EXTENDING TOPOLOGICAL CONSTRAINT THEORY TO GLASS NETWORKS WITH HYDROGEN BONDING

by

Austin Bloom

A thesis submitted to the faculty of The University of North Carolina at Charlotte in partial fulfillment of the requirements for the degree of Master of Science in Applied Physics

Charlotte

2022

Approved by:

Dr. Donald Jacobs

Dr. Susan Trammell

Dr. Irina Nesmelova

©2022 Austin Bloom ALL RIGHTS RESERVED

ABSTRACT

AUSTIN W. BLOOM. Extending Topological Constraint Theory to Glass Networks with Hydrogen Bonding. (Under the direction of DR. DONALD JACOBS)

Topological constraint theory has been successful in providing a simple model for the liquid-glass transition in covalent bond bending networks. However, it is limited to networks where all bonds are quenched and possess equivalent potential well depths. This limitation makes it difficult to model glass networks with heterogeneous bonds such as hydrogen bond networks. By adding a single adjustable parameter to model hydrogen bonds, topological constraint theory can be extended to model glass networks with both quenched (covalent) and flickering (hydrogen) bonds. This parameter is the time scale of observation, and it is implemented by calculating hydrogen bond probabilities by time averaging over the recent past using data from molecular dynamics simulation. The rigidity properties are quantified with the pebble game algorithm for body-bar networks to identify rigid subgraphs. Molecular dynamics trajectory data is mapped into a generic graph topology for a rigidity analysis. The hydrogen bond dynamics are characterized with correlation functions, and a method is presented for determining the optimal geometrical hydrogen bond definition. Further, it will be shown that spatial-temporal correlations present in molecular dynamics simulations shift the rigidity percolation threshold to lower mean coordination numbers. It is found from rigidity percolation and scaling theory that a second order rigidity transition is driving the liquid-glass transition, and an analysis of the β critical exponent suggests that covalent bond bending networks belong to the same universality class as thermal phase transitions.

ACKNOWLEDGEMENTS

I would like to acknowledge my advisor, Dr. Donald Jacobs, for his guidance in this project. I would also like to acknowledge Dr. Susan Trammell and Dr. Irina Nesmelova for serving on my thesis committee and giving valuable criticism and Nikita Nikulsin who laid the foundation which made this thesis possible. Finally, I would like to acknowledge the University of North Carolina at Charlotte for providing the high performance computing cluster on which I ran thousands of hours of simulations.

TABLE OF CONTENTS

LIST OF TABLES	vii
LIST OF FIGURES	viii
LIST OF ABBREVIATIONS	xiv
CHAPTER 1: Introduction	1
1.1. The Ergodic Hypothesis	7
1.2. Molecular Dynamics Simulation	8
CHAPTER 2: Theory	11
2.1. Theory of Phase Transitions	11
2.1.1. Connection to Percolation Theory	13
2.2. Rigidity Theory	14
CHAPTER 3: Methods	22
3.1. Simulation Protocols	22
3.1.1. Volume Hysteresis	23
3.2. Hydrogen Bond Analysis	24
3.3. Pebble Game algorithm	29
3.4. Procedure	32
CHAPTER 4: Results and Analysis	35
4.1. Volume Hysteresis	35
4.2. Hydrogen Bond Analysis	37
4.2.1. Hydrogen Bond PDF distributions	37
4.2.2. Hydrogen Bond Correlation Functions	39

		vi
4.3. Rigidi	ty Analysis	52
4.3.1.	Time-Correlated Pebble Game	52
4.3.2.	Cluster Distributions	64
4.3.3.	Uniform-Probability Pebble Game	77
4.3.4.	Critical Exponents	88
CHAPTER 5:	Conclusion	91
REFERENCES	3	96

TABLE 3.1: Example of short label.

- TABLE 4.1: Results of the H-bond analysis for CMHP at 270 K. The left three columns and right three columns are calculated from simulations which output data every 6 ps and 0.1 ps, respectively. θ_c and τ are the cutoff angle and H-bond lifetime defined in the text. The number of H-bonds, N_h , refers to $\langle h(0) \rangle$ defined in the H-bond correlation function (equation 3.1).
- TABLE 4.2: Results of the H-bond analysis for CMHP at 300 K. The left three columns and right three columns are calculated from simulations which output data every 6 ps and 0.1 ps, respectively. θ_c and τ are the cutoff angle and H-bond lifetime defined in the text. The number of H-bonds, N_h , refers to $\langle h(0) \rangle$ defined in the H-bond correlation function (equation 3.1). A – indicates that data was not collected.
- TABLE 4.3: Results of the H-bond analysis for CDHP at 270 K. The left three columns and right three columns are calculated from simulations which output data every 6 ps and 0.1 ps, respectively. θ_c and τ are the cutoff angle and H-bond lifetime defined in the text. The number of H-bonds, N_h , refers to $\langle h(0) \rangle$ defined in the H-bond correlation function (equation 3.1). A - indicates that data was not collected.
- TABLE 4.4: Results of the H-bond analysis for CDHP at 300 K. The left three columns and right three columns are calculated from simulations which output data every 6 ps and 0.1 ps, respectively. θ_c and τ are the cutoff angle and H-bond lifetime defined in the text. The number of H-bonds, N_h , refers to $\langle h(0) \rangle$ defined in the H-bond correlation function (equation 3.1). A indicates that data was not collected.
- TABLE 4.5: Critical exponents for CMHP at 300 K. The critical exponent 90 β is calculated from a linear regression of $P \sim (p p_c)^{\beta}$ plotted on a logarithmic scale. r is the Pearson correlation coefficient for the fit. Only the first 20 data points after p_c were used in the fit.

26

48

49

50

51

LIST OF FIGURES

 $\mathbf{2}$

FIGURE 1.1: This sketch shows the typical hysteresis behaviour of the heat capacity of vitrified liquids. The inflection point is defined to be the glass transition temperature.	2
FIGURE 1.2: This sketch demonstrates the strong/fragile classification of glass forming materials. Linearity indicates arrhenius behaviour and deviations from linearity indicate non-arrhenius behaviour (or non-exponential relaxations).	4
FIGURE 2.1: Panels (a) through (d) show the evolution from floppy to rigid of a simple generic graph. In panel (a), there are two inde- pendent motions associated with the horizontal bars which can pivot	17

- around the joints connected by the vertical bar. The addition of an independent constraint in (b) removes one floppy mode. The addition of the dashed edge in (c) now turn the four vertices into a single rigid cluster. The thick dashed line in (d) is redundant because the graph has already undergone a rigidity transition.
- FIGURE 3.1: θ is the HDA angle. H-Bonds are restricted within a cone with cutoff angle θ_c . HDA angles outside of this cone are not considered H-Bonds.
- FIGURE 3.2: A demonstration of the pebble game on a generic network. Independent (redundant) bonds are shown with solid (dashed) lines that are (are not) covered by a pebble. Large (filled, open) circles denote (anchored, free) pebbles on (bonds, sites). The two closest pebbles to a given site are tethered to that site. Small (filled, open) circles denote sites belonging to (one, more than one) rigid cluster. Overconstrained bonds are shown with heavy dark lines. Shaded regions denote 2d rigid bodies. (a) Five free pebbles indicate five floppy modes until a new bond is added and tested for independence. A fourth free pebble is found via the path traced by arrows. (b) The added bond is independent and thus covered. There are now six rigid clusters and four floppy modes. Figure and caption reproduced from reference [1]

31 FIGURE 3.3: Figure reproduced from [2]

FIGURE 4.1: The number of trehalose and water molecules as a function of concentration. Solid (dashed) lines refer to CDHP (CMHP) while black (red) curves refer to water (trehalose). The water and trehalose curves intersect at ~ 96 percent solute by weight.

30

35

25

FIGURE 4.2: Volume hysteresis data for (a) CMHP and (b) CDHP at different temperatures.	36
FIGURE 4.3: Shown above are the H-Bond donor-acceptor angle distributions taken from a single frame. The number at the red dot is the angle of the first minimum of the distribution. Note the prominent peaks succeeding the first which correspond to the non-participatory water hydrogen.	37
FIGURE 4.4: See caption of figure 4.3. Notice the absence of prominent secondary and tertiary peaks as compared to when water is a donor.	38
FIGURE 4.5: Choline-Choline interactions are excluded because the PDF is generated from too few data points to make any meaningful inferences.	39
FIGURE 4.6: The log correlation functions for choline-phosphate and water-phosphate donor-acceptor H-Bond pairs at 72.5 percent solute by weight are shown as a function of time over the entire simulation for different angle cutoffs. The insets show the middle of the plots where the decay is exponential. One can see that there is a slight angle dependence which allows a geometric cutoff to be made. Also note the faster decay in (b) compared to (a) despite the fact that wa- ter neighboring phosphate is not considered a plasticizer in CMHP systems. These plots were calculated with no distance cutoff.	40
FIGURE 4.7: The lifetimes calculated from equation 3.2 are shown for the CMHP system at 300 K and 72.5 percent solute by weight. (a) shows the lifetime for each H-Bonded donor-acceptor pair as a function of cutoff angle with no HA distance restriction. (b) shows the lifetimes with no angle restriction. Notice the uptick in the inset of (b) at around 2.9 angstroms which corresponds to the second hydrogen in water being incorrectly sampled as an H-Bond.	41
FIGURE 4.8: See the caption of 4.7. Shown are the lifetime plots for the CMHP system at 300 K and 89 percent solute by weight.	42
FIGURE 4.9: H-bond lifetime data as a function of cutoff angle is shown for CMHP and CDHP at 270 K and 300 K each. The concentration (in % solute by weight) for each panel is: (a) 62 %; (b) 65 %; (c) 60 %; and (d) 65 %. These concentrations are all below the glass transition concentration.	43

ix

FIGURE 4.10: H-bond lifetime data as a function of cutoff angle is shown for CMHP and CDHP at 270 K and 300 K each. The concentration (in % solute by weight) for each panel is: (a) 71 %; (b) 71 %; (c) 69.5 %; and (d) 74 %. These concentrations are all in the vicinity of the glass transition concentration.	45
FIGURE 4.11: H-bond lifetime data as a function of cutoff angle is shown for CMHP and CDHP at 270 K and 300 K each. The concentration (in % solute by weight) for each panel is: (a) 80 %; (b) 89 %; (c) 84.5 %; and (d) 89 %. These concentrations are all above the glass transition concentration.	47
FIGURE 4.12: Data for the number of vertices (atoms) in the network after pruning is shown as a function of window size for various concentrations.	53
FIGURE 4.13: Data for the number of floppy modes normalized per de- gree of freedom is shown as a function of window size for various concentrations. The number of degrees of freedom is defined as three times the number of vertices.	55
FIGURE 4.14: Data for the RCS is shown as a function of window size for various concentrations.	56
FIGURE 4.15: Data for the mean coordination number is shown as a function of window size for various concentrations.	57
FIGURE 4.16: Data for the fraction of sites in the largest cluster is shown as a function of window size for various concentrations.	58
FIGURE 4.17: Data for the number of floppy modes normalized per degree of freedom is shown as a function of mean coordination number for various concentrations. Each point along each curve corresponds to a different window size.	60
FIGURE 4.18: Data for the RCS is shown as a function of mean coor- dination number for various concentrations. Each point along each curve corresponds to a different window size.	61
FIGURE 4.19: Data for the fraction of sites in the largest cluster is shown as a function of mean coordination number for various concentrations. Each point along each curve corresponds to a different window size.	62

