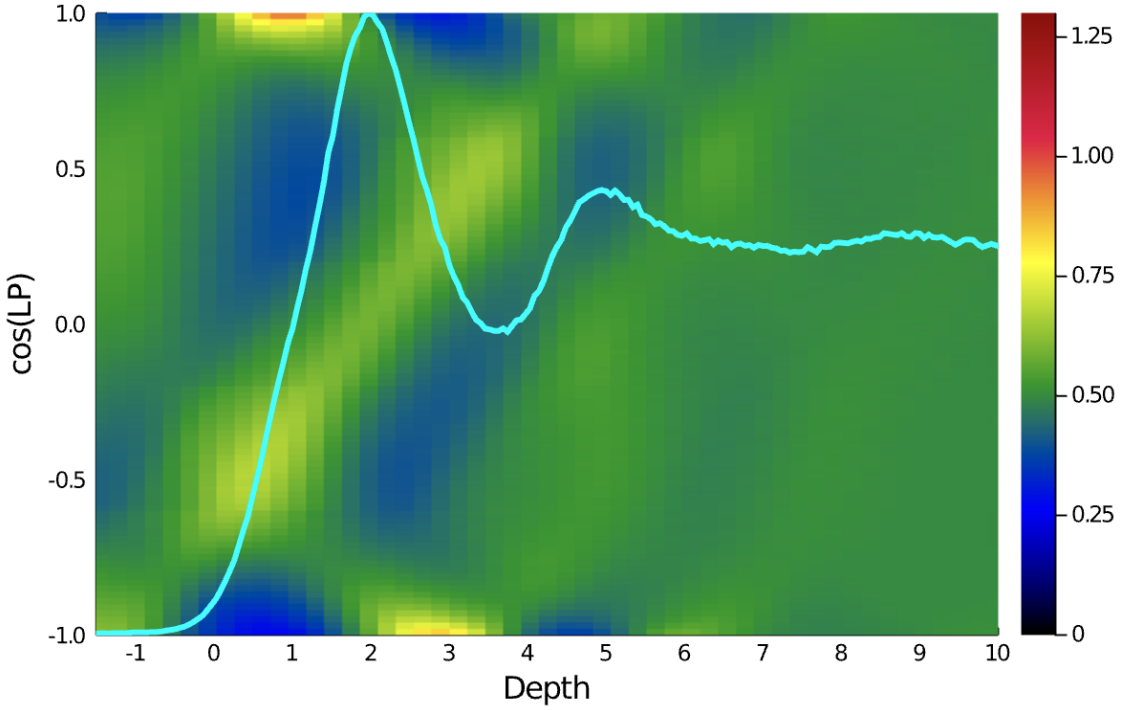

Figure 3.10: Mean field surface prediction for water LP direction.

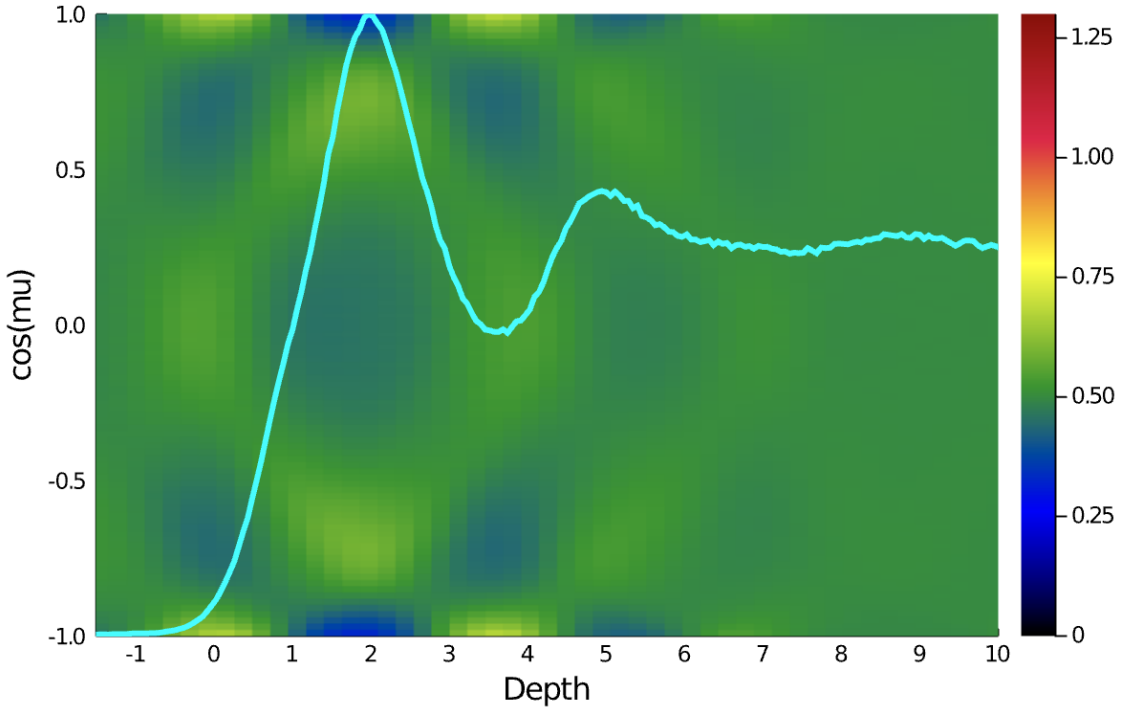

Figure 3.11: Mean field surface prediction for water dipole direction.

In general, predictions of the mean field surface match MD simulation measurements, especially for SPC/E. There is a surprisingly poor match to SPC/E dipole, which highlights one of the largest shortcomings of this model. Given the strong match between OH and LP geometries, it is very surprising that the dipole near the surface is poorly predicted. There must be extended dipole interactions across water's surface that is not captured by the density mean field and immediate neighbor orientations. It is also likely that non-tetrahedral defects play a significant role in the final dipole structure.

Modifications to the water mean field surface model

As with the 2-dimensional MB water model in Section 3.3, modifications to the water mean field surface model can distinguish the role of various pieces of surface water interactions.

Density modifications

As in Section 3.3, the 3-dimensional density profile can be substituted with a monotonic sigmoid, which does not exhibit any layering.

In this case, a sigmoidal density profile leads to a significant change in the orientational outputs. While the main surface features are still vaguely present, the information quickly

washes out toward bulk. Even the surface features change substantially. After multiple self-referencing iterations, the structure in Figure 3.12 remains essentially unchanged, implying that neighbor interactions are small and unimportant in this formulation, unlike what was observed in 2-dimensions. In the case of 3-dimensional water, it appears that the details of layering are important to predicting surface behavior.

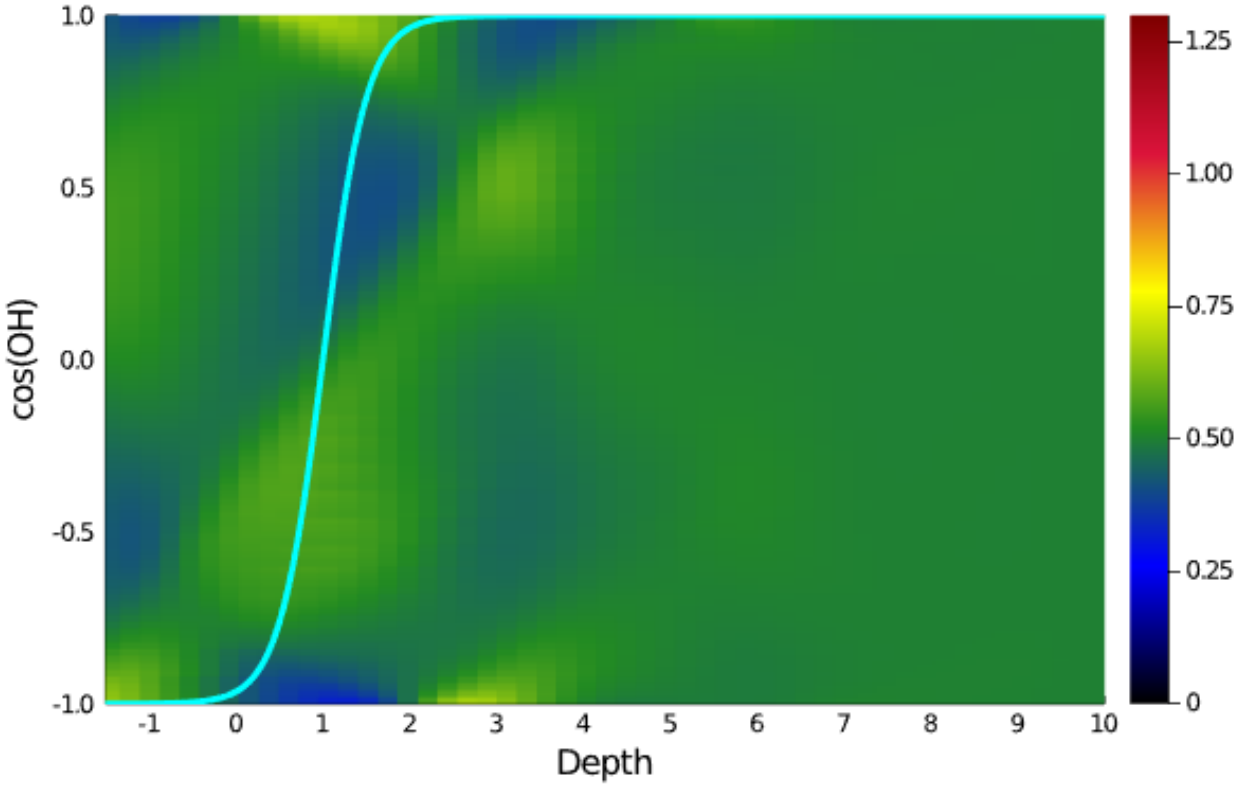

Figure 3.12: Mean field surface calculation for water given a sigmoidal density. Water's surface layering is apparently an important component in determining water orientations.

Varying neighbor orientations

Similar to the 2-dimensional case in Section 3.3, interactions between neighbors can be hypothetically turned off. In the 3-dimensional case, the neighbor interactions are not strictly as essential as the layering.

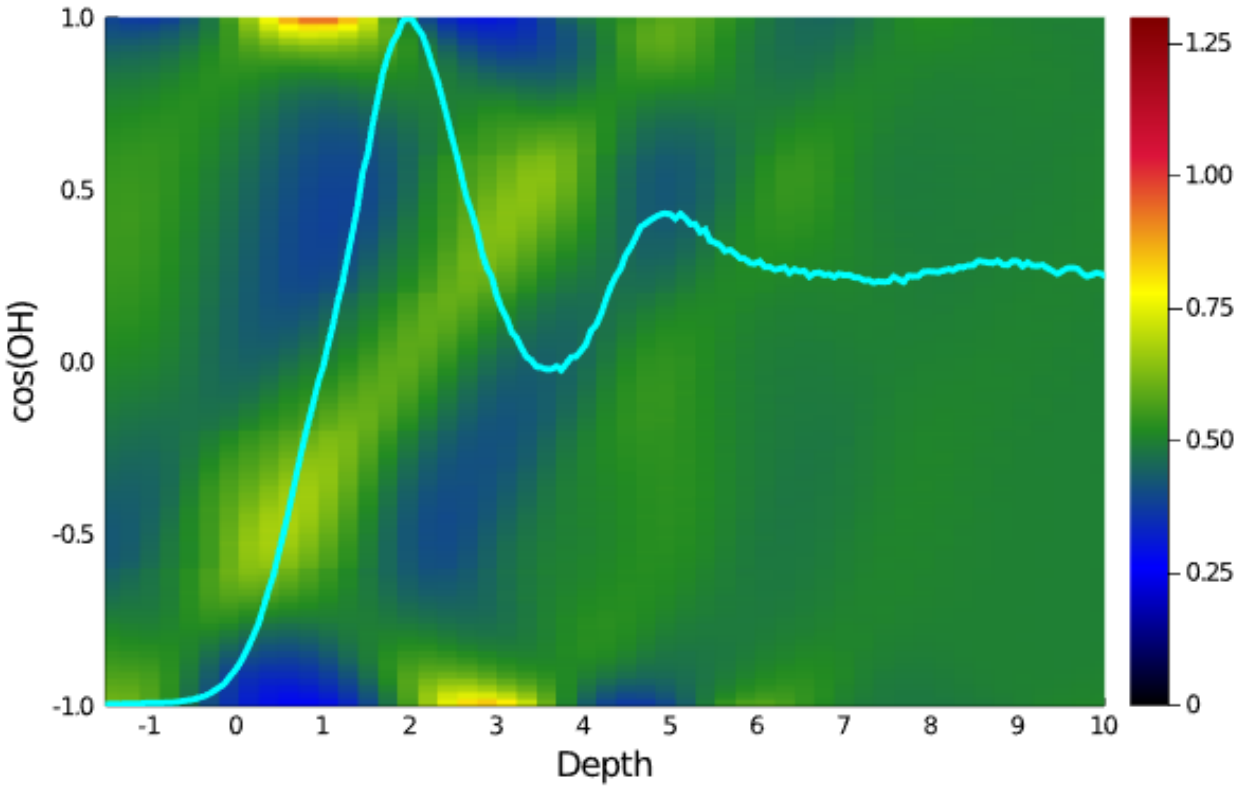

Figure 3.13: Mean field surface calculation for water without explicitly referencing neighbor orientations.