х

FIGURE 4.20: Data for the mean coordination number is shown as a function of concentration for various window sizes.	63
FIGURE 4.21: Data for the number of floppy modes normalized per degree of freedom is shown as a function of concentration for various window sizes.	65
FIGURE 4.22: Data for the fraction of sites in the largest cluster is shown as a function of concentration for various window sizes.	66
FIGURE 4.23: Data for the number of sites in the largest rigid cluster is shown as a function of window size on a semi-log scale for various concentrations of CMHP at 270 K. The histograms are broken down by the type of molecule participating in the largest cluster, but only atoms within molecules are counted not whole molecules.	67
FIGURE 4.24: Data for the number of sites in the largest rigid cluster is shown as a function of window size on a semi-log scale for various concentrations of CMHP at 300 K. The histograms are broken down by the type of molecule participating in the largest cluster, but only atoms within molecules are counted not whole molecules.	68
FIGURE 4.25: Data for the number of sites in the largest rigid cluster is shown as a function of window size on a semi-log scale for various concentrations of CDHP at 270 K. The histograms are broken down by the type of molecule participating in the largest cluster, but only atoms within molecules are counted not whole molecules.	69
FIGURE 4.26: Data for the number of sites in the largest rigid cluster is shown as a function of window size on a semi-log scale for various concentrations of CDHP at 300 K. The histograms are broken down by the type of molecule participating in the largest cluster, but only atoms within molecules are counted not whole molecules.	70
FIGURE 4.27: Data for the distribution of rigid clusters for various con- centrations and window sizes of CMHP at 270 K.	71
FIGURE 4.28: Data for the distribution of rigid clusters for various con- centrations and window sizes of CMHP at 300 K.	72
FIGURE 4.29: Data for the distribution of rigid clusters for various con- centrations and window sizes of CDHP at 270 K.	73

xi

	xii
FIGURE 4.30: Data for the distribution of rigid clusters for various con- centrations and window sizes of CDHP at 270 K.	74
FIGURE 4.31: Data for the mean coordination number is shown as a function of H-bond probability.	78
FIGURE 4.32: Data for the number of Floppy modes normalized per degree of freedom is shown as a function of H-bond probability.	79
FIGURE 4.33: Data for the RCS is shown as a function of H-bond probability.	80
FIGURE 4.34: Data for the fraction of sites in the largest rigid cluster is shown as a function of H-bond probability.	82
FIGURE 4.35: Data for the number of floppy modes normalized per degree of freedom is shown as a function of mean coordination number. Each data point along each curve corresponds to a different uniform H-bond probability.	83
FIGURE 4.36: Data for the RCS is shown as a function of mean coordi- nation number. Each data point along each curve corresponds to a different uniform H-bond probability.	84
FIGURE 4.37: Data for the fraction of sites in the largest cluster is shown as a function of mean coordination number. Each data point along each curve corresponds to a different uniform H-bond probability.	85
FIGURE 4.38: Comparison between the time-correlated PG and the uni- form probability PG. The fraction of sites in the largest cluster are shown for various concentrations for CMHP and CDHP at 270 K and 300 K each. There is a distinct shift in the rigidity percolation threshold with time correlations causing rigidity to percolate at lower MCNs.	86
FIGURE 4.39: Hydrogen bond probabilities calculated from simulation data. The difference between the uniform H-bond probability at the percolation threshold (as determined by the RCS peak) of the UPPG and the average H-bond probability from simulation data is plotted. When $p_{hb} > p_t$ ($p_{hb} < p_t$), the average H-bond probability is higher (lower) than the percolation threshold and the system is predicted to be rigid (floppy). The large (small) symbols refer to simulation data sampled every 6 ps (0.1 ps).	87

FIGURE 4.40: Data for the fraction of sites in the percolating rigid cluster is shown as a function of $p - p_c$ on a logarithmic scale, where p is the uniform H-bond probability, and p_c is the approximate cutoff for the existence of the percolating cluster as determined from the peaks of the RCS in figure (4.33). Each data point along each curve corresponds to a different uniform H-bond probability. 89

LIST OF ABBREVIATIONS

- CDHP: Choline Dihydrogen Phosphate
- CMHP: Choline Monohydrogen Phosphate
- H-bond: Hydrogen Bond
- HA: Hydrogen-Acceptor distance
- HDA: Hydrogen-Donor-Acceptor angle
- MD: Molecular Dynamics
- PG: Pebble Game
- RCS: Rigid Cluster Susceptibility
- RFM: Reactive Flux Method
- TCPG: Time-correlated Pebble Game
- TCT: Topological Constraint Theory
- UPPG: Uniform Probability Pebble Game
- VPG: Virtual Pebble Game
- XRD: X-Ray Diffraction

•

CHAPTER 1: Introduction

Since the first X-ray diffraction (XRD) studies, considerable progress has been made in untangling the phenomenology of the glassy, or vitreous, state. It has been known since Kauzmann that the liquid-glass transition is largely a relaxation phenomenon, not a thermodynamic phase transition [3]. Zachariasen noted that certain oxides experienced a glass transition while others did not. He postulated that compositions of the form, $A_m O_n$, would be an economical route for systematically studying the glassy state by modulating the constituent components [4]. This insight - coupled with Kauzmann's view that the glass transition is an entropy crisis - would later lead Phillips to propose a topological constraint theory (TCT) of glasses which focuses solely on counting degrees of freedom. Despite its almost naive simplicity, TCT is in excellent agreement with experiment [5]. The collaborations between Thorpe and Phillips utilized Maxwell's research in rigidity theory and furnished the mathematical methods of percolation theory to extend TCT [6] [7] [8]. An important contribution to TCT is a graph-theoretic algorithm known as the pebble game (PG) which was developed by Jacobs and Thorpe [1]. The PG is built on Laman's theorem in graph theory which allows one to unambiguously identify rigid subgraphs in two dimensions (the pebble game algorithm was then generalized further to three dimensions). The PG has been successful in elucidating the effects of rigidity in two dimensional lattices [1] [9] [10] [11] [12] [13], chalcogenide glasses [5] [2], the dynamics of proteins [14] [15], and in suggesting a possible mechanism for the Boson peak [16].

The foregoing represents a small fraction of the work done in understanding the liquid-glass transition. Other theoretical efforts involve free volume theory [17] and mode coupling theory [18] [19]. Experimentally, the mechanical and thermodynamic

properties are of great interest and are commonly measured via dynamic mechanical analysis (DMA) and differential scanning calorimetry (DSC), respectively. These measurements capture a multitude of out-of-equilibrium effects, such as hysteresis and nonexponential relaxation, in the vicinity of the glass transition temperature, T_g . Heat capacity measurements made through DSC are one of the most cited techniques for identifying T_g . C_p curves as a function of temperature will exhibit a drop as the sample is quenched, and the inflection point of this heat capacity curve corresponds to T_g . After the rapid quench, and if the sample is reheated, hysteresis effects cause the heat capacity to take a slightly different path in the vicinity of the inflection point before C_p increases to its value in the liquid state. In addition, the thermal expansion coefficient also shows hysteresis upon reheating and cooling of the sample. While a reduction in C_p upon cooling is usually evidence of a thermodynamic (first order) phase transition, this is not the case with vitrification. Instead, the reduction in C_p may be attributed to a reduction in configurational entropy which consequently freezes degrees of freedom so that they may no longer contribute to the heat capacity.



Figure 1.1: This sketch shows the typical hysteresis behaviour of the heat capacity of vitrified liquids. The inflection point is defined to be the glass transition temperature.

The relaxation time can be measured from the viscosity and the shear modulus, via $\tau = \eta/G_{\infty}$ [20], refractive index measurements [21], Brillouin scattering studies [22], and others. There are a number of equations which link together the viscosity, relaxation time, shear modulus, fragility, and other quantities of interest [18]. Among the most commonly used is the Vogel-Fulcher equation which is typically successful in accounting for small to intermediate deviations from Arrhenius behavior [18]. The extent of this deviation can be seen with Angell's classification of strong and fragile glasses (see figure 1.2), with the former representing Arrhenius glasses and the latter non-arrhenius glasses. This classification provides an excellent framework in which to model different glass forming materials. In particular, all glasses belonging to a specific category will generally exhibit the same relaxation functional form, mechanical properties, and thermodynamic properties. For instance, the Vogel-Fulcher equation tends to fail for fragile glasses, but successfully models strong glasses. In addition, fragile glasses will experience a greater deformation to a shear stress than strong glasses, and fragile glasses. The extent to which a glass is deemed fragile may be quantified by its fragility index [23],

$$M = \left(\frac{\mathrm{dlog}\eta}{\mathrm{d}T_g/T}\right)\Big|_{T=T_g},\tag{1.1}$$

which is the derivative of the log of viscosity with respect to T_g/T and evaluated at the glass transition temperature. One can see from figure (1.2) that the slope near T_g becomes large for fragile glasses, but remains invariant for strong glasses. Thus, large values for equation (1.1) indicate fragile glasses.

Before proceeding, it should be noted that the determination of T_g , and the related mechanical and thermodynamic quantities is not absolute. Kauzmann pioneered the recognition that the liquid-glass transition is actually a failure of the experiment for not making measurements on the appropriate time scales. Extending the time scale of the experiment, then, will often "improve" the data since the system has been given additional time to relax to equilibrium. The data that is measured in experiment



Figure 1.2: This sketch demonstrates the strong/fragile classification of glass forming materials. Linearity indicates arrhenius behaviour and deviations from linearity indicate non-arrhenius behaviour (or non-exponential relaxations).

depends on the cooling rate, so the determination of T_g is ultimately arbitrary. For consistency, most researchers use a cooling rate of 10 K/min. This commonly used value has resulted in general agreement of T_g to within 2 K among different labs [20]. While the exact value of T_g may be arbitrary, there is still a maximum and minimum allowable T_g . This is because T_g asymptotically approaches an upper bound and lower bound as quench rates approach infinity and zero, respectively [24]. The time scale dependence of the liquid-glass transition will be an important consideration later on in this work.

Certainly, a great deal of effort has been expended to characterize and to identify the mechanism of the glass transition; yet, there still remain many questions after decades of study. The central question in this work is "to what extent does rigidity contribute to structural arrest in the liquid-glass transition of fragile glasses?" In the case of strong glass formers such as chalcogenide glasses, there is overwhelming theoretical and experimental evidence that network rigidity is a key factor. Will the same be true of multi-component solutions with hydrogen bonding (H-bonding) and ionic interactions?

Previously, ternary sugar/salt solutions of trehalose, choline monohydrogen or dihydrogen phosphate (CMHP or CDHP), and water as well as binary sugar/salt mixtures of trehalose and sodium or phosphate have been studied in molecular dynamics (MD) simulation because of their promising biological applications [25] [26]. Interest in these systems, however, extend beyond the preservation of biologics because they pose intriguing questions in fundamental physics. In ternary systems containing water, it was discovered that the percolation threshold of water is a lower bound to the glass transition concentration, C_g . Interestingly, the oft-disputed free volume theory accounts well for this observation because it considers the percolation of free volume to be responsible for the onset of the glass transition. The problem encountered with free volume theory is the difficulty in defining free volume, but this was circumvented by regarding all water molecules as liquid-like in CDHP systems, while only water molecules which do not neighbor phosphate are considered liquid-like in CMHP systems.

There has been considerable interest in the glass transition of trehalose/water solutions since Angell first observed that the cryoprotective ability of various disaccharides exactly matched the ranking of their glass transition temperatures. Like any glass, there is experimental evidence for a broad spectrum of relaxation time scales involving trehalose-solvent interactions. The mean squared displacement (MSD) in such systems can be decomposed into vibrational and relaxational components, $\langle u^2 \rangle = \langle u^2 \rangle_v + \langle u^2 \rangle_r$, which represents dynamics on fast and slow time scales, respectively [27] [28]. In this paradigm, particles which jump rapidly between two sites of different free energy constitute the vibrational term, while the relaxational term arises when the jumping process exceeds a certain displacement. While this model is corroborated by experiment, there is still a continuous range of time scales which are relevant in the glassy state. This observation is important for considering how empirical measurements of glass depend on time scale. The work in CMHP and CDHP systems already gave hints that this would be a good direction in which to proceed, as evidenced by volume hysteresis effects observed in CDHP and CMHP systems that were attributed to frozen-in strain in the trehalose matrix. This work will extend the previous studies of CMHP and CDHP systems by analyzing their rigidity properties. However, complications arise in applying TCT to fragile glass forming materials with H-bonding. The ambiguity inherent in the definition of an H-bond requires us to regard it as an adjustable parameter. Unlike the covalent bond network that is quenched, the network formed by H-bonds is not static but is subject to thermal energy fluctuations. The problem of accounting for the formation and breaking of Hbonds will be addressed through a careful examination of H-bond lifetimes. This work differs in significant ways from previous applications of the PG. The PG is typically run on lattices that have undergone random site or bond dilution until a specified fraction of sites and bonds remain. This approach fails to account for the tendency of glasses to self-organize in order to minimize energy cost, so self-organization or equilibration algorithms have been developed to remedy this [12] [10] [29]. A natural extension of this line of research is through MD simulation so that we let "nature" equilibrate the lattice. The PG has already been applied to MD-equilibrated systems in two dimensions [30], to strong glasses in three dimensions [31] [32], and to protein systems with H-bond networks, but to our knowledge, it has never been applied to bulk multi-component H-bonded systems in three dimensions.

We propose to address these problems by incorporating a time scale dependence into the PG by time averaging over the recent past with a sliding window. Whereas all bonds are regarded as "equally rigid" in a covalent bond bending network, this approach, which uses the probability that an H-bond is present, makes it possible to account for the heterogeneity of H-bonds with respect to the time scale on which observations are made. Because all H-bond data originates from MD simulations, the spatial-temporal correlations which are experimentally present in glass networks may now be reproduced within the theoretical framework of rigidity theory.

1.1 The Ergodic Hypothesis

There are many ways of characterizing and defining a glass. At a minimum, a glass must lack long-range order, as determined by XRD. One may also rely on mechanical measurements, to wit, a glass must have a viscosity greater than 10^{12} poise [17]; it must experience a sharp drop of the shear modulus at the glass transition [24]; and a glass experiences large amplitude molecular motion on picosecond time scales, while the macroscopic response is on a time scale of milliseconds to seconds [24]. Perhaps the most fundamental feature of a glass is that a glass violates the ergodic hypothesis. Originated by Boltzmann, the ergodic hypothesis states that the time average of an observable equals its ensemble average in the limit that time goes to infinity [33]. An ergodic system will explore the entire phase space after a sufficiently long amount of time, while systems which are not ergodic will remain confined to a "compartment" within the phase space. This difference can be envisioned by considering many different independent microstates (which occupy a compartment within the phase space) of the same system and taking this set of microstates as a set of initial conditions. In an ergodic system, each of these initial microstates are free to evolve (and will evolve) into any of the other microstates given enough time. In a non-ergodic system, each of the initial microstates are trapped within their respective compartments and are separated from other microstates by an energy barrier.