Likewise, the bonding enforcement can be based off of an angular width and may ignore the details of energy, as in Section 3.3. Because there is no significant difference, no plot is included here.

3.5 Dipole trends

The mean field surface model requires some level of asymmetry to predict dipole trends. Because of its symmetric geometry, MB water does not have a meaningful way of modeling dipole behavior. Water does have a dipole, but a fully symmetric tetrahedral model predicts a symmetric dipole structure.

The dipole distribution in 3.11 is used to compute the expectation value at each depth in Figure 3.14,

$$\langle m_z \rangle (d) = \int_{-1}^{1} d\cos(\theta_\mu)\cos(\theta_\mu) P(\cos(\theta_\mu)|d). \tag{3.74}$$

As used here, $m_z = \mu_z/|\mu|$ is the dimensionless dipole vector component aligned to the surface normal.

If the OH and LP vectors are not distinguishable, as in a symmetric tetrahedron, then there is no way to distinguish μ from $-\mu$ and the expectation is trivially

$$\langle m_z \rangle = 0. \tag{3.75}$$

The dipole variance is computed

$$\left\langle (\delta m_z)^2 \right\rangle (d) = \int_{-1}^{1} d\cos(\theta_\mu) \cos^2(\theta_\mu) P(\cos(\theta_\mu)|d) - \langle m_z \rangle^2 (d). \tag{3.76}$$

In the case of a dipole diffusing in an isotropic environment, the variance is ideally

$$\langle m_z^2 \rangle = \frac{\cos^3(\theta)}{3\Omega}\Big|_{\cos(\theta)=-1}^{\cos(\theta)=1} = \frac{1}{3} \tag{3.77}$$

$$\int_{-1}^{1} d\cos(\theta) P(\cos(\theta)) = \frac{1}{\Omega} \int_{-1}^{1} d\cos(\theta) = 1. \tag{3.78}$$

It is unexpected that despite measurable (albeit symmetric) dipole structure, which is qualitatively similar to the asymmetric case in Figure 3.11, the dipole variance of a rigid tetrahedron within this surface model does not differ significantly from the ideal case, $\langle m_z^2 \rangle = \frac{1}{3}$.

To create asymmetry, donor and acceptor bonding flexibility use the respective spring constants $k_{\gamma,d} = 1$ and acceptor $k_{\gamma,a} = 1000$, which is analogous to SPC/E by requiring a near perfect match along the hydrogen direction, but allowing a very wide mismatch for bonding around the lone pairs. The resulting dipole prediction, Figure 3.14, differs substantially from the dipole in SPC/E measurements, Figures 2.23 and 2.24.

Accordingly, it seems clear that a mismatch in donor/acceptor bonding flexibility is not sufficient to predict surface dipole behavior. Possibly, the mean field approximation averages through important higher-order dipole behavior of the surface-most waters. Three-body bonding interactions, like those proposed by Shin and Willard[21], may also be central to dipole trends. Non-tetrahedral surface defects in water could explain the shortcomings of the largely tetrahedral mean field surface model.

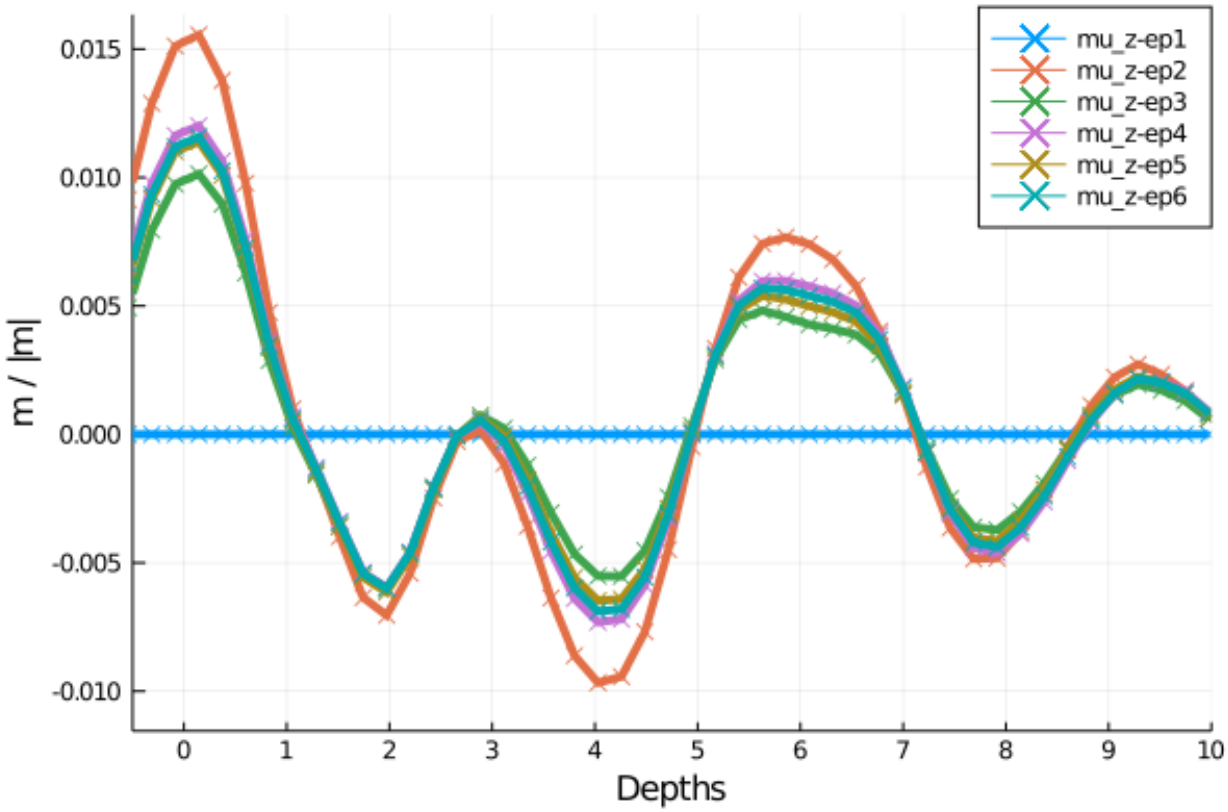

Figure 3.14: Mean field surface dipole predictions across iterations (ep1 to ep6) converge to a series of dipole values different from simulation measurements.

Summary

The mean field surface model of water can make predictions about water's surface based off only a few critical assumptions about bonding behavior. Hydrogen bonding is the leading reason for water interfacial structure, as the top-most water molecules orient to maximize the number of bonds. Reasonable changes to bonding flexibility and bond asymmetry make almost no change to the model predictions. The principles of the mean field surface model can be applied to other strongly geometric liquid surfaces. The model breaks down when predicting higher-order neighbor effects that are averaged out in the mean field, notably in failing to predict dipole behavior. The dipole structure at the interface must accordingly be a result of more than a simple maximization of the immediate hydrogen bonding environment.

Chapter 4

Dielectric Response of Surface Water

Overview

Dielectric continuum theory (DCT) is a method to simplify complex electrostatic information into a mathematically accessible model of material polarization that is continuous in space and valid at a range of length scales. Atomistic molecular models often use atomic partial charges, but these are already a simplification of the more complicated quantum mechanical behavior of electrons around atomic nuclei. DCT simplifies charge interactions yet another step by averaging most charges into a dielectric field that is smoother and easier to manage than point-wise atomic partial charges. DCT is a macroscopic model that averages atomic-scale information, and so necessarily omits details at small scales. Conscious to the loss of molecular detail, DCT can make limited predictions at length scales corresponding to a handful of water molecules.

One formulation of DCT for liquids, the Gaussian field model of fluids, hinges upon the correlation between fluid dipoles, $\chi(\boldsymbol{r}, \boldsymbol{r}')$. Large deviations in mean molecular dipole are limited in point charge models of water, which leads to finite size effects. Theoretical predictions that use $\chi(\boldsymbol{r}, \boldsymbol{r}')$ match simulation measurements both in bulk liquid and can be successfully extended to the water interface.

4.1 Motivation for dielectric continuum theory

A homogeneous collection of charges with some freedom to reorganize can be caricatured with a single macroscopic parameter, the dielectric constant ϵ. This constant, sometimes more generally termed the electric permittivity, quantifies the response of a material (the dielectric) to an electric field. For example, a molecular dynamics simulation of liquid water consists of a collection of positive hydrogen charges and negative oxygen charges, all of which will experience forces from an external electric field. While other constraints are present, like molecular bonding and size exclusion, the atom-by-atom charges are able to move in response to an electric field. The rearrangement creates a local field as molecular dipoles

align contrary to and therefore partially attenuate the applied field. The atomistic details of the response can be very complex, but when taken as an average on a macroscopic scale the response to an external electric field is well characterized by the dielectric constant.

DCT addresses molecular simulation needs

Dielectric theory has many applications. Molecular simulations often approximate a liquid as a uniform dielectric to estimate either the fluid's response to an electric field or the free energy of ion solvation.

An external field $\mathcal{E}$ is attenuated by a dielectric material with dielectric constant ϵ, yielding an electric field $\mathcal{E}_{inner}$ inside the material that is a uniform scaling of the external field,

$$\mathcal{E}_{inner} = \frac{1}{\epsilon}\mathcal{E}. \tag{4.1}$$

The free energy of solvation of an ion can be estimated with the Born equation, proposed by Max Born in 1920,

$$\Delta G = -\frac{q^2}{8\pi\epsilon_0 r_0}\left(1 - \frac{1}{\epsilon}\right). \tag{4.2}$$

Equation 4.2 combines the permittivity of free space ϵ_0, the radius of the ion r_0, the charge of the ion q, and the dielectric constant of the solvent ϵ. For water, with $\epsilon_W \approx 80$ at STP, the Born solvation equation implies a strongly favorable free energy of solvation for all atomic and small-molecule ions.

Separately or in combination, these two relations are sometimes used to replace the explicit numerical simulation of a solvent like water. For example, a molecular dynamics simulation with implicit solvent may approximate the interaction of two charged solutes or charged functional groups with Equation 4.1. The solvation energy of a more complicated molecule can be estimated through the use of Equation 4.2. DCT is useful because it simplifies complicated molecular interactions, but at the cost of accuracy and much atomic detail.

The method of images used to describe interfacial ion solvation is a solution to DCT given an interfacial boundary conditions, see ref [4]. More specific use-cases often require modifications to the DCT model, sometimes including measurements from molecular simulations.

Limits of DCT

Given the simplicity and flexibility of a model like DCT, there are many cases when DCT does not perform well as a model. Simple modifications can be made to fix some shortcomings. DCT works best in macroscopic systems where relative fluctuations in average

quantities become small. At atomic lengthscales or the timescales of a molecular simulation, a highly schematic description like the dielectric constant will inevitably break down, but the limitations may be predictable and possible to remedy.

Extremes in electric fields lead to a dielectric breakdown. Water, for example, has a theoretical maximum to its ability to polarize. If an electric field is applied to a slab of water, the slab will respond in a mostly linear way for small to moderate electric fields, as liquid water molecules become increasingly likely to align dipoles with the electric field during regular thermal motion. A rigid water model like SPC/E, however, will eventually align totally with the field and reach a limit of response [35, 36]. The dielectric susceptibility will change substantially as water has nears a "dielectric saturation" point, near $\mathcal{E} = 1$ V/nm for SPC[35], when the water dipoles reach a limit to their collective ability to orient with the field. Rigid water models are unable to further polarize because there is no other mechanism to do so.

Non-rigid models of water can distort oxygen and hydrogen bonding in response to a field, but molecular models in general are not intended to handle such an extreme perturbation. Physical water exposed to a strong electric field hypothetically has three modes of dielectric response, starting with dipole alignment, then shifting to significant molecular distortions, and culminating in ionization of the component oxygen and hydrogen atoms. Accordingly, there are likely three drastically different but valid modes in water's dielectric constant, depending on the strength of a field. In practice, familiar water structure disappears with electric fields exceeding approximately $\mathcal{E} = 1$ V/nm[37, 38], where macroscopic droplets will distort and lose cohesion.

The introduction of non-electrostatic constraints also changes the otherwise uniform dielectric of a liquid. For example, a force which limits water's ability to reorient its dipole will limit water's ability to align to an electric field, thereby reducing the dielectric constant. This constraint could take the form of a solute (probably with an electric field of its own) or a non-polar interface with another substance. A phase boundary, like the water liquid-vapor interface of Chapter 2, could theoretically result in a new constraint that either limits or frees water orientations strongly enough to modify the local dielectric constant.

A significant body of research has attempted to assign a spatially varying dielectric permittivity to water as interfacial interactions begin to constrain water to behave differently than in bulk[10, 28, 30, 39–42]. Chapters 2 and 3 predict that the dielectric behavior is unlikely to be affected more than 10 Å from the surface. The remainder of this chapter attempts to characterize the dielectric permittivity at water's surface by sampling small sub-volumes near the water interface and shows that DCT faithfully describes microscopic polarization at length scales beyond a few molecular diameters, and that rigid water models' polarization limit reduces the fluctuations in dipole that determine the dielectric constant.

4.2 The DCT Gaussian field model of fluids

Understanding the dielectric fluctuations of water

Molecular water is already simplified in classical molecular simulations as a group of 3-5 point charges that are constrained by a combination of bond and angle rules. All rigid point-charge water models can be reasonably expressed as a dipole, $\boldsymbol{\mu}$ that translates and rotates in space with the regular motion of a liquid. The molecular dipole moment of a rigid water model has a constant magnitude, 2.351 Debye for SPC/E.

The interaction energy of a dipole with a field $\mathcal{E}$ is

$$E(\boldsymbol{\mu}) = -\boldsymbol{\mu} \cdot \mathcal{E}. \tag{4.3}$$

In a system of bulk water, a tagged water molecule will see a net electric field from its neighbors, $\mathcal{E}_{neib}(\boldsymbol{r})$. Because bulk liquid is isotropic, there is no net average field in time or space

$$\langle \mathcal{E}_{neib} \rangle = 0 \tag{4.4}$$

but at any moment in time each water molecule will be pushed to orient according to the instantaneous local field created by its environment. The scale of fluctuations in the liquid environment sets the scale of the dielectric constant.

To bridge the scale of individual molecules with DCT, bulk water can be divided into cells, each with total dipole $\boldsymbol{m_r}$, see Figures 4.1 and 4.2. Every cell is able to interact with its neighboring cells which can also be summarized with a local field $\mathcal{E}_{neib}(\boldsymbol{r})$. As cells are made larger, microscopic trends disappear and resemble the macroscopic behavior assumed by DCT. Song, Chandler, and Marcus [1–3] suggest that the cell size should be larger than the correlation length of molecular water-water dipole interactions. According to the fluctuation-dissipation theorem, the dielectric constant is related to the fluctuations of the local dipole,

$$(\epsilon - 1) \sim \langle \boldsymbol{m}_{\boldsymbol{r}}^2 \rangle \,. \tag{4.5}$$

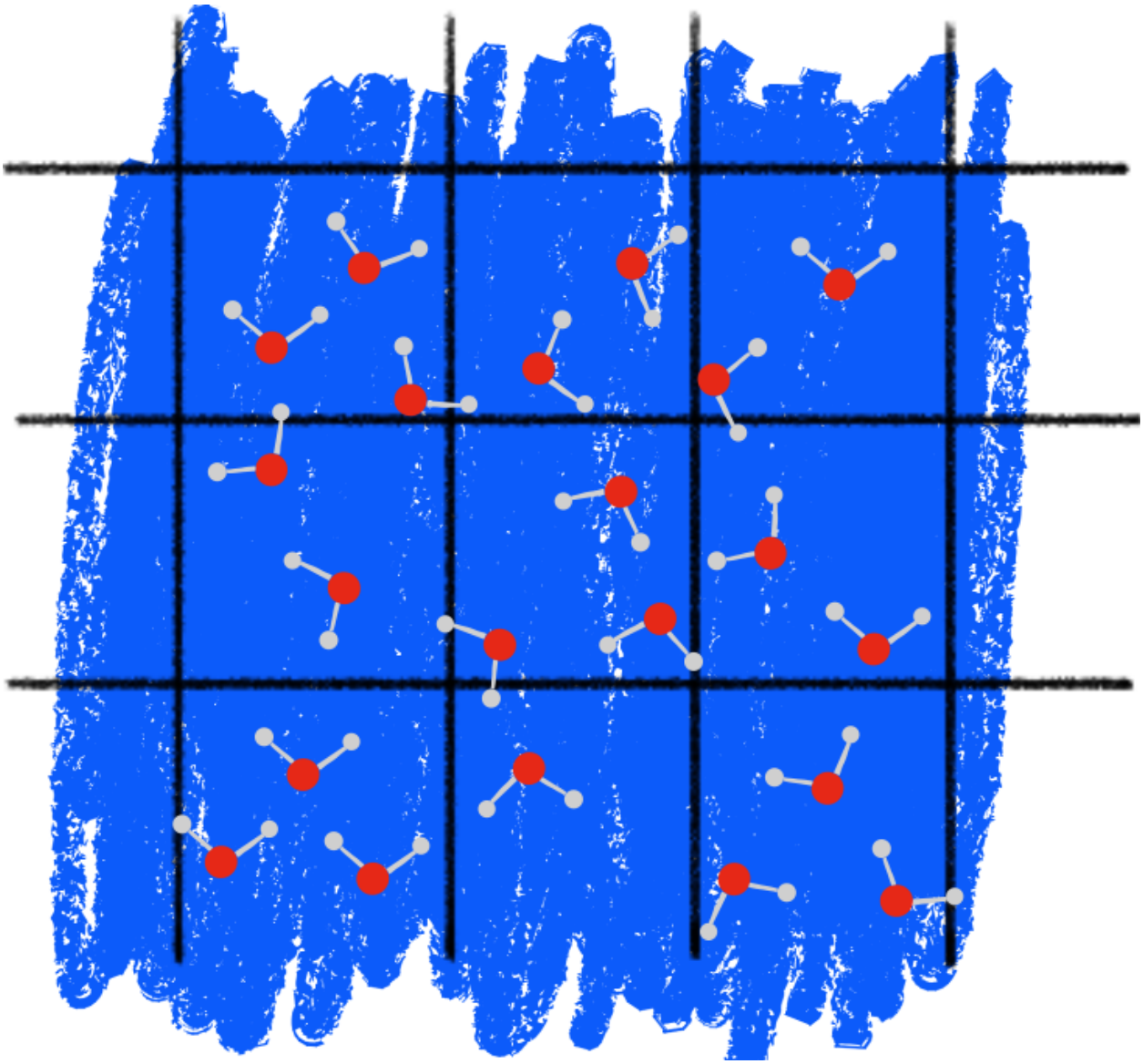

Figure 4.1: Diagram of liquid water with a grid illustrating a division of water molecules into local dipole domains.

Formulation of the theory

Song, Chandler, and Marcus[1–3], building off previous work by Wertheim[43], proposed a Gaussian field model to relate dielectric continuum theory to the polarizability of an appropriately sized cell of water.

According to Gaussian Field Theory, the Hamiltonian of the system takes a form similar

to a set of interacting harmonic oscillators,

$$\beta H = \frac{1}{2}\int d\boldsymbol{r}\int d\boldsymbol{r}'\boldsymbol{m}_r\cdot\boldsymbol{\chi}^{-1}(\boldsymbol{r},\boldsymbol{r}')\cdot\boldsymbol{m}_{r'}. \tag{4.6}$$

Given a cell at $\boldsymbol{r}$ with net dipole $\boldsymbol{m}_r$, the response function describing its correlation with another cell at $\boldsymbol{r}'$ is written[2]

$$\boldsymbol{\chi}(\boldsymbol{r},\boldsymbol{r}') = \frac{\epsilon-1}{4\pi\epsilon}\left[\epsilon\delta(\boldsymbol{r}-\boldsymbol{r}')\boldsymbol{I} - \frac{\epsilon-1}{4\pi}\nabla\nabla'\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\right] \tag{4.7}$$

where $\boldsymbol{I}$ is the identity matrix. $\boldsymbol{m}_r$ absorbs the units of dipole density,

$$\boldsymbol{m}_r = \frac{1}{|\boldsymbol{\mu}|}\int_v d\boldsymbol{r}\boldsymbol{\mu}(\boldsymbol{r}). \tag{4.8}$$

The response function $\boldsymbol{\chi}$ is related to the correlation function through the Fluctuation-Dissipation Theorem,

$$\boldsymbol{\chi} = \langle \boldsymbol{m}_r\boldsymbol{m}_{r'}\rangle, \tag{4.9}$$

which will be measured computationally.

Shaffer[4] offers an in-depth derivation of the above expressions, which are valid for an infinite and continuous dipolar material.

A semi-infinite dielectric

Shaffer[4] adapts the Gaussian Field model to a semi-infinite dielectric, which corresponds to the water liquid-vapor interface. Using the common DCT method of images for a charge near an interface, the response function in Equation 4.7 is modified with an added a term corresponding to the non-uniform coordinate $\boldsymbol{r}'\cdot\hat{z} = z'$,

$$\boldsymbol{\chi}(\boldsymbol{r},\boldsymbol{r}') = \frac{\epsilon-1}{4\pi\epsilon}\left[\epsilon\delta(\boldsymbol{r}-\boldsymbol{r}')\boldsymbol{I} - \frac{\epsilon-1}{4\pi}\nabla\nabla'\left(\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|} + \frac{\epsilon-1}{\epsilon+1}\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'+2z'\hat{z}|}\right)\right]. \tag{4.10}$$

Equation 4.10 can also be written more succinctly

$$\boldsymbol{\chi}(\boldsymbol{r},\boldsymbol{r}') = \frac{\epsilon-1}{4\pi\epsilon}\left[\delta(\boldsymbol{r}-\boldsymbol{r}')\boldsymbol{I} - \nabla'\langle\boldsymbol{m}_r\rangle_{r'}\right], \tag{4.11}$$

where

$$\langle\boldsymbol{m}_r\rangle_{r'} = \frac{\epsilon-1}{4\pi}\mathcal{E}(\boldsymbol{r}) \tag{4.12}$$

$$\mathcal{E}(\boldsymbol{r}) = -\frac{q}{\epsilon}\nabla\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|} - \frac{q_{image}}{\epsilon}\nabla\frac{1}{|\boldsymbol{r}-\boldsymbol{r}_{image}|} \tag{4.13}$$

$$q_{image} = \frac{\epsilon-1}{\epsilon+1}q \tag{4.14}$$

$$\boldsymbol{r}_{image} = \boldsymbol{r}' + 2z'\hat{z}. \tag{4.15}$$

The following sections will attempt to measure the dielectric permittivity of water from simulation as it relates to Equations 4.7 and 4.10.

4.3 Relating DCT to water simulation

The response function of Equation 4.9 characterizes the correlation of dipole density at two points in space. An approximate value can be measured from molecular simulation by taking an average over a small volume. A sub-volume v of water molecules will have a total dipole

$$\boldsymbol{M}_v = \int_v d\boldsymbol{r}\ \boldsymbol{m}_{\boldsymbol{r}}. \tag{4.16}$$

Its correlation with the dipole of another sub-volume v' is a sum of point-wise dipole correlations, see Figure 4.2,

$$\langle (\boldsymbol{M}_v)(\boldsymbol{M}_{v'}) \rangle = \int_v d\boldsymbol{r} \int_{v'} d\boldsymbol{r}' \left\langle \boldsymbol{m}_{\boldsymbol{r}} \boldsymbol{m}_{\boldsymbol{r}'} \right\rangle. \tag{4.17}$$

Since the average dipole of a given cell in isotropic water is zero, the correlation can be also be written

$$\delta \boldsymbol{M}_v = \boldsymbol{M}_v - \langle \boldsymbol{M}_v \rangle \tag{4.18}$$

$$\langle \boldsymbol{M}_v \rangle = 0 \tag{4.19}$$

$$\langle (\delta \boldsymbol{M}_v)(\delta \boldsymbol{M}_{v'}) \rangle = \langle (\boldsymbol{M}_v)(\boldsymbol{M}_{v'}) \rangle \tag{4.20}$$

This covariance can be measured readily with molecular simulation.

For the self-correlation of a single vector component in bulk (the z-component is chosen here), the sub-volume approach to Gaussian field theory implies

$$\langle (\delta M_{v,z})^2 \rangle = \int_v \int_{v'} d\boldsymbol{r} d\boldsymbol{r}' \chi(\boldsymbol{r} - \boldsymbol{r}') \tag{4.21}$$

$$= v \frac{(\epsilon - 1)(2\epsilon - 1)}{3\epsilon}. \tag{4.22}$$

The quantity $\boldsymbol{M}_v(\boldsymbol{r})$ sums the dipoles of water molecules whose center lies inside a cuboid or spherical sub-volume v centered at location $\boldsymbol{r}$. Computationally, there is a trade-off between cell volume, spatial resolution, and finite-size effects. Density correlations stemming from a drastic change in water density span about 10 Å (see Figure 2.2), and electrostatic correlations in response to an electric field are potentially even longer because electrostatic interactions decay as $1/r$. Given these lengthscales, it would seem unlikely that a sub-volume of less than 10 Å would be decorrelated enough from edge-to-edge to provide a meaningful average dipole $\langle \boldsymbol{M}_v \rangle$. Figure 4.3 shows the slow convergence of $\frac{1}{v} \langle (\delta \boldsymbol{M}_v)^2 \rangle$ with increasing v, which deviates noticeably from DCT predictions even for relatively large sub-volumes $v = L^3$. It will be shown that sub-volumes containing only a few water molecules are sufficient to capture most dielectric behavior after correction.