This description is reminiscent of how one might describe a glass [3] [18], and indeed, the notion of ergodicity-breaking is quite useful in understanding the glassy state. Ergodicity-breaking can be invoked in free volume theory which states in simple terms that a glass network experiences a reduction in mobility because there is not a percolating cluster of free volume (liquid-like molecules) which allows molecules to freely move past each other without an energy cost [17] [25]. Likewise, from the perspective of rigidity theory, ergodicity is broken due to formation of large, extended rigid clusters, and in particular, a percolating rigid cluster which prevents the system from exploring its phase space. At a conceptual level, free volume theory and rigidity theory (as applied to glass networks) complement each other, for the former attributes the glass transition to the loss of percolation of free volume, while the latter attributes the glass transition to the percolation of a rigid cluster of molecules. Using both viewpoints, one arrives at the same conclusion that there is a threshold at which degrees of freedom become frozen out and the ergodic hypothesis is thus violated.

1.2 Molecular Dynamics Simulation

MD simulations solve the many-body problem of an ensemble of molecules by applying Newton's equations of motion. As opposed to Monte Carlo methods which only compute the equilibrium properties of a system, MD simulations reveal the time evolution of a system as well. The leapfrog or Verlet algorithms are usually used to numerically solve the equations of motions in MD simulations because they provide stable calculations over an indefinite amount of time and are relatively fast. According to the Verlet algorithm, the position and velocity for each particle in the system is updated as [34],

$$x_i(n+1) \approx 2x_i(n) - x_i(n-1) + a_{i,x}(n)(\Delta t)^2$$
 (1.2a)

$$v_{i,x} \approx \frac{x_i(n+1) - x_i(n-1)}{2\Delta t},\tag{1.2b}$$

where x_i , $v_{i,x}$, and $a_{i,x}$ are the position, velocity, and acceleration of the i^{th} particle, respectively. The time step, Δt , is usually small (on the order of femtoseconds) in order to obtain physically realistic trajectories. The acceleration is calculated from the gradient of the potential energy which is divided into bonded and non-bonded terms. The bonded terms are divided into bond stretching, bond bending, and dihedral angle forces, while the non-bonded terms are the short-range Lennard-Jones interactions and long-range coulombic force. The general form of the potential is given by [35],

$$U(\mathbf{r}^{N}) = \sum_{d} k_{d}(d - d_{0})^{2}$$

$$+ \sum_{\theta} k_{\theta}(\theta - \theta_{0})^{2}$$

$$+ \sum_{\chi} (1 + \cos(n\chi - \delta))$$

$$+ \sum_{\phi} k_{\phi}(\phi - \phi_{0})^{2}$$

$$+ \sum_{i,j} \epsilon_{i,j} \left(\left(\frac{r_{i,j}^{0}}{r_{i,j}}\right)^{12} - \left(\frac{r_{i,j}^{0}}{r_{i,j}}\right)^{6} + \frac{q_{i}q_{j}}{\epsilon_{l}r_{i,j}} \right), \qquad (1.3)$$

where d is the distance between two covalently bonded atoms, θ is the angle associated with three covalently bonded atoms, χ is the proper dihedral angle, and ϕ is the improper dihedral angle. The last term contains the Lennard-Jones and Coulombic interactions where $\epsilon_{i,j}$ is related to the Lennard-Jones potential well depth, $r_{i,j}$ is the separation distance between two atoms, and $r_{i,j}^0$ is the equilibrium separation. The last term in the last sum is the Coulombic potential where q_i and q_j are the charges of atoms i and j and ϵ_l is the effective dielectric constant. The first, second, and fourth sums come from Hooke's law, while the third sum models the potential energy of a dihedral angle between two planes which are free to rotate.

Most experiments seek to ascertain the glass transition temperature by rapidly quenching the sample. Indeed, a rapid quench is necessary in order to form a glass, but the simulations performed in this work are at constant temperature. Instead of running MD simulations which have different initial and final temperatures to form the glass, the glass is formed during energy minimization which represents a quench from infinite temperature. This is because the starting the configuration is generated by randomly placing molecules in the simulation box. Molecules which have an equally likely chance of occupying any location, irrespective of their neighboring molecules, can be said to be at a very high temperature. The role of energy minimization is to reshape the potential energy surface by rearranging the molecules so as to achieve a starting configuration which is physically realistic. Without energy minimization, the system will "explode" due to the unphysical starting configuration which could only exist at very high temperatures. Energy minimization iteratively adjusts atomic coordinates until the maximum force falls below an acceptable threshold, and the resulting system is in a local energy minimum. The ability to escape from that minimum during the following NPT and NVT equilibrations will determine if the system is a glass or liquid (non-ergodic or ergodic).

CHAPTER 2: Theory

This chapter will provide a brief and cursory exposition of the relevant theory. First, Landau's theory will briefly be discussed, and the connection will be made to percolation theory. Then, rigidity theory will be introduced which is necessary for understanding the pebble game algorithm.

2.1 Theory of Phase Transitions

Among the first phase transitions discovered which do not involve a latent heat, but still exhibit critical phenomena, is that of liquid helium when it transitions to a superfluid at a critical temperature. This anomalous result prompted Ehrenfest to devise a classification scheme of phase transitions which are defined based on the existence of singularities in the derivatives of free energy [36]. Soon thereafter, Landau generalized this idea to include any transition between a symmetrical and unsymmetrical phase [37]. Landau's phenomenological theory expresses the free energy as a power series in even powers of an order parameter. The order parameter may itself be a function of temperature or some other quantity which has a critical value that distinguishes between two phases. In this theory, the order parameter, m_0 , is zero in the symmetric phase and non-zero in the unsymmetric phase. The manner in which m_0 approaches zero is of great importance, as this will determine crucial properties of the system, such as the universality class to which a transition belongs in the case of second order transitions. When m_0 discontinuously approaches zero, the transition is said to be first order, while the transition is second order if it continuously approaches zero. The first order phase transition is the "usual" transition which involves a latent heat. In contrast, the thermodynamic potentials vary continuously over the transition point in a second order transition and therefore involve no latent heat. The free energy can be expanded as,

$$\psi(t, m_0) = q(t) + r(t)m_0^2 + s(t)m_0^4 + \dots, \qquad (2.1)$$

where $t = (T - T_c)/T_c$, and $|t| \ll 1$. The coefficients are themselves power series in t,

$$q(t) = \sum_{k>0} q_k t^k \tag{2.2a}$$

$$r(t) = \sum_{k>0} r_k t^k \tag{2.2b}$$

$$s(t) = \sum_{k \ge 0} s_k t^k \tag{2.2c}$$

but only the first one or two terms will be retained in the asymptotic expression for free energy. Equation (2.1) is an expansion of the free energy around the critical point, so it is only valid for small deviations of temperature from the critical temperature. The first order term is absent because it represents the perturbation Hamiltonian associated with an ordering field, but equation (2.1) assumes there is no external perturbation. The equilibrium value of the order parameter is found by minimizing the free energy by taking the derivative of ψ with respect to m_0 . Doing so yields,

$$r(t)m_0 + 2s(t)m_0^3 = 0. (2.3)$$

There are only two physical solutions: one corresponds to the symmetrical phase, $m_0 = 0$ (t > 0), and the other corresponds to the unsymmetrical phase,

$$|m_0| = \left(\frac{-r(t)}{2s(t)}\right)^{1/2} \approx \left(\frac{r_1|t|}{2s_0}\right)^{1/2}, \tag{2.4}$$

for t < 0. The exponent in equation (2.4) is physically significant and is known as a critical exponent because it describes the power law behavior in the vicinity of the second order transition. In particular equation 2.4 gives the β critical exponent which describes order parameter in the vicinity of the transition.

2.1.1 Connection to Percolation Theory

Consider a periodic lattice where each site on this lattice may either be occupied or unoccupied with a probability p. On this lattice, clusters of occupied sites will form. A cluster is defined as a continuous set of occupied sites such that there exists at least one path between any two sites in the cluster which does not pass through an unoccupied site. The aim of percolation theory is to mathematically characterize the clusters which form as the lattice topology and probability are varied.

Percolation theory is built on simple axioms, but in terms of characterizing a phase transition it bears great similarities to Landau theory. As the site occupation probability increases, the cluster sizes in the lattice will also increase, and there is a critical probability, p_c , where the lattice will experience a transition. When $p < p_c$, all of the clusters are finite, but when $p > p_c$, there exists an infinite (percolating) cluster (or a cluster which spans the length of the lattice in the case of a finite lattice). The range of probabilities around p_c will accordingly display critical phenomena.

The cluster number, n_s , is defined as the probability that a cluster of size s exists in the lattice. Percolation theory is typically concerned with the scaling relations of the k^{th} moment of n_s ,

$$M_k = \sum_s s^k n_s, \tag{2.5}$$

and in particular, the first three moments are given considerable interest. M_0 , M_1 , and M_2 , represent the number of clusters, the strength of the infinite cluster, and the average cluster size, respectively. The strength of the infinite cluster is denoted by P and represents the probability for any arbitrary site to be in the infinite cluster, while the mean cluster size is, $S \propto M_2$.

The manner in which the order parameter, P, goes to zero at the percolation threshold depends on the scaling exponents which usually belong to a university class that is in turn determined by the dimensionality of the lattice. One lattice which admits exact solutions is the Bethe lattice which is a tree where each site has the same coordination number, z. For the Bethe lattice when z = 3, it can be proven that $P \propto (p - p_c)$, giving a critical exponent of $\beta = 1$. The Bethe lattice is amenable to exact analytical solutions because it has no closed loops. Therefore, each branch of the lattice is statistically independent of other branches, and the probability of a site not being connected to a branch in the percolating cluster can be calculated. The Bethe lattice is a special case because it is infinite dimensional. For finite dimensions, the order parameter has the general form,

$$P \propto (p - p_c)^{\beta}. \tag{2.6}$$

The discussion thus far has focused on connectivity (or scalar) percolation. Rigidity (or vector) percolation implies connectivity percolation, but also imposes an additional requirement. This will be discussed in the next section. It is also important to note that generic rigidity percolation must not take place on a geometrically periodic lattice (topologically periodic is okay). Such non-periodic systems are convenient to study computationally.

2.2 Rigidity Theory

Quantitatively characterizing rigidity in networks dates back to Maxwell who was interested in assessing the structural stability of trusses [23]. In the theory of network rigidity, one considers an idealized framework of completely rigid bars which are connected at frictionless joints so that the only possible motion arises from rotations about a joint or bar. Given an infinite lattice in two dimensions, it is not always obvious if deformations of the framework are permitted. By considering a d-dimensional bar and joint framework with N joints (nodes) and N_c bars (constraints), Maxwell demonstrated that the number of floppy modes is given by $F = dN - N_C - d(d+1)/2$. For d = 2, such a framework becomes rigid when $N_c = 2N - 3$ [8]. This result is exact for a homogeneous lattice, but it is approximately true if each joint has on average N_c constraints. This mean field approximation is known as Maxwell counting.

Colloquially, the concept of rigidity is self-explanatory and implies some resistance to elastic deformation. Mathematically, rigidity is a well defined concept by employing linear algebra. For certain types of rigidity (first order stable and structures with generic positions of joints) the linear algebra can be formalized in terms of constraint counting using graph theory in two dimensions. A graph, G, consists of a vertex set V and an edge set E which is comprised of unordered pairs from V. A framework, (V, E, \mathbf{p}) , may be constructed by embedding G into an m dimensional Euclidean space such that each vertex is assigned a coordinate in space using $\mathbf{p} = \mathbf{p}_1, ..., \mathbf{p}_n$. A deformation or finite flexing of the framework is a continuous family of realizations such that the coordinate of each vertex is a differentiable function of time and the distance between two edge sharing vertices at \mathbf{p}_i and \mathbf{p}_j remains fixed, according to [38]

$$(\mathbf{p}_i(t) - \mathbf{p}_j(t)) \cdot (\mathbf{p}_i(t) - \mathbf{p}_j(t)) = c_{ij}.$$
(2.7)

A deformation of the framework is said to be trivial if the distance between any two points remains fixed regardless if they share an edge. If all deformations are trivial, the framework is said to be rigid. These trivial motions correspond to one rotational and two translational degrees of freedom in \mathbf{R}^2 and three rotational and three translational degrees of freedom in \mathbf{R}^3 . It is often easier to work with the derivative of equation 2.7,

$$(\mathbf{p}\prime_i(t) - \mathbf{p}\prime_j(t)) \cdot (\mathbf{p}_i(t) - \mathbf{p}_j(t)) = 0, \qquad (2.8)$$

whence one may define the notion of infinitesimal rigidity and introduce the rigidity The solutions to equation 2.8 define the set of all infinitesimal motions, matrix. and a framework is said to be infinitesimally rigid if there exist only m(m+1)/2solutions (corresponding to the rigid body motions). The rigidity matrix has E rows and mV columns and thus encapsulates the relative locations of all vertices and their connections within a framework. Each row corresponds to an edge and has 2m nonzero entries which are the difference in coordinates of the adjoining vertices. The number of linearly independent edges is the rank of the matrix. An edge is said to be redundant if the rank of the rigidity matrix is invariant upon its removal, while an edge is said to be independent if removing it decreases the rank. The foregoing linear algebra approach will always be sufficient to determine the rigidity of a framework. A graph, unlike a framework, contains only connectivity information and requires a graph-theoretic approach. A graph is generically rigid in *m*-dimensional space if there exists an embedding **p** such that all frameworks, (V, E, \mathbf{p}) , in the neighborhood of **p** are infinitesimally rigid. Therefore, infinitesimal rigidity implies generic rigidity.

The concept of an independent or redundant edge is integral to the analysis methods presented here, but they are defined within a somewhat abstruse formalism. In two dimensions, the difference between a redundant and independent edge can be easily seen with a few illustrative examples. Figure 2.1 shows some of the possible generic graph representations for a graph with four vertices. Panels (a) and (b) both have more than zero floppy modes (not including the three rigid body motions in the plane), while (c) and (d) have zero floppy modes and are therefore rigid. The redundant constraint is specified in panel (d) as the heavy dashed line, but in principle one could



Figure 2.1: Panels (a) through (d) show the evolution from floppy to rigid of a simple generic graph. In panel (a), there are two independent motions associated with the horizontal bars which can pivot around the joints connected by the vertical bar. The addition of an independent constraint in (b) removes one floppy mode. The addition of the dashed edge in (c) now turn the four vertices into a single rigid cluster. The thick dashed line in (d) is redundant because the graph has already undergone a rigidity transition.

remove any edge without adding floppy modes. The heavy dashed line is identified as redundant because it was the last edge added to the graph. In the mean field approximation, the number of floppy modes is found with, F = mN - C, where Nis the number of nodes in the graph and C is the number of constraints. Applying this equation to the graph realizations in figure 2.1 (and subtracting out three trivial degrees of freedom), one can easily calculate that (a) has two floppy modes, (b) has one floppy mode, and (c) has zero floppy modes. While calculating (d), however, one can see that this equation has an obvious flaw because the constraints cannot be increased

after the isostatic point, for then F becomes negative. This problem is addressed in the mean field approximation by taking max(F,0) which is reasonably accurate if constraints are approximately homogeneously distributed. Strictly speaking, finding max(F,0) is a combinatorial problem which must be applied to all subgraphs, but it can be applied globally if most constraints are uniformly distributed. If a precise count of floppy modes is desired, one must proceed by identifying the R redundant constraints in the framework so that the number of floppy modes becomes F = mN - mNC+R. Numerical methods exist for this purpose, but they usually involve calculating the rank of the rigidity matrix which can become computationally expensive for large frameworks. The PG algorithm was designed to solve this problem. The PG is built on Laman's theorem in graph theory which states that a graph embedded in two dimensions does not have a redundant bond iff no subgraph containing n sites and b bonds violates $b \leq 2n - 3$. Thus, any graph with more than two nodes can be recursively checked for redundant bonds. Because this result is true for a generic framework, it is not necessary to embed a graph in \mathbf{R}^2 and compute a rigidity matrix. It is only necessary to know the topology of a generic graph representation which makes the PG is substantially faster than other algorithms. The mechanics of the PG will be discussed in the methods section.

Covalent bond bending networks may be modeled with central force constraints having a harmonic potential, and bond bending constraints associated with the dihedral angle twist which have a Kirkwood or Keating potential. From the perspective of the PG, the central force covalent bonds form the graph G, and the bond bending constraints (not explicitly in G) are found in G^2 . A body-bar multi-graph can be constructed from nodes which have six degrees of freedom (representing bodies) which are connected by sets of five bars (instead of one bar). It has been proven that the body-bar multi-graph G is isomorphic to the graph G^2 . This convenient result makes it possible to construct the square of G without having to manually input the edges connecting next nearest neighbors. Moreover, this mapping will enable the assignment of fractional weight to central force connections in the VPG.

One of the most important quantities in rigidity percolation is the rigid cluster susceptibility (RCS), for this indicates the bond or site concentration where the lattice transitions from floppy to rigid. This transition occurs when a percolating rigid cluster spans the system. Following standard percolation theory, the RCS is expressed as [39],

$$\chi = \frac{1}{N_c - 1} \sum_{c=1}^{N_c - 1} (N_v(c) - \langle N_v \rangle_{reduced})^2,$$
(2.9)

were N_c is the number of rigid clusters, $N_v(c)$ is the number of vertices within the c-th cluster, and $\langle N_v \rangle_{reduced}$ is the average cluster size, subtracting out the largest cluster. Likewise, the largest rigid cluster is not included in the sum. The rigidity transition can be located by plotting equation (2.9) as a function of the fluctuating edge probability or another parameter which controls the connectivity and edge weight (e.g., concentration). The peak in this plot corresponds to the rigidity transition. A similar quantity is defined in the case of connectivity percolation. The effectiveness of equation (2.9) in identifying the existence of a percolating cluster can be understood by considering the cluster distributions before, during, and after the transition. Before the transition, there are many small clusters distributed throughout the network, but none are percolating so that the variance in cluster size is close to zero. During the transition, there is a more heterogeneous distribution of cluster sizes (and therefore a large variance), and the largest clusters are "competing" to be the percolating cluster. After the transition, all of the largest clusters have been consolidated into the percolating cluster which is now embedded in a "sea" of floppy inclusions made of smaller clusters. Because equation (2.9) subtracts out the largest cluster, the variance in the cluster size again becomes approximately zero.

While locating the peak in the RCS gives an accurate estimate of the rigidity

transition, Maxwell counting is still a useful approximation which can be used as a reference for the results of the PG. Three dimensional chalcogenide glasses have the form $\text{Ge}_x \text{As}_y \text{Se}_{1-x-y}$ where Germanium, Arsenic, and Selenium represent atoms which are 4-, 3-, and 2-fold coordinated, respectively. For an *r*-fold coordinated atom, each central force constraint is shared by two atoms giving r/2 central force constraints, and there are 2r - 3 angular constraints. The total number of constraints is [7],

$$n_c = \sum_{r=2}^{4} n_r [r/2 + 2r - 3], \qquad (2.10)$$

where n_r is the number of r-fold coordinated atoms. The total degrees of freedom in a chalcogenide glass with N atoms is 3N, so Maxwell counting predicts the number of floppy modes to be (normalized by degree of freedom),

$$f = \frac{3N - \sum_{r=2}^{4} n_r [r/2 + 2r - 3]}{3N}.$$
 (2.11)

Simplifying this equation, one obtains $f = 2 - 5\langle r \rangle/6$. The lattice becomes globally rigid when f = 0 which implies $\langle r \rangle = 2.4$ at the rigidity transition. Typically, there are correlations in the lattice which cause the transition to occur slightly below 2.4.

TCT agrees very well with experiment [40] [41], and it has become a useful tool because it offers a simple way to conceptualize the loss of ergodicity in the glassy state by balancing degrees of freedom with constraints. TCT is a course-grain approach which subsumes the complexities of the system under a single parameter – namely, the mean coordination number which acts as a topological order parameter. A more comprehensive approach would involve investigating the potential energy hypersurface or using the full machinery of statistical mechanics. TCT demonstrates that it is possible to ignore these complexities while still arriving at a physically coherent (but incomplete) description of the glassy state.

An objection which is commonly raised against TCT is that it is valid only at

zero Kelvin (all constraints are unbroken). Several methods have been proposed to include a temperature dependence in TCT, the simplest of which uses a twostate thermodynamic function, q(t), so that all bonds are intact at low temperature (q(t) = 1), and all bonds are broken at high temperature (q(t) = 0). By treating the thermal fluctuations which break and form bonds as an equilibrium reaction, a form for q(t) at finite temperature may be derived [23],

$$\ln K_e = \frac{\Delta G^*}{T} = \ln \frac{q(t)}{1 - q(t)},$$
(2.12)

whence,

$$q(t) = \frac{e^{\Delta G^*/T}}{1 + e^{\Delta G^*/T}},$$
(2.13)

where ΔG^* is the equilibrium Gibb's free energy, and K_e is the equilibrium constant.

Other attempts to expand the scope of TCT include incorporating the effects of water [42]. In silicate glass networks, it has recently been suggested that water cannot percolate rigidity [41], so water must act as a network modifier. This occurs when water reacts with silica and forms hydroxyl groups between network forming atoms thereby breaking constraints in the network. However, in networks with proteins, water molecules bonded to the surface of proteins may be rigid. In this work, it will be shown that water molecules can act as rigid constraints in network glasses, but there is a dependence on time of observation and time correlations within the H-bond network.

CHAPTER 3: Methods

3.1 Simulation Protocols

The simulation protocols are similar to those that were used in the work of Nikulsin, et al [25]. CDHP and CMHP systems were studied at 270 and 300 Kelvin each in a 2:1 stoichiometric ratio of trehalose to salt. All simulations used GROMACS and the OPLS-AA forcefield in conjunction with the Transferable Intermolecular Potential 4-point (TIP4P) water model. A time step of 2 fs is used, and data is saved every 3000 steps for the data collection runs and 500 steps for the equilibration runs. The systems were initialized with Packmol which randomly placed all molecules within a 1331, 1728, or 2197 nm^3 simulation box (corresponding to side lengths of 11, 12, or 13 nm, respectively). The number of molecules for each concentration is determined by requiring that the sum of the specific volumes of all molecules is approximately 1000 nm³. The specific volumes were calculated from experimental densities. Following Packmol, an energy minimization was performed with a steepest descent algorithm until the greatest force in the system converged to below 500 kJ mol⁻¹ nm⁻¹. If the energy minimization trapped the system in a local minimum without converging, the simulation is restarted from the initial step with Packmol. Following the energy minimization, an NPT (constant pressure) equilibration was performed by coupling the system to a velocity rescaling thermostat set to either 270 K or 300 K with a coupling constant of 0.1 ps. The system was coupled to a constant pressure bath via a Parrinello-Rahman barostat set to 1 bar with a coupling constant of 2 ps or 10 ps for the 2016 and 2020 versions of GROMACS, respectively and an isothermal compressibility of 4.5×10^{-5} bar⁻¹ (the isothermal compressibility of water). This equilibration was performed until the volume converged. The NPT equilibration is
said to converge if a linear regression of the last 100 ps of volume data has a slope whose absolute value is less than or equal to $0.02 \text{ nm}^3 \text{ ps}^{-1}$. If this condition is not met, another simulation is started which uses the final condition of the previous simulation as its initial condition. In this manner, the simulation can be extended until convergence is reached. After convergence, the barostat was removed and an NVT (constant volume) equilibration run was performed for 3 ns. The final data was collected from a 30 ns production run. The pressure coupling constant was again chosen to be 2 ps or 10 ps depending on the version of GROMACS.

3.1.1 Volume Hysteresis

Due to the presence of non-ergodicity, hysteresis is a hallmark of the glassy state [24], [33]. In CMHP and CDHP systems, a volume hysteresis effect has been observed which has been used to identify the glass transition [25]. The volume hysteresis method works as follows. During equilibration in the NPT ensemble, the final volume (after an elapsed time) is dependent on the initial volume if the system is a glass, but this is not true in the liquid state. That is to say, a system is a liquid when many different, independent initial microstates at different volumes are prepared, and each of these microstates converge to the same macrostate at the end of the equilibration (e.g., they converge to the same volume). If the system is a glass, then the system will converge to different final volumes for different initial volumes. This volume relaxation or hysteresis effect is analogous to what is observed experimentally in dynamic mechanical analysis.

Moreover, hysteresis is one implication of non-ergodic systems. Such systems are history dependent because the set of possibilities for the current microstate is limited by the initial microstate. This occurs when a system is confined to within a compartment of its phase space. Thus, when many different samples of the same system are initialized to different volumes, each of these initial configurations remain separated in phase space for the duration of the simulation.