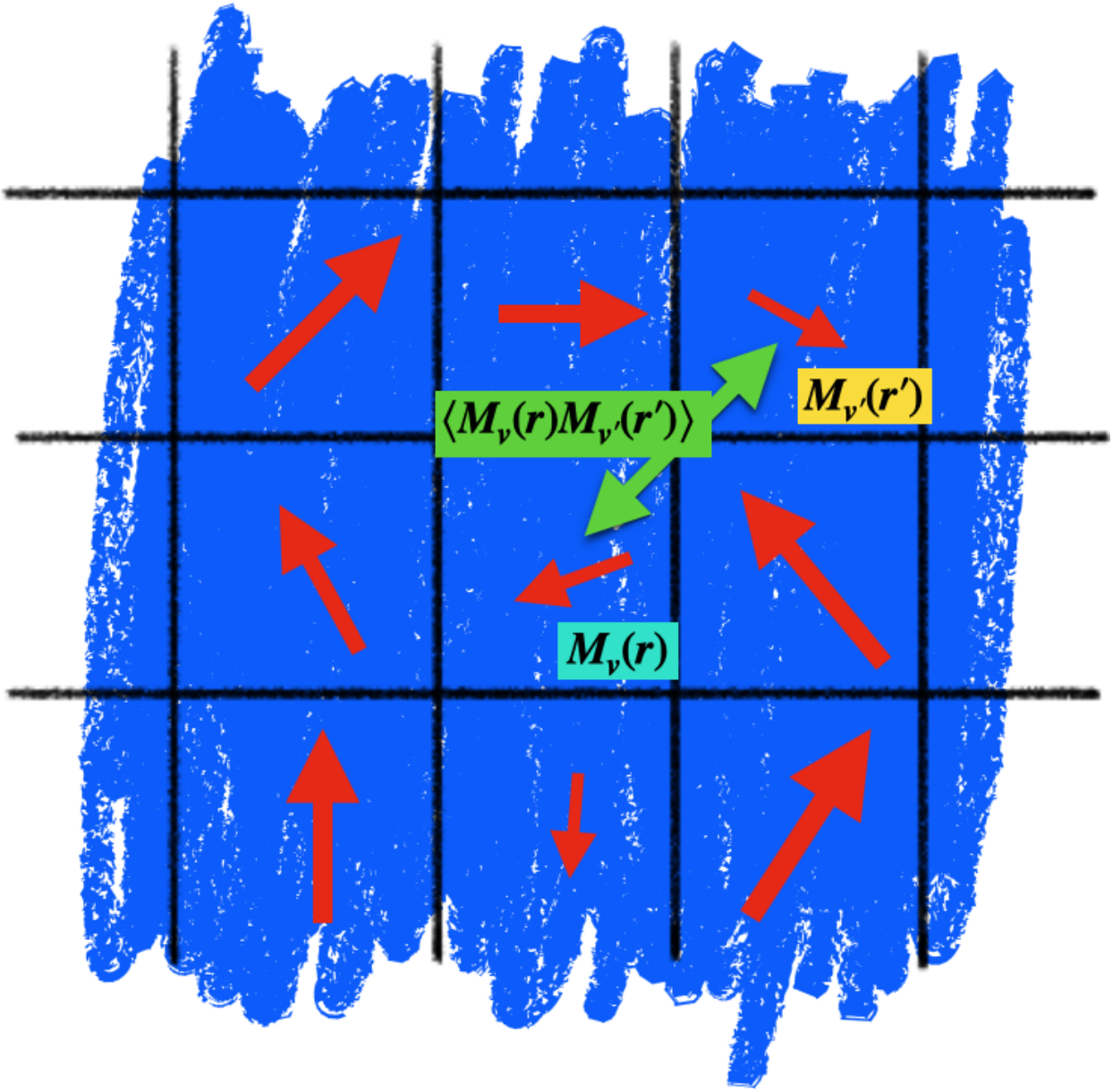

Figure 4.2: Schematic of the correlation between two sub-volumes with dipoles $\boldsymbol{M}_v(\boldsymbol{r})$ and $\boldsymbol{M}_{v'}(\boldsymbol{r}')$.

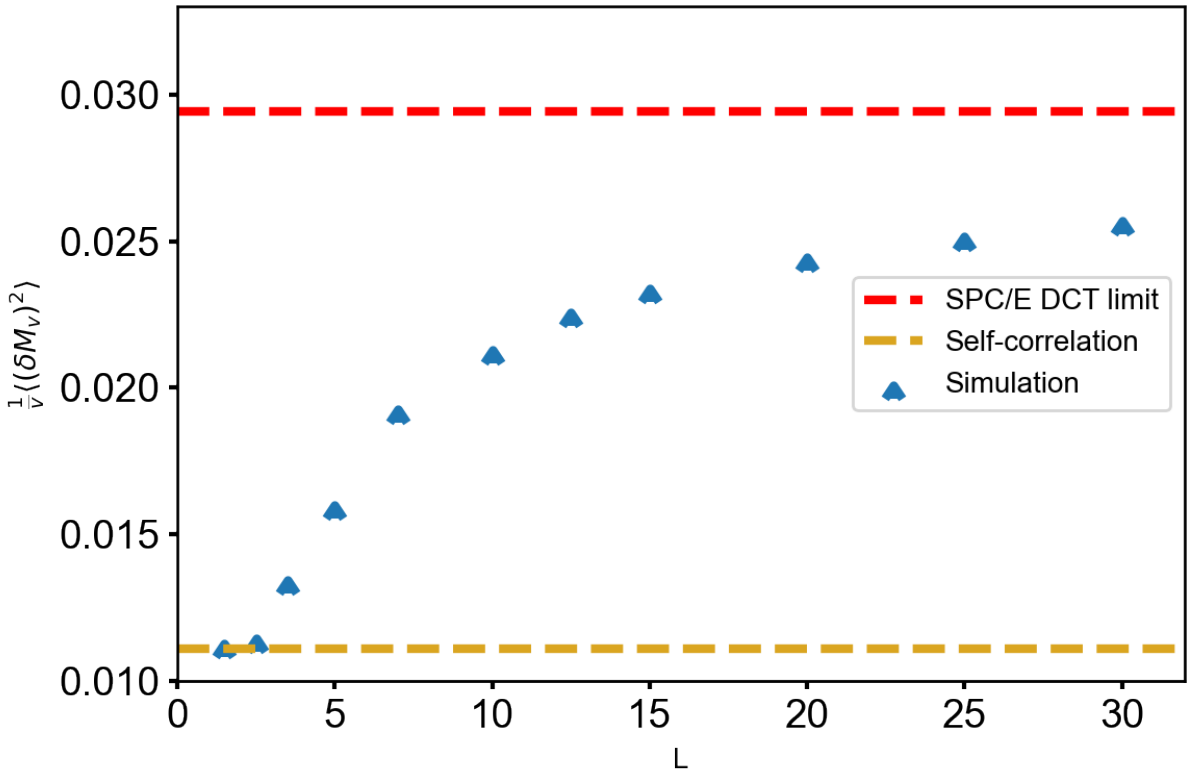

Figure 4.3: The sub-volume variances show slow convergence meeting DCT expectations, limits from Equations 4.22 and 4.25.

For sub-volumes so small as to contain at most one molecule, the variance of M_v is determined by the fluctuations of a freely rotating molecule. A freely rotating dipole will split its average vector components evenly through three dimensions, meaning

$$\langle \mu_z \rangle = 0, \tag{4.23}$$

$$\langle \mu_z^2 \rangle = \frac{1}{3}|\boldsymbol{\mu}|^2, \tag{4.24}$$

compare Equation 3.77. The small sub-volume limit in dipole variance should match that of a single dipole scaled by density,

$$\langle (\delta \boldsymbol{M}_{v,z})^2 \rangle = \frac{\rho_{SPC}}{3} \approx \frac{1}{90} \text{Å}^{-3}. \tag{4.25}$$

Deviations from the DCT prediction in Figure 4.3 are due to the finite size of the sub-volumes, since large volumes asymptotically approach Gaussian field model DCT expectations. Section 4.4 will show that there is a simple correction to account for finite size effects that stems from the dielectric saturation (maximum polarization) of the point-charge water models.

Measurements of χ from simulation

For SPC/E at 298 K, the following sections show estimates of the response function $\boldsymbol{\chi}$ using the correlations between various geometries of sub-volumes, i.e. according to Equation 4.17.

Given two sub-volumes v and v' located at r and r', their displacement is represented

$$\boldsymbol{u} = \boldsymbol{r}' - \boldsymbol{r}, \tag{4.26}$$

$$d = |\boldsymbol{r}' - \boldsymbol{r}|, \tag{4.27}$$

$$\hat{u} = \frac{\boldsymbol{u}}{d}. \tag{4.28}$$

The correlation matrix $\langle \boldsymbol{M}_v(\boldsymbol{r})\boldsymbol{M}_{v'}(\boldsymbol{r}')\rangle$ has nine elements, though only six are unique due to symmetry. Each matrix element corresponds to a mixture of the three spatial dimensions $\hat{x}$, $\hat{y}$, and $\hat{z}$. The mixed terms $\hat{x}\hat{y}$, $\hat{x}\hat{z}$, and $\hat{y}\hat{z}$ vanish in bulk due to symmetry.

In a bulk system, it is convenient to only consider $\hat{u}$ that coincide with a lab-frame axis, though this is not required. The coordinates may also shift from lab-frame Cartesian to a new coordinate system that aligns with $\hat{u}$, illustrated in Figure 4.4.

$$\boldsymbol{M}_{lab} = (\boldsymbol{M}\cdot\hat{x}, \boldsymbol{M}\cdot\hat{y}, \boldsymbol{M}\cdot\hat{z}) \tag{4.29}$$

$$\boldsymbol{M}_{relative} = (\boldsymbol{M}\cdot\hat{u}, \boldsymbol{M}\cdot\hat{u}_{\perp,1}, \boldsymbol{M}\cdot\hat{u}_{\perp,2}). \tag{4.30}$$

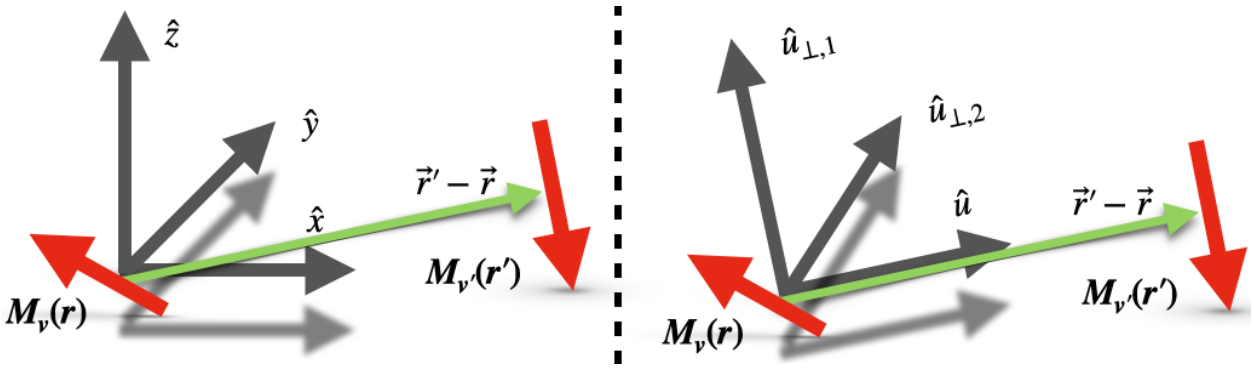

Figure 4.4: Representation of the displacement vector for two dipole vectors. It is favorable for the dipoles to align along $\hat{u}$ but counter-align along the two perpendicular components.

The axis alignment with $\hat{u}$ is important because in an isotropic dipolar liquid correlations along $\hat{u}$ are non-negative, $\langle \delta M_v \delta M_{v'} \rangle_u \geq 0$. Dipole alignment is energetically favorable head-to-tail. The dipole components along $\hat{u}$ therefore orient fully inline. The two dipole components along either $\hat{u}_\perp$, conversely, are favored when counter-aligned head-to-tail, see Figure 4.5.

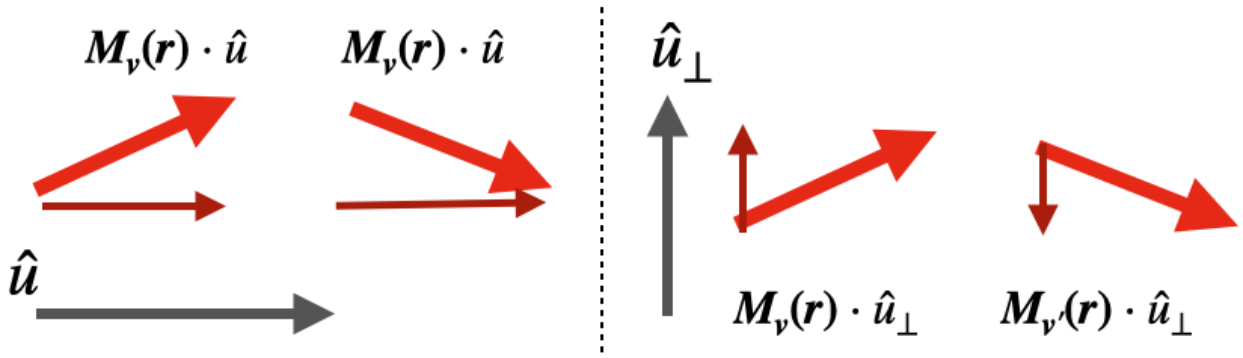

Figure 4.5: Dipoles separated along $\hat{u}$ will tend to align head-to-tail, which is parallel along $\hat{u}$ but is anti-parallel along both $\hat{u}_\perp$. The means that dipole correlations $\langle(\delta \boldsymbol{M}_v)(\delta \boldsymbol{M}_{v'})\rangle$ along $\hat{u}$ will be positive while correlations along either $\hat{u}_\perp$ will be negative.