3.2 Hydrogen Bond Analysis

The nature of hydrogen bonding has been an elusive and controversial subject for decades [43]. Only recently has it generally become accepted that hydrogen bonds are inherently quantum mechanical and not only electrostatic interactions [44] . However, it is often sufficient to regard them as electrostatic interactions of varying strength. This is a useful approximation to make as most MD simulations only consider Lennard-Jones interactions and coloumbic forces when applying the equations of motion. Experimental techniques to analyze H-bond networks have steadily progressed and continue to yield new insights [44]. XRD data are sensitive to the positions of electronegative acceptors and donors [45]. Neutron diffraction and Raman scattering respectively give information about the positions of individual nuclei, and the local geometry of molecules participating in H-bonds [46]. In studies involving water, it is common to use deuterium substitution in combination with Raman scattering in order to probe the subtle local structure of water and its interactions with solute. X-ray absorption experiments excite core electrons of donor or acceptor atoms to probe the local micro-environment. Empirical potential structure refinement is a computational approach which works by iteratively adjusting partial structure factors until a fit to experimental data is achieved in order to make an improved estimate of the potential from the forcefield in MD simulations [47]. This technique has been used to probe fine grain water-trehalose interactions. Clearly, a great deal of effort has been invested in just the study of pure water (and interactions with solute) which still remains somewhat of a mystery. It is then not difficult to understand and appreciate that studies of multi-component solutions are experimentally onerous.

The aforementioned experimental methods are useful for a static analysis of hydrogen bond networks such as understanding the geometrical arrangement of molecules participating in H-bonds. Understanding the dynamics of H-bonds remains an experimental challenge [48]. Some efforts include the measurement of the mean square



Figure 3.1: θ is the HDA angle. H-Bonds are restricted within a cone with cutoff angle θ_c . HDA angles outside of this cone are not considered H-Bonds.

displacement which can be measured indirectly via the Debye-Waller factor in neutron scattering [49], the re-orientational auto-correlation function which can be measured from NMR, and ion-solvation shell lifetimes which have been measured with ultrafast mid-infrared spectroscopy [50]. Experiments which make indirect dynamical measurements tend to be sources of controversy because of the difficulty of interpretation. In light of the experimental difficulty in discerning the complex interactions involving H-bonds in multi-component solutions, the utility of MD simulation becomes readily apparent.

There exist many attempts to characterize and define hydrogen bonds in the literature including geometrical, topological, energetic, and occupancy definitions [44] [51] [52] [43]. This study uses a geometrical definition with both an angle and distance cutoff and is shown in figure 3.1. The cutoffs (discussed later) are in general agreement with those reported elsewhere [53] [51]. Rigidity percolation is a long range phenomenon and can suddenly occur with the placement of just one bond, so it is necessary to have H-bond definitions which are representative of the underlying system dynamics. Increasing an angle cutoff by just a few degrees can allow many more H-bonds to be included and therefore radically shift the rigidity transition. There is evidence that all hydrogens participate in a bond, and definitions have been proposed to record all such bonds [52]. Intuitively, the weak H-bonds would not appreciably act as a hard constraint for network rigidity, and it is more sensible to ignore them. Thus, a geometric criteria provides a simple and rational way to control the selection of relevant H-bonds while filtering out weak H-bonds which are too far from linearity. Moreover, it has been shown that a geometrical criterion gives better agreement with experimental H-bond lifetimes than an energetic criterion [54]. Though HA distances and HDA angles tend to be highly correlated [55], one must still include a reasonable distance cutoff in order to avoid sampling distant hydrogen bonds which happen to be close to linearity. Arbitrary cutoffs are undesirable in this study since one could choose where the rigidity transition occurs, irrespective of the physics which drives the transition.

Table 3.1: Listed in the table are the H-bonds considered in the system. Donors are rows, and acceptors are columns. "Y" means that the H-bonds between those species are counted, while "N" means they are not counted.

	Choline	Phosphate	Trehalose	Water
Choline	Ν	Y	Y	Y
Phosphate	Ν	Ν	Ν	Ν
Trehalose	Y	Y	Y	Y
Water	Y	Y	Y	Y

Before performing any calculations, it is wise to rule out those species which cannot participate in an H-bond. The nitrogen atoms in choline are bonded to four carbons and cannot form H-bonds. All H-bonds are therefore between oxygen donors and acceptors in choline, phosphate, trehalose, and water. Because of the importance of choosing non-arbitrary geometrical cutoffs, each pairwise H-bond was categorized based on the molecules involved. This gives 16 possible combinations of donors and acceptors between the four molecules. One could consider even more types by sub-categorizing oxygens in trehalose and phosphate, but this route quickly becomes bogged down by details which are not necessarily important to rigidity calculations. Table 3.1 summarizes the H-bonds considered in CMHP and CDHP systems.

The hydrogen bond analysis given here has been reported elsewhere in the literature using the reactive flux method (RFM), but the current work differs in some important ways. First, as opposed to the RFM, we only define a simple hydrogen bond correlation function [56],

$$C(t) = \frac{\langle h(0)h(t)\rangle}{\langle h(0)\rangle},\tag{3.1}$$

which has the usual definition that h(t) = 1 if an H-bond is present and h(t) = 0 otherwise. We also impose no criteria on h(t) which requires, for example, that a bond must first exist for a pre-defined amount of time before it is counted. This is also known as an intermittent correlation function. In this sense, the correlation function defined here is in its simplest form, but this is adequate for the current work since the objective is not to perform a rigorous analysis of the hydrogen bond dynamics. To ensure equation 3.1 is calculated for an equilibrated system, we calculate the correlation functions backwards in time starting from the final frame to the initial frame. Doing this will prevent out-of-equilibrium effects at the beginning of the simulation (e.g. volume relaxation) from dominating C(t). The hydrogen bond lifetime is given as,

$$\tau = \int_0^\infty C(t)dt. \tag{3.2}$$

The calculation of τ is crucial to this H-bond analysis, as it will guide in the determination of a non-arbitrary cutoff and give insight into the apparent persistence times of H-bonds. It must be emphasized that the H-bond lifetimes should not be interpreted as a measure of the relative strengths of different H-bonds because this definition neglects the effect of diffusivity on the persistence of an H-bond. That is to say, large molecules may remain neighboring for a long period of time because of their low diffusivity and not because they are connected by a strong H-bond. Likewise, small molecules with otherwise high diffusivity may appear to have a long lifetime only because they are trapped in cages formed by neighboring large molecules. The RFM is designed to accommodate for the competing effects of diffusivity and hydrogen bond strength which undoubtedly occur in ternary sugar/salt solutions. Nevertheless, the objective here is not to provide a detailed description of the H-bond network but to determine the role of rigidity in driving the glass transition. The lifetime as defined here can be regarded as a first approximation to characterize when a constraint is present or not without distinguishing between the diffusivity and hydrogen bond potential well depth. This is still a good approximation to make because the question of why an H-bond is persistent is not relevant to the rigidity properties. All that matters is that the H-bond exists and restricts the flexibility of the network.

Similar to the procedures used in other work which use the RFM [48] [57], an angle cutoff was determined by plotting H-bond lifetimes on the ordinate against different angle cutoffs on the abscissa. It is hypothesized that these plots should have H-bond lifetimes which are monotonically increasing to a maximum and then monotonically decreasing. This is because H-bonds which are too close to linearity are entropically unfavorable, and a cutoff angle which is too restrictive will neglect persistent Hbonds which ought to be counted. Similarly, a cutoff which is not restrictive enough is energetically unfavorable and will record too many transient bonds. Both of these scenarios will drive down the lifetime; thus the maximum lifetime is a natural indicator of the maximum allowable HDA angle. Data will be shown in the next chapter which confirms this hypothesis. One advantage of the lifetime definition provided here over the RFM is the ease of interpretation which allows one to predict that there should indeed be a maximum lifetime. In order to separate the effects of the angle and distance cutoffs, lifetimes were calculated by varying the angle cutoff with no restriction on distance, and then the angle cutoff was varied with a distance cutoff of 2.9 angstroms. The 2.9 angstrom cutoff was determined by varying the distance cutoff with no angle restriction. These results will be elaborated on in the preliminary data section.

The angle cutoffs for each concentration are determined by performing a weighted average over all of the allowed H-bonds,

$$\theta_c = \frac{\sum_i N_i \theta_i}{\sum_i N_i},\tag{3.3}$$

where N_i is the number of H-bonds recorded for the cutoff angle, θ_i , which maximizes equation 3.2. The sum is over the i^{th} H-bond pair recorded in table 3.1 for a given concentration. The H-bond lifetime over all H-bonds is also calculated as a weighted average,

$$\langle \tau \rangle = \frac{\sum_{i} N_{i} \tau_{i}}{\sum_{i} N_{i}}.$$
(3.4)

H-bond distributions were also found from a single frame at different concentrations. These were informative, but they could not be used to determine a cutoff since we could not rule out underlying distributions resulting from random packing, van der waals interactions, or other effects. They were calculated using a PDF estimator [58] and will be presented in the results section.

3.3 Pebble Game algorithm

The inspiration for the name, "pebble game algorithm", comes from how it operates. In two dimensions, two pebbles are attached to each node (one for each translational degree of freedom) and must remain attached to their initial nodes for the duration of the game. While on a node, a pebble is considered to be free, and it can be placed on any edge incident to that node. Once an edge has been covered by a pebble, a pebble must remain on that edge for the duration of the game. The game is played by trying to cover each independent edge in the graph with one free pebble until all independent edges are covered or no more free pebbles can be found. Before a pebble can be placed on an edge, the edge must be tested for independence. An edge is independent if all four pebbles can be freed in order to cover the edge; otherwise, the edge is redundant. The number of remaining pebbles is the number of floppy modes in the network. Though each pebble must remain tethered to its initial node, a free pebble on one end of the network can be used to cover an edge on the other end by shuffling pebbles along a connecting path while ensuring that each covered edge along that path remains covered. This is possible because each edge has two adjoining nodes, but only one node can donate a pebble to cover the edge. Thus, the donating node can reclaim its pebble if and only if the opposing node can donate one of its pebbles to cover that edge. Figure 3.2 is an illustration of the PG algorithm. In it, a bond is being tested for independence, and a search is conducted



Figure 3.2: A demonstration of the pebble game on a generic network. Independent (redundant) bonds are shown with solid (dashed) lines that are (are not) covered by a pebble. Large (filled, open) circles denote (anchored, free) pebbles on (bonds, sites). The two closest pebbles to a given site are tethered to that site. Small (filled, open) circles denote sites belonging to (one, more than one) rigid cluster. Overconstrained bonds are shown with heavy dark lines. Shaded regions denote 2d rigid bodies. (a) Five free pebbles indicate five floppy modes until a new bond is added and tested for independence. A fourth free pebble is found via the path traced by arrows. (b) The added bond is independent and thus covered. There are now six rigid clusters and four floppy modes. Figure and caption reproduced from reference [1]

to free up a fourth pebble across the test bond. Since a fourth pebble could be found, the bond is independent and then covered with a pebble.

In two dimensions, all floppy motions arise from bars pivoting around vertices in the plane, and the PG is able to detect these floppy modes. The extension to three dimensions is more complicated because then there are floppy modes which arise from rigid clusters pivoting around a shared bar. Panel (c) in figure 2.1, for instance, would not be rigid in three dimensions because there are two rigid clusters (the upper and lower triangle) which can pivot out of the plane of the page about the dashed diagonal line. The addition of the heavy, dashed line in figure (d) would constrain this floppy mode, and the graph in panel (d) is rigid in both two and three dimensions. An edge which is shared by two separate rigid clusters that have no other connecting edges is said to be a hinge joint. It is not hinge joints which cause Laman's theorem to break down in three dimensions; rather, Laman's theorem fails to account for implied hinge joints. An implied hinge joint exists between two vertices not connected by an edge, but which share any two rigid clusters. As such, the two vertices remain a fixed distance apart while the surrounding rigid clusters behave as if they are connected by a hinge joint between the two vertices. This is demonstrated by the so-called "banana graphs" illustrated in figure 3.3. The three dimensional pebble game (with three



Figure 3.3: Figure reproduced from [2]

pebbles per node) would incorrectly predict this banana graph to be rigid. However, there are three rigid clusters which are free to rotate about implied hinges, and the three subgraphs are connected at vertices 1, 2, and 3 to form a fourth rigid cluster. It has been proven that the only violations of Laman's theorem in three dimensions are due to implied hinge joints [2], so the three dimensional PG can still be used under special conditions. While figure 3.3 is an obvious example of implied hinge joints, banana graphs are not so easily identified in general. It would seem that the PG is only useful in three dimensions if it is possible identify those graphs in advance which prohibit implied hinge joints. Fortunately, there exists a class of graphs where this is always true. Given a graph G, its square is denoted by G^2 and includes the vertex and edge set of G plus edges between all next nearest neighbors. Laman's theorem will always extend to three dimensions for this special class of squared graphs. Thus, any graph can be constructed to satisfy Laman's theorem in three dimensions by taking the square of any other graph G with more than two vertices. This may seem like a restrictive requirement with sparse applications, but squared graphs are in fact bond bending networks and are fortuitously used to model covalent and hydrogen bonds.