Figures in the following sections will use a set of similar conventions: the distance between sub-volumes centers d is in units of the cube side-length L and the estimated correlation $\boldsymbol{\chi}(\boldsymbol{r}, \boldsymbol{r}') \sim \langle(\delta \boldsymbol{M}_v)(\delta \boldsymbol{M}_{v'})\rangle$ is normalized with reference to the bulk self-correlation measurement for the same volume. Setting the bulk variance $\langle(\delta \boldsymbol{M}_v)^2\rangle$ as a reference corrects for most finite size effects without attempting to account for their source. Unless otherwise noted, figures are all in units of the SPC/E dipole.

Bulk water

In Figure 4.6, a range of sub-volume correlations are plotted with colors representing $\hat{u}$, $\hat{u}_{\perp,1}$, and $\hat{u}_{\perp,2}$. Sub-volume dipole correlations show the expected behavior, with significant self-correlation in the range $d < L$ (when the sample cubes overlap) that decreases smoothly with d as the two $\hat{u}_\perp$ components move negative before settling at zero for large d.

The results are largely unchanged when sub-volumes are spherical rather than cubic (not shown), and differ mainly in line shape in the self-correlation region of overlapping spheres spanning $d = 0$ to $d = 1$. The shape of the two sub-volumes appears largely unimportant.

The bulk-liquid behavior at $d = 0$ is used as a reference,

$$\langle(\delta \boldsymbol{M}_v)(\delta \boldsymbol{M}_{v'})\rangle_{(d=0)} = \langle(\delta \boldsymbol{M}_v)^2\rangle \,. \tag{4.31}$$

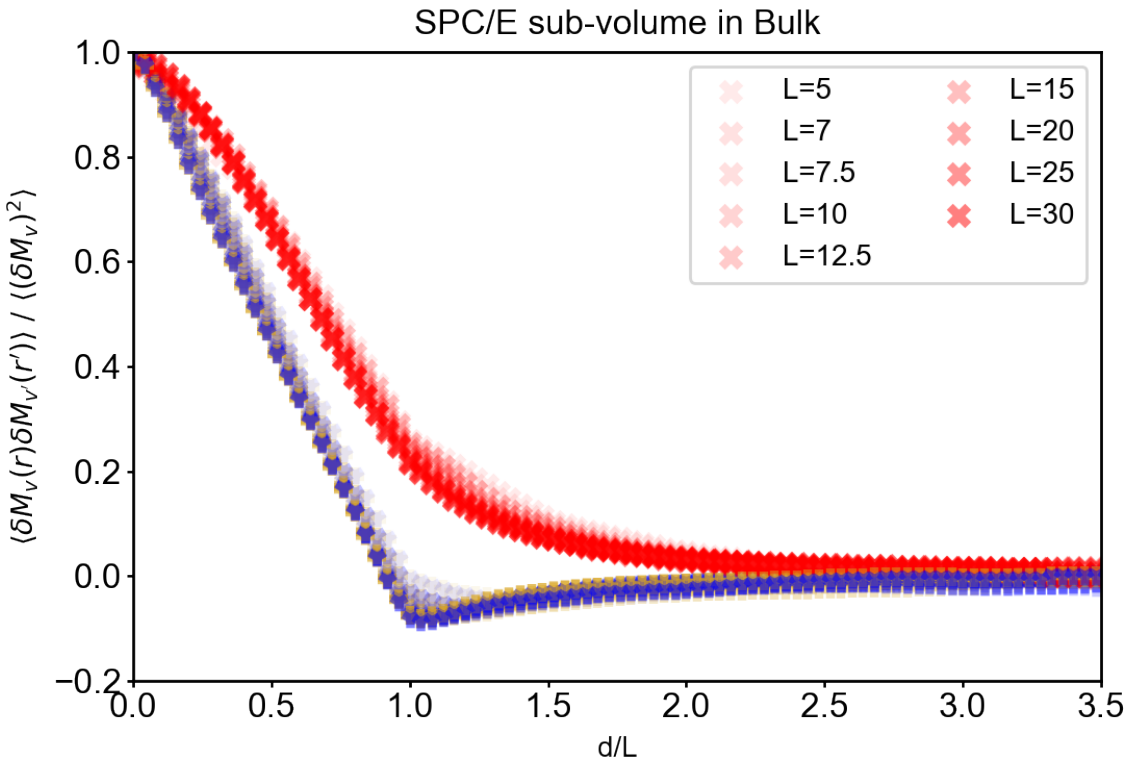

Figure 4.6: Cubic sub-volume correlations for bulk liquid SPC/E under STP. Different markers of the same color denote dipole component correlations along $\hat{u}$ (red), $\hat{u}_{\perp,1}$(blue), and $\hat{u}_{\perp,2}$(olive). The trends match expectations for bulk liquid with distinct behavior for $\hat{u}$ and $\hat{u}_{\perp}$. Sub-volume side lengths are measured in Å.

Figure 4.7 demonstrates a general agreement between a range of sub-volume sizes with predictions from DCT. The results in Figure 4.7 show that the sub-volume approach is an effective way of connecting DCT with simulations of bulk water and is consistent within surprisingly small sub-volumes when referenced to $\langle(\delta \boldsymbol{M}_v)^2\rangle$. The agreement holds in spite of water's tendency to be highly structured. The smallest sub-volume reported, $L = 5$ Å in Figures 4.6 and 4.7 contains on average 4.17 SPC/E molecules at STP.

The agreement of simulation with DCT (solid lines) in Figure 4.7 for small sub-volumes should indicate that DCT adds meaningful information even at molecular length scales.

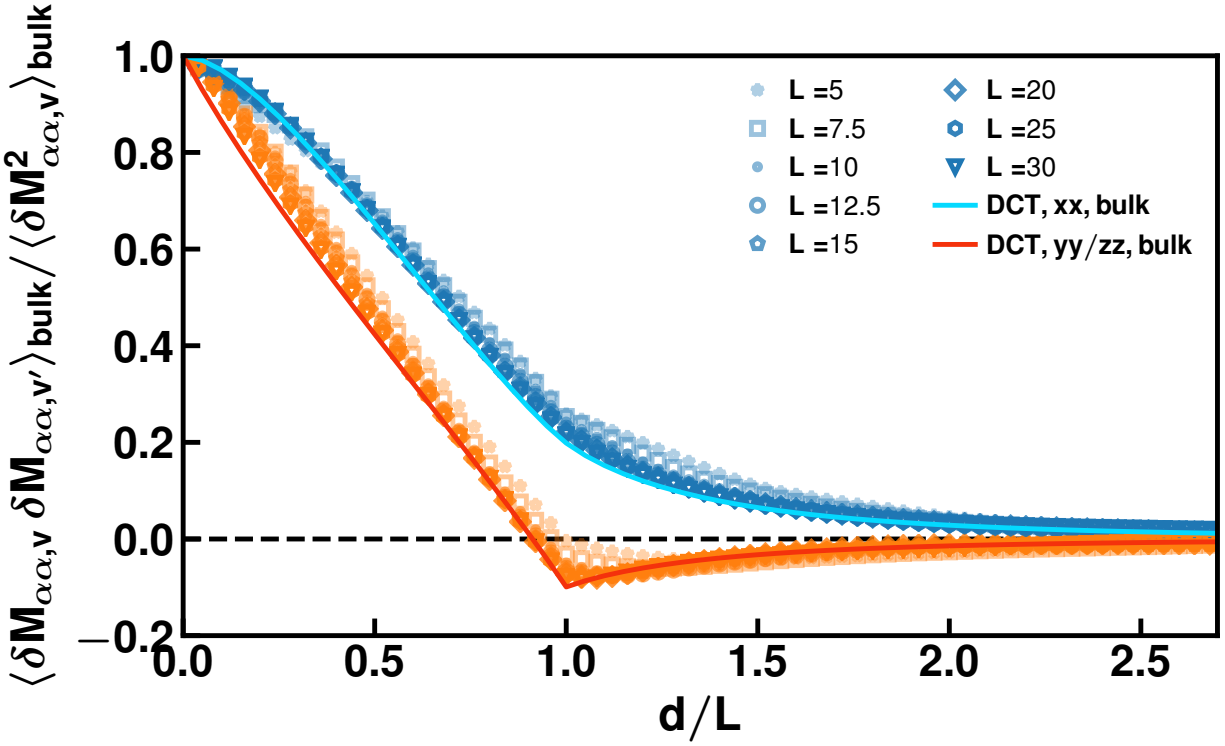

Figure 4.7: Sub-volume dipole correlations referenced by self-correlation for a range of box sizes and according to theory. The agreement between sizes and agreement with theory is good when sub-volume correlations are referenced to the sub-volume self-correlation ($d = 0$) term. Solid lines come from DCT predictions, Equation 4.7.

Interfacial water

Extension of the Gaussian field model of DCT to an interface, Equation 4.11, can also be validated through simulation measurements made at the water interface.

The introduction of an interface breaks the symmetry of bulk and makes the displacement vector $\boldsymbol{u} = \boldsymbol{r}' - \boldsymbol{r}$ insufficient to describe the relative location of sub-volumes. For convenience, $\boldsymbol{M}$ is again expressed in Cartesian coordinates and $z = 0$ is set at the interface, which is flat when projected to depth from the instantaneous interface, see Chapter 1 and ref[5]. Either one or both sub-volumes are pinned to the interface, illustrated in Figure 4.8, such that $\hat{u}$ always aligns with one of the lab-frame axes.

It has been suggested[10, 28, 30, 39–42] that the dielectric of water carries a directional correction when measured either parallel or perpendicular to the surface, usually expressed $\epsilon_{\parallel}$ and $\epsilon_{\perp}$, because the interface breaks isotropy. The various methods for making a directional and depth-dependent dielectric $\vec{\epsilon}(z)$ correction do not all agree.

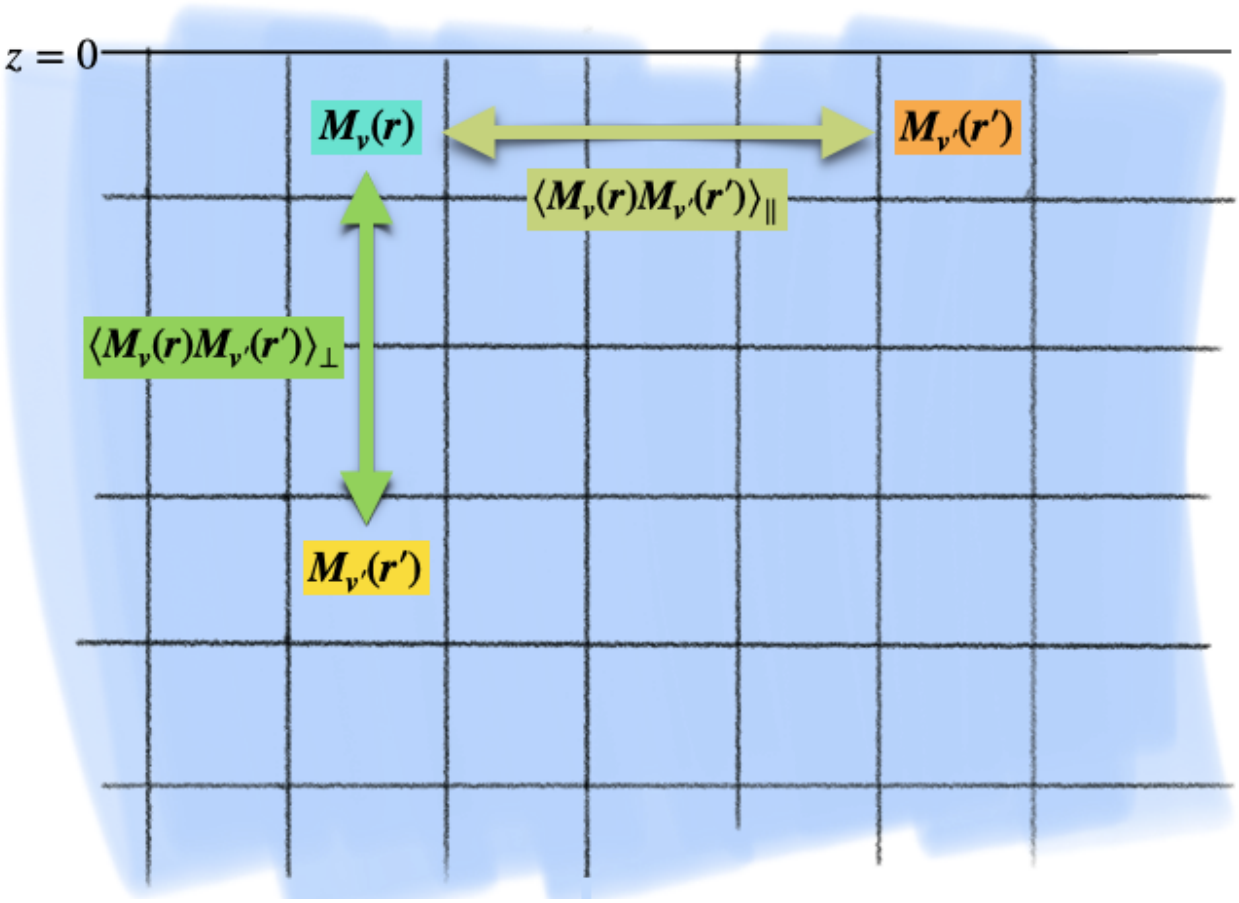

Figure 4.8: With a cell at $\boldsymbol{r}$ located at the interface ($z = 0$), the second cell can be displaced in parallel or perpendicular to the interface. Because water's surface breaks the symmetry of bulk, the net dipole behavior in each cell should vary with depth.

In Figures 4.9 and 4.10, measurements of the parallel and perpendicular correlations are shown to vary in a largely expected way. As with bulk, the trends for sub-volume correlations measured at the interface can be scaled to match with Gaussian field model DCT predictions using Equation 4.10.

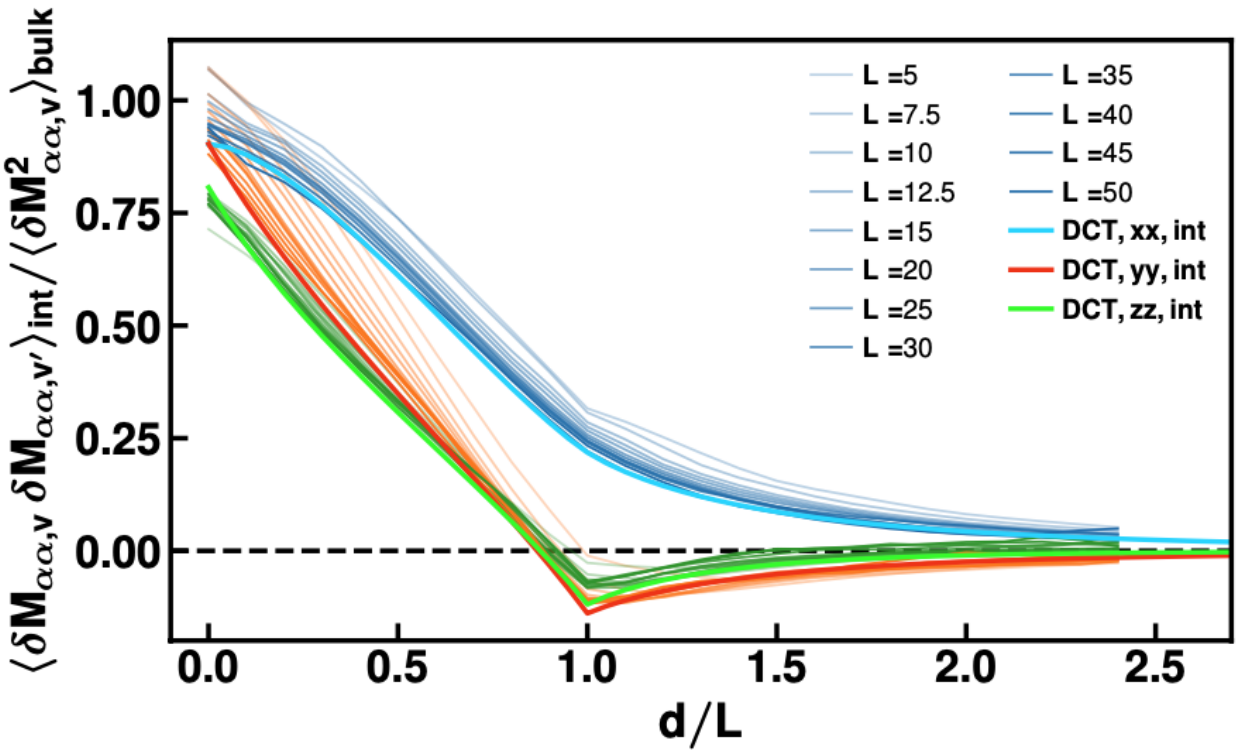

Figure 4.9: Correlations in sub-volume dipole when both cells are located at the air-water interface, $(z = 0, z' = 0)$ and displacements $\boldsymbol{u}$ are made along the xy-plane. Fluctuations in the z-direction (green) are suppressed relative to bulk, which is predicted by DCT. Agreement of sub-volume correlations measured in simulation (faint) with theory (dark) is good when sub-volume correlations are referenced to bulk measurements. Darker lines come from Gaussian field model DCT predictions, Equation 4.10.

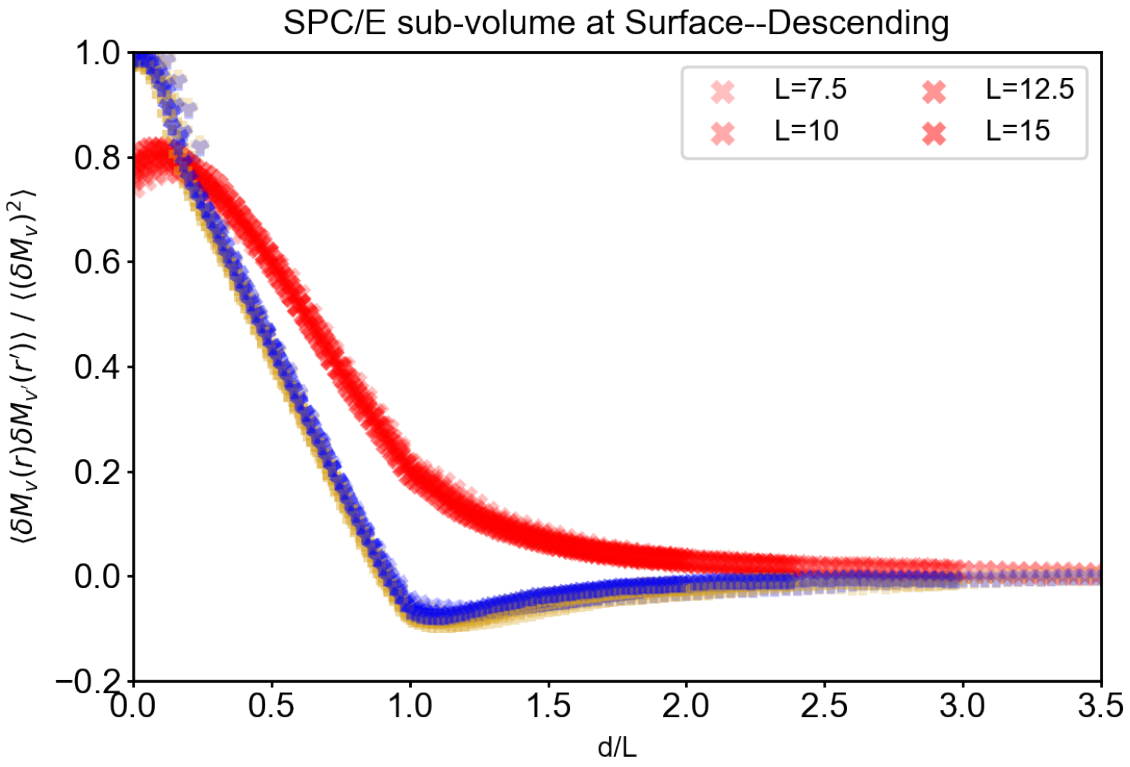

Figure 4.10: Correlations in sub-volume dipole when one cell is located at the surface and the other cell is descending into bulk liquid, ($z = 0, z' < 0$). Since the displacement vector is now along $\hat{z}$, the associated fluctuations in M_z (red) are reduced. Red-$\hat{u}$-z, Olive-$\hat{u}_\perp$-x, Blue-$\hat{u}_\perp$-y

From this perspective, the dielectric permittivity of water behaves predictably according to DCT, even at the interface, when considering cubic sub-volumes at least 5 Å across. It seems possible that the orientational structure at the surface of water may play some role when considering the remaining error of sub-nanometer volumes pinned to the surface.

Application to the dumbbell fluid

The trends in Section 4.3 are not limited to water. The dumbbell fluid is a generic dipole fluid described in ref[10] and shows similar behavior when analyzed under DCT. The dumbbell fluid was proposed as a water-scale dipole that does not share the strongly directional hydrogen bonding or the complicating higher order dipole (quadrupole) of water. It is, by design, a nearly point dipole that can bridge analytical and molecular simulation methods. Figures 4.11 to 4.13 show simulation measurements of the sub-volume dipole correlations.

Measurements of the dumbbell fluid near the interface are more prone to noise because the dumbbell fluid forms a diffuse, extended interface. It is likely that small displacements in Figures 4.12 and 4.13 are poorly converged, but larger displacements match more closely with expectations.

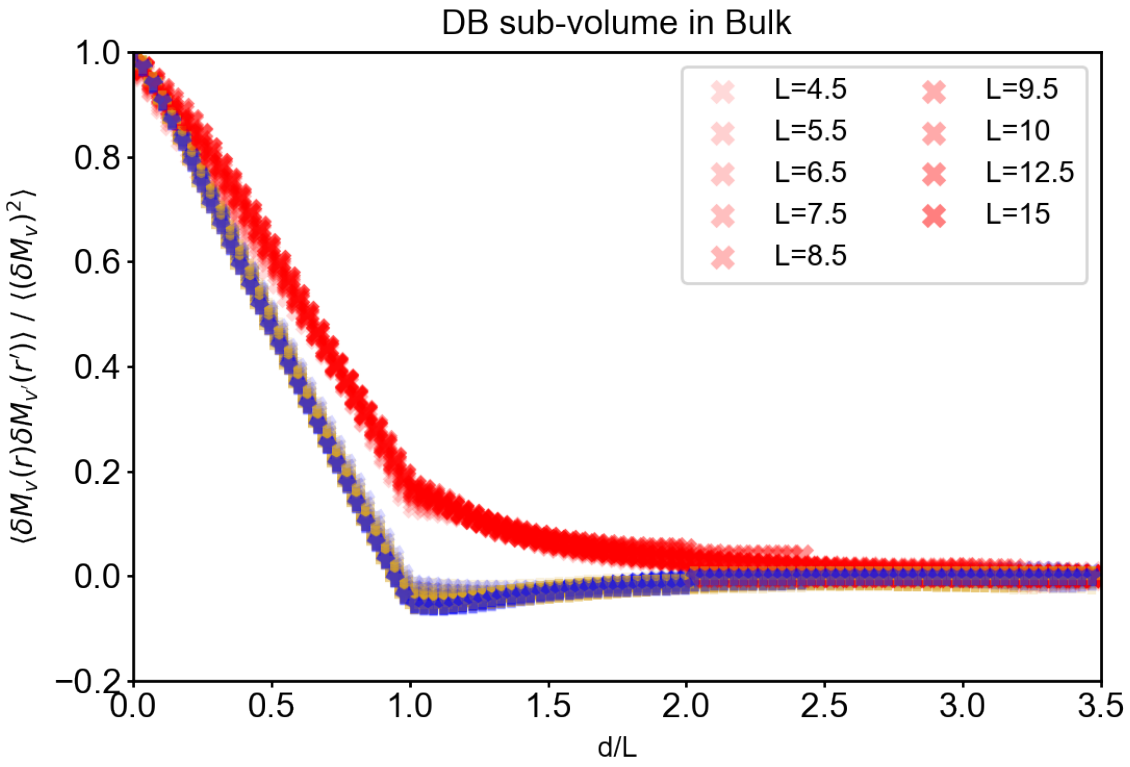

Figure 4.11: Sub-volume dipole correlations for the dumbbell fluid measured in the bulk liquid phase. Results here are very similar to those measured in SPC/E. Different markers of the same color denote correlations along $\hat{u}$ (red), $\hat{u}_{\perp,1}$(blue), and $\hat{u}_{\perp,2}$(olive).