3.4 Procedure

Rigidity theory has been applied to covalent bond bending networks at zero temperature, and some progress has been made in incorporating a temperature dependence. Recently, the heterogeneity of some covalent bonds has been recognized and accounted for by analyzing the variance of bond angles from MD simulations. However, all work which analyzes glass networks using the framework of rigidity theory has been confined to glasses with covalent bond bending networks. The purpose of the H-bond analysis is to definitively identify all H-bonds and then to appropriately weight them by how rigid they are with respect to a given time scale of observation. This is accomplished by calculating an edge weight for each H-bond in the system such that $0 \le e_{ij} \le 1$ and then performing a Monte Carlo simulation to obtain the average rigidity properties from the generated graph realizations. The edge weight represents the probability of an H-bond existing over a window of time, where the window acts acts as a probe across time scales. A small window size will include the vibrational term of the MSD while a large window size will include both the vibrational and relaxational terms. That is, as the window size becomes smaller, rapidly fluctuating H-bonds will be assigned greater probabilities, and the fast dynamics of the system will be included in the system topology; conversely, a large window size will only include H-bonds with a long lifetime thereby filtering out the vibrational term. Modulating the window size, then, will reveal the time scale of importance in order to affect a rigidity transition.

The procedure can be divided into two parts: (i) raw MD simulation data must be converted to a generic graph representation, and (ii) the PG will analyze this topology to output the rigidity properties. Part (i) is the purpose of the H-bond analysis, while part (ii) will address the central question posed in the introduction.

The first step of part (i) is iterating over the coordinate data frame by frame and finding the nearest neighbors of every H-bond capable atom. The nearest neighbors are found by partitioning the simulation box into 3.5 angstrom boxes and iterating over all atoms within each box. This cutoff is quoted very frequently in the literature and comes from the radial distribution function calculated from XRD. The next step involves iterating over neighboring atoms and checking if they satisfy the distance and angle cutoffs. In this model, the only requirement for an H-bond to exist is that it satisfies the H-bond angle and distance cutoff for at least one frame of the simulation. Once an H-bond between donor atom i and accepting atom j has been detected, $h_{ij}(t_n) = 1$ if the H-bond exists at time t_n and $h_{ij}(t_n) = 0$ otherwise. Therefore, for a hydrogen bond recorded at time step t_n , the edge weight is given by,

$$e_{ij}(t_n) = \frac{1}{w} \sum_{k=n-w}^n h_{ij}(t_k),$$
(3.5)

for a window size of w frames. The step size is one frame as implied by the sum in equation 3.5, but the window size is left as an adjustable parameter in order to probe

the structural relaxation on different time scales.

After the edge weights (H-bond probabilities) have been calculated for every Hbond over all frames under consideration, the inter-molecular connections can then be input into the network. All constraints have five bars, and a Monte Carlo simulation must be performed in which many different graph realizations (microstates) are generated, and the H-bonds for each trial are input with a probability e_{ij} .

The PG which is performed on graphs which are generated from H-bond probabilities according to equation (3.5) shall be referred to as the time-correlated PG (TCPG). Another mean field approximation involves uniformly assigning probabilities to all H-bonds in the system and thus removing any time-correlations introduced by the sliding window. This shall be referred to as the uniform-probability PG (UPPG). The motivation for this is to sweep across a wide range of probabilities which may not be accessible from simulation data in order to identify if the transition is possible, and to approximate the H-bond probability in order to affect a transition. In addition, the UPPG will be used to test for the effect of temporal correlations which are introduced by the TCPG. Because of the high resolution afforded by a sweep across probabilities, it is possible to compute critical exponents from the UPPG. For the TCPG and UPPG, the data for each window size and probability is generated from 5000 and 500 Monte Carlo simulations, respectively. This chapter will present the results of the volume hysteresis simulations, the hydrogen bond data, and the rigidity calculations. From this, new insights into the molecular clustering of CMHP and CDHP systems will be gleaned, and evidence for rigidity percolation as an additional mechanism in the glass transition will be presented.



Figure 4.1: The number of trehalose and water molecules as a function of concentration. Solid (dashed) lines refer to CDHP (CMHP) while black (red) curves refer to water (trehalose). The water and trehalose curves intersect at ~ 96 percent solute by weight.

4.1 Volume Hysteresis

As described in the methods chapter, the glass transition concentration can be identified by initializing several different systems at different volumes and then running those systems in the NPT ensemble until each one converges. According to this volume hysteresis method, a particular concentration is said to be a glass if all separately prepared systems converge to different final volumes; if they converge to the same final volume, that concentration is predicted to be a liquid. This is quantified by averaging together the final volumes and calculating the standard deviations. Figure 4.2 shows the volume hysteresis results for CMHP and CDHP at different temperatures. The glass transition concentration occurs when the the standard deviation (error bars) increase suddenly which indicates variation in the final volumes. Moreover, the average volume tends to decrease until C_g , at which point it begins to non-monotonically increase thereby resulting in a volume minimum.



Figure 4.2: Volume hysteresis data for (a) CMHP and (b) CDHP at different temperatures.

Figure 4.2 plots the average final volume as a function of concentration. From this it can be seen that CMHP experiences a transition at 69.5 and 72.5 percent solute by weight for 270 K and 300 K, respectively. In CDHP the transitions occur at 71, 81.5, and 85 percent solute by weight for 270 K, 300 K, and 330 K, respectively. These values are determined to within 1.5 percent solute by weight.



Figure 4.3: Shown above are the H-Bond donor-acceptor angle distributions taken from a single frame. The number at the red dot is the angle of the first minimum of the distribution. Note the prominent peaks succeeding the first which correspond to the non-participatory water hydrogen.

4.2 Hydrogen Bond Analysis

4.2.1 Hydrogen Bond PDF distributions

H-Bond distributions for 72.5 percent solute by weight are given in figures 4.3, 4.4, and 4.5. The distributions for phosphate donors (not shown) indicated no preferential angle, and were monotonically increasing in a Gaussian-like distribution. Cholinecholine interactions are not shown because there were too few instances to generate a detailed distribution, and this is indeed expected of cation-cation interactions. All



Figure 4.4: See caption of figure 4.3. Notice the absence of prominent secondary and tertiary peaks as compared to when water is a donor.

other distributions displayed interesting features, and seemed to indicate a preference for linearity. Distributions involving water tend to show a prominent second peak which is due to the opposing hydrogen in water (not participating in the bond).

PDFs at other concentrations showed similar trends except at very low solute concentrations where there were too few instances of some H-bonds between solute molecules. Though these distributions seem to indicate that a cutoff angle between 25 and 45 degrees is appropriate, we can not rule out underlying distributions. The H-bond lifetime data shown below, however, will corroborate the results of the PDF distribution because τ_{max} occurs in the range $25 < \theta_c < 45$ degrees in almost all cases.



Figure 4.5: Choline-Choline interactions are excluded because the PDF is generated from too few data points to make any meaningful inferences.

4.2.2 Hydrogen Bond Correlation Functions

Cutoff angles and distances were found from the method described in section 2.4. In all cases, with a few exceptions, the H-bond lifetime plots monotonically increased to a maximum distance or angle cutoff and then decreased. The lifetime plots were bumpy at low concentrations, but this is a result of statistical noise because there were too few solute molecules. This noise can be smoothed out by calculating the correlation functions (shown in figure 4.6) starting at $t_f - n\tau_{max}$, where t_f is the time of the final frame in the simulation, τ_{max} is the maximum lifetime calculated for the plot being smoothed, and n is an integer. Correlation functions are then calculated for different values of n so that multiple lifetimes can be obtained for the same H-Bond pair at a given concentration and cutoff angle, and the smoothed lifetime plots are calculated from averaging these new lifetimes. The temporal difference of τ_{max} between initial states of equation 3.1 will ensure that each $h(t_f - n\tau_{max})$ is statistically independent for each additional C(t) (for smoothing) calculated. This procedure will tend to cause a decrease in τ because equation 3.2 is calculated over a smaller time interval, but this effect is small compared to the fluctuations in the lifetime plots, and in all cases, the plots were sufficiently smoothed to definitively identify a cutoff angle.



Figure 4.6: The log correlation functions for choline-phosphate and water-phosphate donor-acceptor H-Bond pairs at 72.5 percent solute by weight are shown as a function of time over the entire simulation for different angle cutoffs. The insets show the middle of the plots where the decay is exponential. One can see that there is a slight angle dependence which allows a geometric cutoff to be made. Also note the faster decay in (b) compared to (a) despite the fact that water neighboring phosphate is not considered a plasticizer in CMHP systems. These plots were calculated with no distance cutoff.

The only other exceptions to this trend are for instances of phosphate donors, choline-choline interactions and for lifetime plots with HA distance on the abscissa. Phosphate is rarely a donor [59] and therefore lacks a preferential angle, and phosphatephosphate as well as choline-choline interactions can be ruled out because we would not expect anion-anion or cation-cation clustering. These plots are only monotonically increasing which indicates that a less restrictive cutoff angle counts H-Bonds which are just as transient as the H-bonds for a smaller θ_c . This supposition is confirmed by the relatively short lifetimes (not shown). The last exception was observed in H-bond lifetime plots for variable distance with no angle restriction. These tended to be smoother because H-bonds were sampled within a sphere of expanding radius instead of a cone with expanding apex angle. For these plots, an uptick was observed around 2.9 angstroms (see panel (b) of figures 4.7 and 4.8) which is due to the opposing hydrogen in water. A reasonable HA distance cutoff would then be 2.9 angstroms in order to exclude the opposing hydrogen in water. As HDA angles and HA distances are strongly correlated, this analysis will focus only on the HDA angle.



Figure 4.7: The lifetimes calculated from equation 3.2 are shown for the CMHP system at 300 K and 72.5 percent solute by weight. (a) shows the lifetime for each H-Bonded donor-acceptor pair as a function of cutoff angle with no HA distance restriction. (b) shows the lifetimes with no angle restriction. Notice the uptick in the inset of (b) at around 2.9 angstroms which corresponds to the second hydrogen in water being incorrectly sampled as an H-Bond.

An important result from this data is that the maximum lifetime occurs at approximately the same cutoff angle at all concentrations and for all H-bond pairs. Using equation 3.3 for all concentrations, it was found that $\theta_c = 37$ is the optimal angle cutoff. Though it is common to use a single geometrical cutoff in H-bond analyses, such close agreement of θ_c was not expected across all concentrations and H-bond pairs. This greatly simplifies the analysis because now the same angle and distance $(r_c = 2.9 \text{ angstroms})$ cutoffs can be used in all systems.



Figure 4.8: See the caption of 4.7. Shown are the lifetime plots for the CMHP system at 300 K and 89 percent solute by weight.

It is interesting that at low concentrations, choline-water donor-acceptor pairs are less strongly bound than water-water H-bonds while choline-water lifetimes surpass water-water lifetimes at high concentrations. The RFM may account for this behaviour because choline (104.17 g/mol) is larger and heavier than water, so its diffusivity should be lower than bulk water at low concentrations. At high concentrations, there is less bulk water and more interfacial water which limits the mobility of water molecules so that they may remain neighboring to choline for longer. Because choline has only one hydroxyl group to form H-bonds, it is incapable of forming closed loops in the rigid network, and the effects of choline on network rigidity will not be manifested in the usual way whereby regions rich with H-bonds will locally form closed loops. Instead, choline may hinder rigidity percolation due to disruptions of the local H-bond network. The conclusions from previous work [25] that the local micro-environment affects the character of inter-molecular interactions and clustering are corroborated by these results.



Figure 4.9: H-bond lifetime data as a function of cutoff angle is shown for CMHP and CDHP at 270 K and 300 K each. The concentration (in % solute by weight) for each panel is: (a) 62 %; (b) 65 %; (c) 60 %; and (d) 65 %. These concentrations are all below the glass transition concentration.

Because of the difficulty in making a one-to-one comparison between CMHP and CDHP systems according to the concentration alone, comparisons will be made depending on if $C < C_g$, $C \approx C_g$, and $C > C_g$, which are shown in figures 4.9, 4.10, and 4.11, respectively. Each figure compares CMHP and CDHP at 270 K and 300 K. A cursory inspection of the data shows that H-bond lifetimes depend on temperature, and CDHP systems are more mobile than CMHP. Moreover, H-bonds involving water molecules tend to be short lived compared to interactions between solute only. These observations follow the same trends as most laws such as the Stokes-Einstein relation, the arrhenius equation, and the Vogel-Fulcher equation, among others, and the RFM may again be invoked to justify the inhomogeneity of H-bond lifetimes.

Previous results indicated a propensity for water and monohydrogen phosphate to cluster, but the data in figures 4.9, 4.10, and 4.11 do not seem significantly different from other water-solute interactions. In the context of the RFM, this is not surprising because most water molecules have very fleeting interactions with phosphate which drives down the lifetime. Nevertheless, there is still a fraction of water molecules which remain neighboring to phosphate for significantly longer than 10 ps. Going from low to high concentrations and high to low temperatures, there is an appreciable increase in the H-bond lifetimes. This trend holds true for all concentrations, temperatures, and systems, as expected. Trehalose-trehalose and trehalose-phosphate H-bonds are the most persistent interactions at all concentrations which corroborates previous claims that volume hysteresis arises from strain being released in the trehalose matrix. Notably, trehalose-phosphate H-bonds surpass the lifetimes of trehalose-trehalose Hbonds at higher concentrations. This suggests that the phosphate anion plays a key part in facilitating the liquid-glass transition beyond its interactions with water. However, it will be shown later that the cause of rigidity percolation cannot be so easily adduced from molecular clustering, as phosphate does not appear with the expected abundance in the percolating cluster.