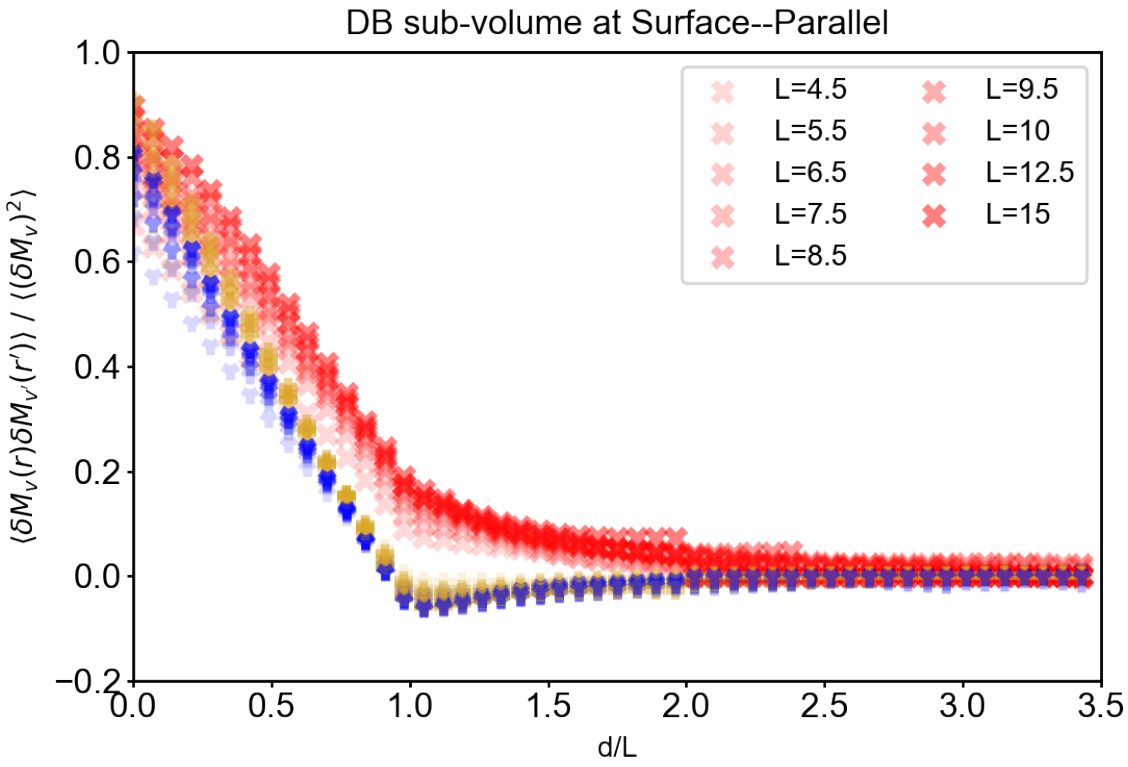

Figure 4.12: Surface dipole correlations of the dumbbell fluid, which largely match behavior seen in water. The dumbbell fluid surface is more diffuse and so the data is less sharp. Red-$\hat{u}$-x, Olive-$\hat{u}_\perp$-y, Blue-$\hat{u}_\perp$-z.

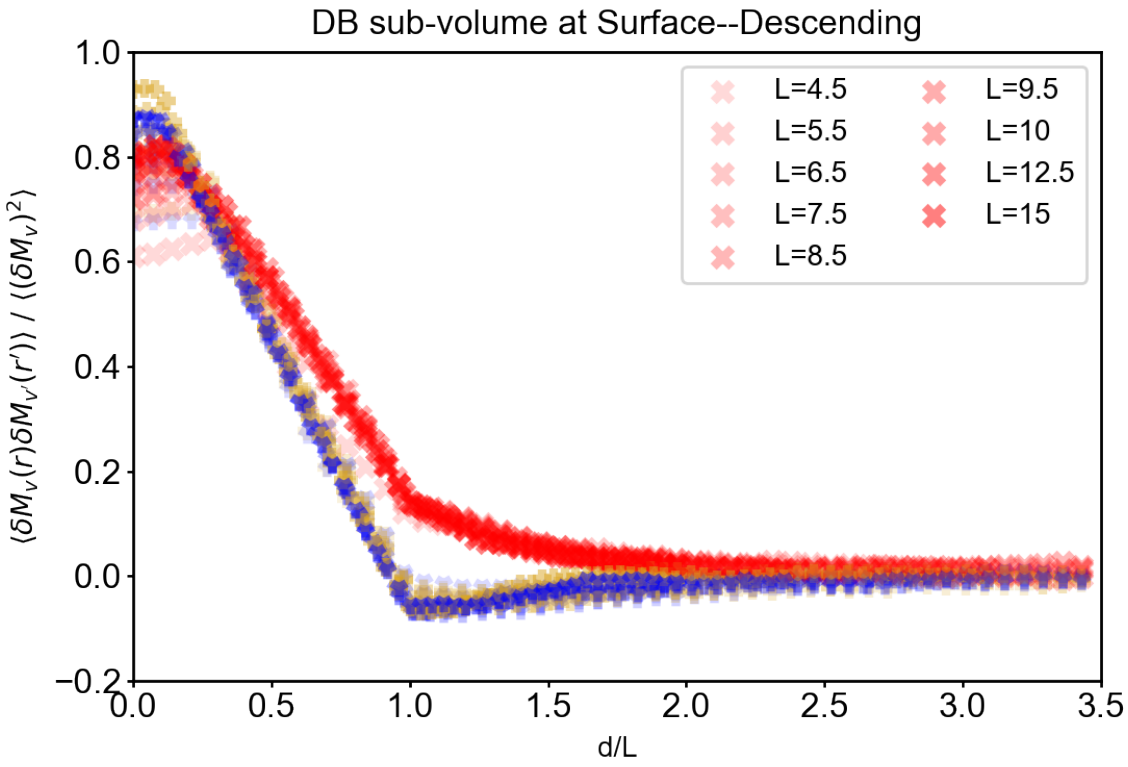

Figure 4.13: Sub-volume dipole correlations of the dumbbell fluid descending from the surface, which mostly matches the behavior seen in water. Red-$\hat{u}$-z, Olive-$\hat{u}_\perp$-x, Blue-$\hat{u}_\perp$-y

4.4 DCT finite size effects explained by large deviations theory

Intro to LDT

Large deviations theory (LDT) characterizes the low-probability extremes of a probability distribution, i.e. explaining the large deviations from the mean. It is a tool for modeling finite size effects because, like other macroscopic models, it also asymptotically converges with large samples.

Working with an extensive macroscopic variable, like the net dipole of a sub-volume,

$$M = \sum_{i=1}^{N} m_i \tag{4.32}$$

$$M = \int_V d\boldsymbol{r} m(\boldsymbol{r}). \tag{4.33}$$

LDT relates the probability of an extensive property with a specific functional form,

$$P(M) \approx e^{-VI(M/V)} \tag{4.34}$$

where the approximate sign indicates model validity at the thermodynamic limit and $I(M/V)$ is called the "rate function" and will be described later.

A cumulant generating function, often used in a statistics context, is defined

$$K(\lambda) := \ln \langle e^{\lambda M} \rangle = \ln Q(\lambda) \tag{4.35}$$

and is useful for generating statistical cumulants, like the mean and variance,

$$\left. \frac{\partial \ln Q(\lambda)}{\partial \lambda} \right|_{\lambda=0} = \langle M \rangle \tag{4.36}$$

$$\left. \frac{\partial^2 \ln Q(\lambda)}{\partial \lambda^2} \right|_{\lambda=0} = \langle \delta M^2 \rangle . \tag{4.37}$$

The cumulant generating function $K(\lambda)$ is extensive and so is often made intensive as the scaled cumulant generating function

$$\phi(\lambda) = \ln Q(\lambda)/V. \tag{4.38}$$

An inverse Laplace transform inverts the moment generating function, $R(\lambda) = e^{K(\lambda)}$, into an expression of probability that uses the scaled cumulant generating function,

$$P(M) = \int d\lambda Q(\lambda) e^{-\lambda M} \tag{4.39}$$

$$= \int d\lambda e^{V(\phi(\lambda) - \lambda \frac{M}{V})}. \tag{4.40}$$

The resulting probability function has a clear dependence on volume, but the intensive piece of the exponential defines the rate function

$$-I(M/V) = \phi(\lambda) - \lambda \frac{M}{V}, \tag{4.41}$$

which is intensive. If M is truly extensive then the rate function will not vary with volume. As $V \to \infty$ in the macroscopic limit, the probability becomes strongly concentrated about a single λ_{max} and the integral can be approximated with a delta function,

$$P(M) \approx e^{-VI(M/V)} \tag{4.42}$$

$$I(M/V) \approx -\frac{\ln(P(M/V))}{V}. \tag{4.43}$$

From another perspective, because the DCT Gaussian Field Model of fluids, Equation 4.6, uses a Gaussian function to model free energy, the probability of a sub-volume dipole fluctuation M is assumed to be regular even for large deviations. Using the Boltzmann weighting

$$P(M) \propto \exp\left[-\frac{M^2}{2\left\langle(\delta M)^2\right\rangle}\right]. \tag{4.44}$$

Re-expressed in terms of large deviations theory by assuming the form of the rate function $I\left(\frac{M}{V}\right)$,

$$P(M) \propto \exp\left[-VI\left(\frac{M}{V}\right)\right], \tag{4.45}$$

which implies

$$I\left(\frac{M}{V}\right) \sim -\frac{\ln P(M/V)}{V} = \frac{1}{V}\left(\frac{M^2}{2\left\langle(\delta M)^2\right\rangle}\right). \tag{4.46}$$

If the rate function is well behaved, it should be parabolic in M.

LDT for finite size effects

The convergence of sub-volume dipole variance is shown Figure 4.3 to be very slow on the Angstrom scale. To the extent that the Gaussian Field model is relevant to SPC/E, the log of the probability distribution $\ln[P(M/V)]$ of Equation 4.46 ought to be quadratic in M and form a parabola. Figures 4.14 and 4.15 demonstrate that the rate function is not fully parabolic because of the maximum of polarizability, i.e. large deviations are asymptotically stunted by dielectric saturation.

Extreme fluctuations in net sub-volume dipole $\boldsymbol{M}/V$ become rare with increasing size, as a larger collective dipole fluctuation becomes energetically unattainable. The tails of the resulting log of probability are impossible to sample through simple simulations. By biasing the net sub-volume dipole in a series of simulations, Figure 4.14 illustrates the limit of extreme fluctuations in sub-volume polarization.

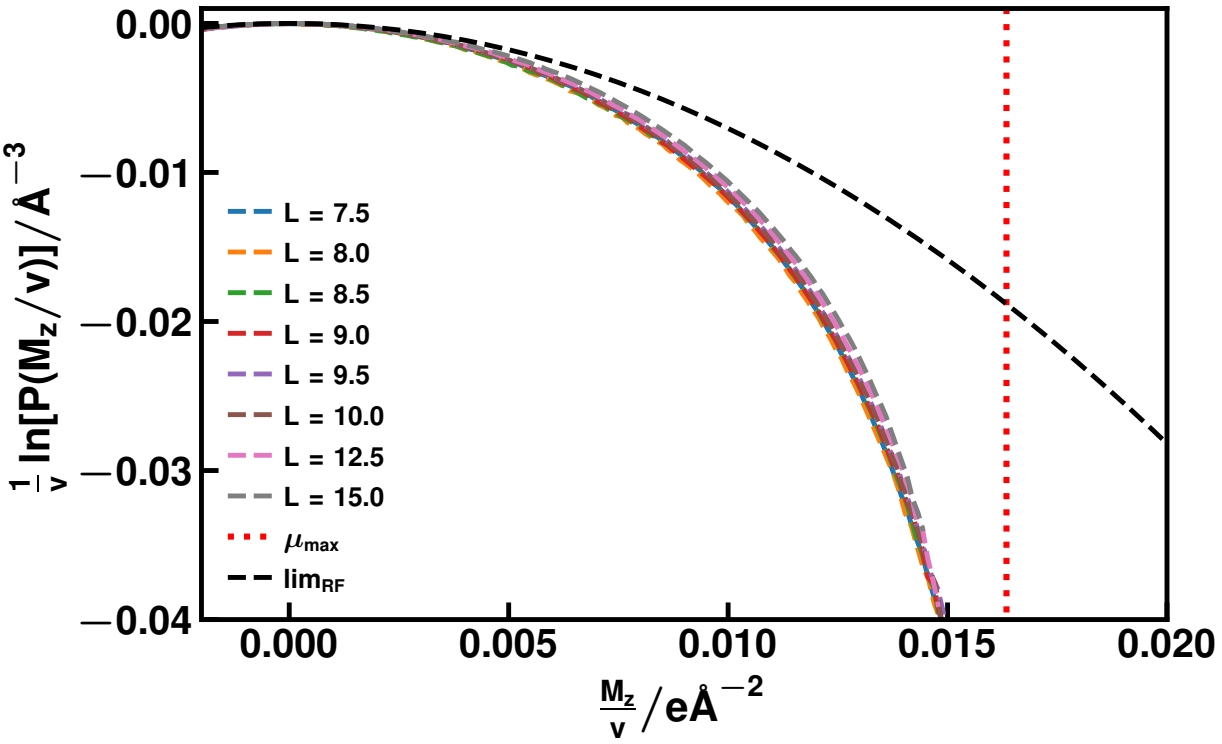

Figure 4.14: Log of probability of sub-volume polarization generated from biased simulations of bulk SPC/E, which is expected to be parabolic according to Equation 4.46. All sub-volumes deviate from quadratic behavior and asymptotically approach the maximum possible polarization for a rigid, non-polarizable liquid model.

The measured data in Figure 4.14 show considerable non-parabolic behavior in the tails of the probability distribution. The plotted data pinch inward when compared to the DCT dashed black line, so the Gaussian Field model–with its parabolic behavior–does not apply when extremes in dipole fluctuations are involved. The "pinching" of parabolas in Figure 4.14 is caused by the asymptotic approach to the maximum polarizability of a sub-volume (dashed red vertical line), which is dielectric saturation. Once a rigid, non-polarizable water model like SPC/E orients its dipole in a given direction, no further increase in molecular polarization is possible. The limit μ_{max} represents a sub-volume at normal liquid density with all SPC/E water molecules fully aligned.

Figure 4.14 explains the connection between the finite size of a sub-volume and the slow convergence of its variance (Figure 4.3). Fluctuations near dielectric saturation are much more likely at small volumes, so the small volumes are least like the DCT limit. While the maximum in polarizability limits the reach of Gaussian field model DCT, it is also shown here that a correction for dielectric saturation can account for almost all of the small sub-volume error.

LDT at the interface

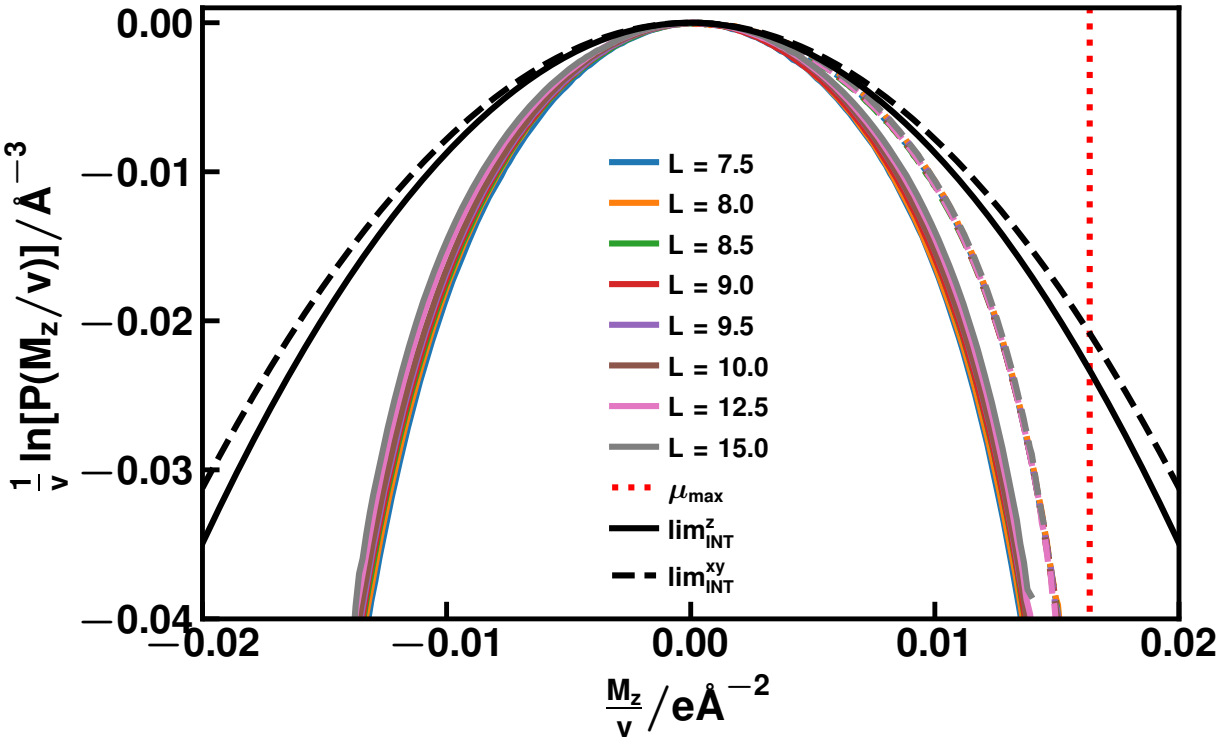

Figure 4.15: Biased measurements of the rate function for different sub-volumes at the interface (colored lines). Included are the expected macroscopic DCT behavior (black) for dipole fluctuations along the plane of the interface (dashed) and perpendicular to the interface (solid). Simulation measurements show an asymptotic limit to large fluctuations in sub-volumes that approach the maximum possible polarization of the sub-volume (dotted red).

The addition of an interface does not significantly change the LDT explanation for finite size effects. Figure 4.15 shows data gathered from biased simulations of an SPC/E liquid/vapor interface. The maximum in polarization suppresses the scale of all fluctuations, and the reduction is most apparent in the unscaled variance of smaller sub-volumes.

Summary

Dielectric Continuum Theory (DCT) is a widely used macroscopic model that simplifies complex electrostatic information. DCT near the scale of molecules may be less effective because the averages that DCT depends on may be limited in scope when only a few molecules

contribute. Non-isotropic constraints, like a liquid/vapor interface, may also interfere with DCT predictions. To address this concern, trends in the average dipole of sub-volumes of water can be measured in molecular dynamics simulations.

The fluctuations in net sub-volume dipole converge to DCT expectations at large sizes, but slowly. Once adjusted for this slow convergence, the correlations between the net dipoles of two sub-volumes are shown to match with DCT predictions for cubes as small as 5 Å across. These trends hold for isotropic liquid water, the water-vapor interface, and for a comparable dipole fluid.

Large deviations in the net dipole of sub-volume are not perfectly described by a Gaussian probability distribution because of a limit in absolute polarization. SPC/E (or any rigid molecular model) can point its dipole in a direction and then no further polarization is possible. Because of this, the tails of the large deviations theory rate function asymptotically approach a dielectric saturation limit, where all dipoles are totally aligned. Rare fluctuations to high net polarization do not strictly follow Gaussian field theory, but the limiting behavior should be expected given dielectric saturation.

These results discourage the notion that liquid water's local permittivity changes significantly on approach to the instantaneous interface with air. In computer simulations the attenuation of probe volume dipole fluctuations is observed near the interface, but the sclae of suppression agrees well with the predictions of DCT.

Bibliography

(1) Chandler, D. *Physical Review E* **1993**, *48*, 2898–2905.

(2) Song, X.; Chandler, D.; Marcus, R. A. *Journal of Physical Chemistry* **1996**, *100*, 11954–11959.

(3) Song, X.; Chandler, D. *Journal of Chemical Physics* **1998**, *108*, 2594–2600.

(4) Shaffer, P. R., *Ion solvation at air-water interfaces*; University of California, Berkeley: 2013.

(5) Willard, A. P.; Chandler, D. *The Journal of Physical Chemistry B* **2010**, *114*, 1954–1958.

(6) Pártay, L. B.; Hantal, G.; Jedlovszky, P.; Vincze, Á.; Horvai, G. *Journal of Computational Chemistry* **2008**, DOI: 10.1002/jcc.20852.

(7) Hantal, G.; Darvas, M.; Pártay, L. B.; Horvai, G.; Jedlovszky, P. *Journal of Physics Condensed Matter* **2010**, *22*, DOI: 10.1088/0953-8984/22/28/284112.

(8) Zhou, T.; McCue, A.; Ghadar, Y.; Bakó, I.; Clark, A. E. *Journal of Physical Chemistry B* **2017**, *121*, 9052–9062.

(9) Sega, M.; Fábián, B.; Jedlovszky, P. *Journal of Chemical Physics* **2015**, *143*, 114709.

(10) Cox, S. J.; Geissler, P. L. *Journal of Chemical Physics* **2018**, DOI: 10.1063/1.5020563.

(11) Ben-Naim, A. *The Journal of Chemical Physics* **1971**, *54*, 3682–3695.

(12) Silverstein, K. A.; Haymet, A. D.; Dill, K. A. *Journal of the American Chemical Society* **1998**, DOI: 10.1021/ja973029k.

(13) Benjamin, I. *Annual Review of Physical Chemistry* **1997**, *48*, 407–451.

(14) Luzar, A.; Chandler, D. *Nature* **1996**, *379*, 55–57.

(15) Vega, C.; McBride, C.; Sanz, E.; Abascal, J. L. F. *Physical Chemistry Chemical Physics* **2005**, *7*, 1450.

(16) Kessler, J.; Elgabarty, H.; Spura, T.; Karhan, K.; Partovi-Azar, P.; Hassanali, A. A.; Kühne, T. D.; Kü, T. D. *Journal of Physical Chemistry B* **2015**, DOI: 10.1021/acs.jpcb.5b04185.

(17) Pezzotti, S.; Serva, A.; Gaigeot, M.-P. P.; Galimberti, D. R.; Gaigeot, M.-P. P. *Journal of Physical Chemistry Letters* **2017**, *8*, 3133–3141.

(18) Shen, Y. R.; Ostroverkhov, V. *Chemical Reviews* **2006**, *106*, 1140–1154.

(19) Goh, M. C.; Hicks, J. M.; Kemnitz, K.; Pinto, G. R.; Bhattacharyya, K.; Eisenthal, K. B.; Heinz, T. F. *Journal of Physical Chemistry* **1988**, *92*, 5074–5075.

(20) Fan, Y.; Chen, X.; Yang, L.; Cremer, P.; Gao, Y. Q. *Journal of Physical Chemistry B* **2009**, *113*, 11672–11679.

(21) Shin, S.; Willard, A. P. *The Journal of Physical Chemistry Letters* **2018**, *9*, 1649–1654.

(22) Wilson, M. A.; Pohorille, A.; Pratt, L. R. *The Journal of Chemical Physics* **1988**, *88*, 3281–3285.

(23) Taylor, R. S.; Dang, L. X.; Garrett, B. C. *Journal of Physical Chemistry* **1996**, *100*, 11720–11725.

(24) Ballenegger, V.; Hansen, J.-P. *The Journal of Chemical Physics The Journal of Chemical Physics The Journal of Chemical Physics* **2005**, *12211*, 114711–911.

(25) Willard, A. P.; Reed, S. K.; Madden, P. A.; Chandler, D. *Faraday Discussions* **2008**, *141*, 423–441.

(26) Chaplin, M. *Water* **2009**, *1*, 1–28.

(27) Kathmann, S. M.; Kuo, I.-F. W.; Mundy, C. J.; Schenter, G. K.; William Kuo, I.-F.; Mundy, C. J.; Schenter, G. K. *J. Phys. Chem. B* **2011**, *115*, 4369–4377.

(28) Bonthuis, D. J.; Gekle, S.; Netz, R. R. *Physical Review Letters* **2011**, *107*, DOI: 10.1103/PhysRevLett.107.166102.

(29) Besford, Q. A.; Liu, M.; Christofferson, A. J. *J. Phys. Chem. B* **2018**, *122*, 50.

(30) Vega, C. *Molecular Physics* **2015**, *113*, 1145–1163.

(31) Ben-Naim, A. *Molecular Physics* **1972**, *24*, 705–721.

(32) Silverstein, K. A.; Dill, K. A.; Haymet, A. D. *Fluid Phase Equilibria* **1998**, *150*, 83–90.

(33) Urbic, T. *Physical Review E* **2017**, *96*, 1–7.

(34) Luzar, A.; Chandler, D. *The Journal of Chemical Physics* **1993**, *98*, 8160–8173.

(35) Alper, H. E.; Levy, R. M. *Journal of Physical Chemistry* **1990**, *94*, 8401–8403.

(36) Nikzad, M.; Azimian, A. R.; Rezaei, M.; Nikzad, S. *The Journal of Chemical Physics* **2017**, *1471*, 204701–174701.

(37) Macky, W. A. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **1931**, *133*, 565–587.

(38) Taylor, S. G. *Royal Society* **1964**, 22–24.

(39) Baer, M. D.; Stern, A. C.; Levin, Y.; Tobias, D. J.; Mundy, C. J. *The Journal of Physical Chemistry Letters* **2012**, *3*, 1565–1570.

(40) Sato, T.; Sasaki, T.; Ohnuki, J.; Umezawa, K.; Takano, M. *Physical Review Letters* **2018**, DOI: 10.1103/PhysRevLett.121.206002.

(41) Loche, P.; Ayaz, C.; Schlaich, A.; Bonthuis, D. J.; Netz, R. R. *Journal of Physical Chemistry Letters* **2018**, *9*, DOI: 10.1021/acs.jpclett.8b02473.

(42) Schlaich, A.; Knapp, E. W.; Netz, R. R. *Physical Review Letters* **2016**, *117*, 048001.

(43) Wertheim, M. S. *Annual Review of Physical Chemistry* **1979**, *30*, 471–501.

(44) Plimpton, S. *Journal of Computational Physics* **1995**, *117*, 1–19.

(45) Martínez, L.; Andrade, R.; Birgin, E. G.; Martínez, J. M. *Journal of Computational Chemistry* **2009**, *30*, 2157–2164.

(46) Humphrey, W.; Dalke, A.; Schulten, K. *Journal of Molecular Graphics* **1996**, *14*, 33–38.

(47) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *Journal of Physical Chemistry* **1987**, *91*, 6269–6271.

(48) Mahoney, M. W.; Jorgensen, W. L. *Journal of Chemical Physics* **2000**, *112*, 8910–8922.

References for tools and methods

Tools and methods that were used but may not have been cited within the body of the dissertation.

- lammps [44]
- packmol [45]
- Visual Molecular Dynamics (VMD) [46]
- SPC/E [47]
- TIP5P [48]
- BN2D [11, 31]