Figure 4.10: H-bond lifetime data as a function of cutoff angle is shown for CMHP and CDHP at 270 K and 300 K each. The concentration (in % solute by weight) for each panel is: (a) 71 %; (b) 71 %; (c) 69.5 %; and (d) 74 %. These concentrations are all in the vicinity of the glass transition concentration.

At concentrations below C_g , the longest lived H-bonds occur between trehalose molecules and trehalose hydroxyl groups donating to phosphate anions. All H-bonds between solute and water molecules are short-lived due to the high diffusivity of water prior to structural arrest. Towards intermediate concentrations in the vicinity of C_g (figure 4.10), there is a broadening spectrum of H-bond lifetimes as degrees of freedom are frozen out. Indeed, these additional relaxations are illustrated in panels (a) and (b) of figure 4.10 which show CMHP at 270 K and 300 K, respectively, but both with a constant concentration of 71 %. One can surmise from this data the system at 270 K is closer to the transition than the same system at 300 K simply by comparing the variance in H-bond lifetimes. Not only does this suggest what the clustering propensities of the various species might be, but it guides our expectations for what ought to be present in the percolating cluster and what species may be primarily responsible for the onset of structural arrest. Another interesting feature of panels (a) and (b) of figure 4.10 is that all H-bond lifetimes scale with approximately the same factor in going from 270 K to 300 K, except for trehalose-phosphate H-bonds which drop significantly. This is likely because phosphate is a smaller molecule compared to trehalose which implies that its diffusivity is more sensitive to temperature variation, as one may conclude from the RFM.

Figure 4.11 shows H-bond lifetimes above C_g where the lifetimes for water-water H-bonds are 100 ps for CMHP and 100 ps for CDHP at 270 K and 50 ps for CDHP at 300 K. The H-bond lifetime plots for each donor-acceptor interaction have clearly become vertically separated in CMHP at high concentrations while the separation is present but not as pronounced in CDHP (panels (c) and (d)). Panels (b) and (d) of figure 4.11 show CMHP and CDHP at 89 % at 300 K, respectively. Though this is not strictly a one-to-one comparison, it is clear that CDHP systems have a higher diffusivity compared to CMHP systems at similar concentrations. It is evident that the H-bond lifetimes alone are informative and may be used to approximate the location of the liquid-glass transition, and to predict which species are most important in the transition. In particular, one may expect trehalose and phosphate to be primarily responsible for rigidity percolation based on the high lifetimes for trehalose-trehalose and trehalose-phosphate H-bonds. In contrast, H-bonds involving water usually have lower lifetimes than H-bonds between two solute molecules, so one may predict that water should not be present in the percolating cluster.



Figure 4.11: H-bond lifetime data as a function of cutoff angle is shown for CMHP and CDHP at 270 K and 300 K each. The concentration (in % solute by weight) for each panel is: (a) 80 %; (b) 89 %; (c) 84.5 %; and (d) 89 %. These concentrations are all above the glass transition concentration.

Table 4.1: Results of the H-bond analysis for CMHP at 270 K. The left three columns and right three columns are calculated from simulations which output data every 6 ps and 0.1 ps, respectively. θ_c and τ are the cutoff angle and H-bond lifetime defined in the text. The number of H-bonds, N_h , refers to $\langle h(0) \rangle$ defined in the H-bond correlation function (equation 3.1).

CMHP at 270 K						
	$\Delta t=6.0$	ps		$\Delta t=0.1$	ps	
Concentration	$ heta_c$ (°)	$\langle \tau \rangle ~(\mathrm{ps})$	N_h	$ heta_c$ (°)	$\langle \tau \rangle$ (ps)	N_h
(% solute by						
weight)						
10	39.76	1.08	36968	37.52	0.63	36133
20	37.48	2.02	33717	37.42	1.22	34020
62	36.35	293.61	24656	37.18	35.5	24791
63.5	35.12	343.04	24420	36.25	40.19	24428
65	35.09	463.49	23852	36.10	46.08	23956
66.5	35.44	567.65	23733	35.60	50.28	23738
68	35.43	628.01	23364	36.69	56.37	23590
71	36.88	748.57	23359	35.86	66.75	23010
72.5	35.83	938.45	23003	34.75	71.11	22689
74	34.58	1189.05	22479	35.71	75.53	22594
80	36.41	1976.12	21565	36.18	99.98	21541

Table 4.2: Results of the H-bond analysis for CMHP at 300 K. The left three columns and right three columns are calculated from simulations which output data every 6 ps and 0.1 ps, respectively. θ_c and τ are the cutoff angle and H-bond lifetime defined in the text. The number of H-bonds, N_h , refers to $\langle h(0) \rangle$ defined in the H-bond correlation function (equation 3.1). A – indicates that data was not collected.

CMHP at 300 K						
	Δt =6.0	ps		$\Delta t=0.1$	ps	
Concentration	$ heta_c$ (°)	$\langle \tau \rangle$ (ps)	N_h	θ_c (°)	$\langle \tau \rangle$ (ps)	N_h
(% solute by						
weight)						
40	38.06	2.90	28771	_	_	_
65	36.77	113.59	23876	_	_	_
68	36.27	184.38	23392	_	_	_
71	36.15	298.22	23105	35.76	39.45	22796
72.5	35.51	418.58	22763	35.51	6.98	22763
74	35.81	462.92	24436	36.73	51.07	22200
77	36.98	973.95	21903	_	_	_
80	36.55	1165.43	22194	36.28	78.41	21554
86	36.88	1826.06	19756	_	_	_
89	37.02	2709.45	18374	_	_	_

Table 4.3: Results of the H-bond analysis for CDHP at 270 K. The left three columns and right three columns are calculated from simulations which output data every 6 ps and 0.1 ps, respectively. θ_c and τ are the cutoff angle and H-bond lifetime defined in the text. The number of H-bonds, N_h , refers to $\langle h(0) \rangle$ defined in the H-bond correlation function (equation 3.1). A - indicates that data was not collected.

CDHP at 270 K						
	$\Delta t=6.0$	ps		$\Delta t=0.1$	ps	
Concentration	$ heta_c$ (°)	$\langle \tau \rangle$ (ps)	N_h	θ_c (°)	$\langle \tau \rangle$ (ps)	N_h
(% solute by						
weight)						
40	37.50	11.08	31890	37.46	4.95	31626
60	36.35	116.49	28115	37.69	21.01	28518
65	37.04	203.62	27761	37.82	29.59	27768
69.5	36.47	420.47	27129	36.46	43.36	26940
71	37.10	513.03	26879	37.08	48.54	26710
72.5	37.88	566.37	26538	38.47	27.83	26521
74	36.57	739.70	26054	37.73	60.75	26332
81.5	38.87	1349.57	25115	38.11	84.21	24814
83	37.69	1486.12	24457	38.54	92.68	24629
84.5	38.18	1815.38	24110	37.59	100.30	23704

Table 4.4: Results of the H-bond analysis for CDHP at 300 K. The left three columns and right three columns are calculated from simulations which output data every 6 ps and 0.1 ps, respectively. θ_c and τ are the cutoff angle and H-bond lifetime defined in the text. The number of H-bonds, N_h , refers to $\langle h(0) \rangle$ defined in the H-bond correlation function (equation 3.1). A – indicates that data was not collected.

CDHP at 300 K						
	Δt =6.0	ps		$\Delta t=0.1$	ps	
Concentration	$ heta_c$ (°)	$\langle \tau \rangle$ (ps)	N_h	θ_c (°)	$\langle \tau \rangle$ (ps)	N_h
(% solute by)						
weight)						
65	38.32	48.73	27505	_	_	_
68	36.36	74.29	26447	38.33	17.54	26889
72.5	37.64	148.49	25952	38.95	26.27	26669
77	37.49	360.67	25443	39.28	40.76	25663
81.5	39.68	707.39	24880	39.03	58.70	24623
83	37.59	863.17	24013	38.47	63.46	24246
84.5	39.05	1028.94	23935	39.62	72.61	23902
89	40.51	1517.72	22476	39.14	88.11	21997

Tables (4.1), (4.2) (4.3), and (4.4) show the weighted average HDA angle cutoff defined in equation (3.3), the weighted average H-bond lifetime in equation (3.4), and the number of H-bonds recorded in the last frame of the simulation. Across all systems, the number of H-bonds decreases for increasing concentrations while the H-bond lifetimes become longer at high concentrations. This indicates that most H-bonds are between highly mobile water molecules at low concentrations. These rapidly flickering H-bonds have a low probability of existence across any window of time so they are likely not included in the generic graph topology. The individual lifetimes which enumerate each type of interaction (table 3.1) show that H-bonds involving water tend to become more persistent at high concentrations and thus show a dependence on the local micro-environment. The lengthening average lifetime is expected for glasses which have a continuous spectrum of relaxation times. As water content is removed, there is a broadened spectrum of relaxations which is associated with an increase in dynamical heterogeneity and a concomitant reduction in entropy which are hallmarks of the glassy state [60].

An important result to notice in the H-bond analysis is that the optimal cutoff angle is remarkably consistent across four dimensions of variability: temperature, chemical composition (CMHP or CDHP), concentration, and time scale of observation (Δt) . Though it is known that a uniform geometrical cutoff tends to be good at identifying H-bonds, it is not *a priori* evident that a single geometrical cutoff can be used in complex, multi-component solutions. The H-bond analysis herein suggests a universality with respect to θ_c . A cutoff angle of $\theta_c = 37$ degrees is thus used to determine the existence of all H-bonds.

4.3 Rigidity Analysis

This section will show the results of the PG algorithm for CMHP and CDHP systems at 270 K and 300 K each. First, the results of the time-correlated PG with a time dependence incorporated via a sliding window will be presented. The rigid cluster distributions will then be examined with an emphasis on the dynamics of the largest or percolating rigid cluster. Following this, data for the uniform-probability PG will be presented, and the beta critical exponents will be extracted. For both the TCPG and UPPG, the data will be presented as a function of window size or probability, respectively, and then data will be presented as a function of the MCN.

4.3.1 Time-Correlated Pebble Game

Data for the number of vertices, floppy modes, RCS, MCN, and largest cluster size will be shown as a function of window size. The same data will again be shown, but as a function of the MCN, with the exception of the number of vertices and MCN.



Figure 4.12: Data for the number of vertices (atoms) in the network after pruning is shown as a function of window size for various concentrations.

Figure 4.12 shows the number of vertices left in the network after pruning as a function of window size. It is important to note the effect that the window size has on the network, because all dangling ends are removed from the network by pruning. The resulting network consists only of closed loops which may be connected by chains of atoms. A larger window size will result in low-probability H-bonds and in turn,

more broken connections which lead to more dangling ends. For this reason, there are fewer atoms in the network as window size increases. Pruning the network is crucial for this rigidity analysis, as dangling ends or chains contribute degrees of freedom which result in the inaccurate determination of the rigidity transition.

The insets in panels (a) - (d) of figure 4.12 show the same data on a logarithmic scale revealing a power law which is due to the decay of H-bonds. Thus, the properties of the pruned network are directly related to the underlying H-bond dynamics. Comparing panels (a) and (b) against (c) and (d), one observes that there are more atoms which contribute to rigidity percolation in CDHP systems than CMHP systems. From this, one may hypothesize that CDHP has more opportunities to form bonds which would increase the MCN. However, the glass transition actually occurs at a higher concentration in CDHP than in CMHP.

Figure 4.13 shows the effect on the floppy modes as window size increases. A larger window size results in more broken bonds, which reduces the number of redundant bonds. Therefore, the number of floppy modes increases as the time scale of observation increases. Again, the insets show a power law behavior which can possibly be attributed to the the decay of H-bonds though it is not obvious that this should be the case.

Figure 4.14 shows the RCS as a function of window size. Percolation theory predicts that there should be a peak in the RCS at the rigidity transition. An important result here is that there are some concentrations which never experience a rigidity transition. For CMHP the lowest concentration at which there is a rigidity transition is between 68 and 71 percent solute by weight for 270 K and 71 percent for 300 K. For CDHP, the lowest concentrations which experience a transition are 69.5 and 74 percent solute by weight for 270 K and 300 K, respectively. Below these concentrations, no rigidity transition will occur for a sampling rate of 6 ps. By using 0.1 ps, the rigidity transition can be shifted to slightly lower concentrations, but the effect is modest. While these



Figure 4.13: Data for the number of floppy modes normalized per degree of freedom is shown as a function of window size for various concentrations. The number of degrees of freedom is defined as three times the number of vertices.

concentrations do not align exactly with the value of C_g as predicted by volume hysteresis, they are close for CMHP and for CDHP at 270 K. In fact, they align quite well with the volume minimums observed in figure 4.2. From this, a compelling narrative emerges: due to the low isothermal compressibility of water, removing water content enables the system to contract during equilibration in the NPT ensemble. There is a threshold, however, where the entropy has been reduced so much that the



Figure 4.14: Data for the RCS is shown as a function of window size for various concentrations.

system can no longer be regarded as a liquid but has now become a glass. It is at this threshold that there is a volume minimum, and as more solute is added (and water removed), the volume must begin to expand to accommodate the increased strain in the system. It is also at this volume minimum that rigidity percolation is first detected in the glass network. That the volume minimum should coincide with the rigidity percolation threshold is evidence that the latter has a crucial role in the liquid-glass transition in these systems. Though there is not perfect agreement, the volume minimums and percolation thresholds agree to within 5 percent solute by weight in all cases. There are many possible explanations for this discrepancy. When the percolating cluster first appears in a second order transition, it is typically highly ramified so that there are still many floppy inclusions in the network. Therefore, the emergence of a percolating cluster does not necessitate a global reduction in mobility, especially if the percolating cluster is transient and constantly being reformed.



Figure 4.15: Data for the mean coordination number is shown as a function of window size for various concentrations.

Figure 4.15 plots the MCN as a function of window size. From figure 4.12 which plots the number of remaining vertices as a function of window size, the MCN should decrease for increasing window sizes. The insets again show a power law decay, which is consistent with all prior data. Though there is a definite dependence on concentration, the vertical shift between concentrations is in the hundredth decimal place. Because $\langle r \rangle$ is accurate to the thousandth decimal place, this shift is significant.



Figure 4.16: Data for the fraction of sites in the largest cluster is shown as a function of window size for various concentrations.
The fraction of sites in the largest cluster is the order parameter of interest in a rigidity transition, and this is plotted in figure 4.16. The trends here are consistent with the trends seen in the number of vertices, floppy modes, and MCN. A larger window size reduces the size of the largest cluster making it more difficult for rigidity to percolate. It is conceivable that a system with strong enough correlations can percolate rigidity at low MCNs, however, CMHP and CDHP simulated in the stoichiometric ratios and temperatures in this work do not appear to have particularly strong correlations. This point is clarified by comparing the TCPG with the UPPG. Panel (b) of 4.16 shows a crossover between 86 % and 89 %, and panel (c) shows a crossover between 81.5 % and 83 %. This effect is possibly due to frustration of the network. Water molecules which occupy equally favorable states both in and out of the percolating cluster will result in a greater variance over all Monte Carlo trials as seen in the figure. Indeed, it will be shown that water is not only present in the percolating cluster but exists in abundant quantities. The presence of frustration is a likely consequence for molecules which are typically liquid-like but forced into configurations which severely limit local mobility.

Figure 4.17 plots the number of floppy modes as a function of the MCN which is the parameter of interest in TCT. Comparing this to figure 4.13, we see a spectacular collapse of the data onto a single curve. Each value of the MCN corresponds to a different window size so that as the window size decreases, the MCN increases. The slope of the curve in the floppy region is the same for all concentrations and is approximately the same between CMHP and CDHP irrespective of the temperature. This shows there is a characteristic universality in the floppy modes with respect to the mean coordination number in CMHP and CDHP systems as observed in other systems [15]. By observing the insets, one can see that that the curves of the individual concentrations slightly depart from one another. This is especially pronounced in panel (b). One possible explanation is that in the floppy region where r 2.4, there are



Figure 4.17: Data for the number of floppy modes normalized per degree of freedom is shown as a function of mean coordination number for various concentrations. Each point along each curve corresponds to a different window size.

no correlations in time because they have been averaged out over the longer period of observation. If floppy modes are removed at the same rate for all systems in the liquid phase, then the curves should appear the same for all liquid systems, and they should intersect the abscissa at approximately 2.4 as predicted by Maxwell constraint counting. After the transition, however, time correlations are more important which will cause some deviation from the mean field approximation.



Figure 4.18: Data for the RCS is shown as a function of mean coordination number for various concentrations. Each point along each curve corresponds to a different window size.

Figure 4.18 shows the RCS as a function of the MCN. Again, the data collapses onto a single curve as compared to the same data plotted as a function of window size in figure 4.14. While Maxwell's mean field approximation would predict the peaks to be centered at $\langle r \rangle = 2.4$, they are all centered at approximately 2.38. Because of the precision in calculating $\langle r \rangle$, this is not an insignificant shift. Evidently, spatialtemporal correlations are present which shift the rigidity transition to lower MCNs.



Figure 4.19: Data for the fraction of sites in the largest cluster is shown as a function of mean coordination number for various concentrations. Each point along each curve corresponds to a different window size.

The size of the percolating cluster when plotted as a function of the window size does not show a clear transition, but figure 4.19 which plots the fraction of sites in the largest cluster as a function of MCN gives a clear phase transition from floppy to rigid. Panels (b) and (c) show anomalous behavior at 89 and 83 percent solute by weight, respectively. As already mentioned, this is believed to indicate frustration in the network related to the behavior of water molecules.



Figure 4.20: Data for the mean coordination number is shown as a function of concentration for various window sizes.

Figure 4.20 shows the dependence of the mean coordination number on concentration. It is interesting that the dependence is linear, as the value of $\langle r \rangle$ ought to depend on changes in enthalpy-entropy competition as concentration increases for a given temperature. At most concentrations (up until ~ 96 % solute by weight) there is an excess of water, so the addition of more solute molecules will result in bonds with a greater potential well depth and increase the enthalpy of the system. Conversely, the presence of larger solute molecules will reduce the overall mobility of the system and reduce the entropy. The concentration is increased by removing water molecules while adding solute; the number of water molecules decreases non-linearly with concentration while the number of solute molecules increases approximately linearly. Because all of the concentrations simulated have an excess of water molecules, this trend would suggest that it is solute which controls the MCN of the system. It has already been made clear that the MCN is the "good" parameter to use in order to clearly observe the transition from floppy to rigid. Therefore, it would be interesting to observe the effect that substituting different disaccharides would have on the MCN for different concentrations.

The result of figure 4.20 is important because the concentration also becomes a good parameter for analyzing the rigidity properties via a linear mapping. Figure 4.21 plots the floppy modes normalized per degree of freedom as a function of concentration. It is clear from this that there is a region where degrees of freedom become frozen out, and where this occurs varies slightly between temperatures and even more so between CMHP and CDHP.

4.3.2 Cluster Distributions

In this section, the distribution of rigid clusters and the constituents of the largest or percolating cluster will be examined. This information is quite revealing because it shows the capability of a system to support a percolating cluster, and it elucidates which components are primarily responsible for rigidity percolation.



Figure 4.21: Data for the number of floppy modes normalized per degree of freedom is shown as a function of concentration for various window sizes.



Figure 4.22: Data for the fraction of sites in the largest cluster is shown as a function of concentration for various window sizes.



Figure 4.23: Data for the number of sites in the largest rigid cluster is shown as a function of window size on a semi-log scale for various concentrations of CMHP at 270 K. The histograms are broken down by the type of molecule participating in the largest cluster, but only atoms within molecules are counted not whole molecules.



Figure 4.24: Data for the number of sites in the largest rigid cluster is shown as a function of window size on a semi-log scale for various concentrations of CMHP at 300 K. The histograms are broken down by the type of molecule participating in the largest cluster, but only atoms within molecules are counted not whole molecules.



Figure 4.25: Data for the number of sites in the largest rigid cluster is shown as a function of window size on a semi-log scale for various concentrations of CDHP at 270 K. The histograms are broken down by the type of molecule participating in the largest cluster, but only atoms within molecules are counted not whole molecules.



Figure 4.26: Data for the number of sites in the largest rigid cluster is shown as a function of window size on a semi-log scale for various concentrations of CDHP at 300 K. The histograms are broken down by the type of molecule participating in the largest cluster, but only atoms within molecules are counted not whole molecules.



Figure 4.27: Data for the distribution of rigid clusters for various concentrations and window sizes of CMHP at 270 K.



Figure 4.28: Data for the distribution of rigid clusters for various concentrations and window sizes of CMHP at 300 K.



Figure 4.29: Data for the distribution of rigid clusters for various concentrations and window sizes of CDHP at 270 K.



Figure 4.30: Data for the distribution of rigid clusters for various concentrations and window sizes of CDHP at 270 K.

Figures 4.23, 4.24, 4.25, and 4.26 are histograms for the average number of atoms belonging to each species in the largest cluster for various window sizes. In all systems, trehalose occurs in the greatest quantity. Surprisingly, water is also abundant in the rigid cluster, and occurs in greater quantities than phosphate or choline. This is seemingly in contradistinction to the H-bond lifetime plots which show that trehalose and phosphate tend to cluster, while most H-bonds involving water are transient. As figure 4.1 shows, there is an excess of water molecules at all simulated concentrations. This implies that water has a space filling tendency which makes it unlikely for any trehalose hydroxyl group to not be participating in a hydrogen bond. The lifetime plots for trehalose-water (donor-acceptor) H-bonds also show that these bonds can persist for ~ 2000 ps at high concentrations. The conclusion from this argument is that water becomes trapped in trehalose cages towards higher concentrations, but it is still not obvious how water appears so prevalently in the percolating rigid cluster. The ability for water to percolate rigidity depends on the geometry of water and the solute to which it is H-bonded, but the data here is insufficient to deduce the mechanism. At the moment, however, we can assume that the mechanism to rigidify water exists and use this fact to glean subtle insights into what drives the liquidglass transition. The trends in the makeup of the largest cluster are very similar between CMHP and CDHP for both 270 and 300 K. At higher concentrations, the largest cluster persists for longer periods of time. If the observation time scale were increased to infinity but the size of the largest cluster remained constant, the system would in fact be a crystal and not a glass. This observation is crucial to understanding the interpretation of the time window defined in this work. Empirically, it is known that a glass will crystallize if the observation time scale is made to be adequately long; the failure of crystallization (as indicated by a decrease in the size of the largest cluster) reveals that the system is a glass because it is not able to crystallize over the given window of observation. Stated a different way, the number of sites in the largest cluster can be taken as an order parameter which differentiates between an ordered and disordered phase. This order parameter is controlled by the observation window, and increasing the time scale of observation will cause a transition to the disordered state if the system is a glass. However, no such transition will occur for a crystal owing to its long range order.

It is already known from prior work that the additional charge on the monohydrogen phosphate anion drastically alters the clustering of molecules when comparing CMHP and CDHP. In CMHP, a third category of water is introduced for water molecules which neighbor phosphate because of the reduced diffusivity of those water molecules. It is this reduction in diffusivity which causes a global loss of mobility in CMHP systems which shifts the liquid-glass transition to lower solute concentrations as compared to CDHP. Upon inspection of the largest cluster histograms, there are few differences between CMHP and CDHP, except for the prevalence of the phosphate anion in the percolating cluster. In CMHP, monohydrogen phosphate appears in the rigid cluster in similar quantities to CDHP until higher concentrations. In panel (e) of figure 4.24 (CMHP at 300 K), the number of choline atoms present in the largest cluster exceeds the number monohydrogen phosphate atoms. This occurs at relatively high solute concentrations (86 %) in CMHP but does not occur in CDHP. This indicates that the clustering characteristics in CMHP are chiefly due to the interaction between water and phosphate. Furthermore, H-bonds between water and phosphate have similar persistence times in both CMHP and CDHP at low and intermediate concentrations (figures 4.9 and 4.10). At high concentrations (figure 4.11), there is a marked difference, and CMHP water-phosphate H-bonds are 6 to 8 times more persistent than the same H-bonds in CDHP. This suggests that at high concentrations, the remaining water is clustering around phosphate in such a way that it is unable to bond to trehalose. This is possibly due to steric hindrance at high solute concentrations.

Figures 4.27, 4.28, 4.29, and 4.30 show the distribution (on a semi-log scale) in size of all rigid clusters detected over the length of the simulations. These figures indicate that the evolution characteristics for CMHP and CDHP are qualitatively the same, and furthermore there is weak (if any) temperature dependence.

4.3.3 Uniform-Probability Pebble Game

The UPPG assigns each H-bond with the same probability as a mean field approximation so that the average probability required for the rigidity transition at any concentration can be approximately determined. This neglects temporal correlations which are captured by the TCPG, but it is still useful to know what probabilities are required in order to affect a transition or if a rigidity transition is possible at all. In addition, comparisons between the TCPG and UPPG will reveal the effect that temporal correlations have on the rigidity transition. It is known that correlations can significantly alter the MCN at the transition point and that these correlations may be traced back to the ring distributions [7]. Therefore, a comparative analysis between the TCPG and UPPG can provide insight into the underlying mechanism of the rigidity transition. In this section, data as a function of uniform H-bond probability will be examined, followed by the same data as a function of the MCN.

Figure 4.31 plots the MCN as a function of the uniform H-bond probability. Note these curves are approximately linear which suggests that the H-bond probability may also be a "good" parameter to use (in addition to the MCN). Higher solute concentrations also tend to have slightly higher MCNs for a given probability. This is in spite of the fact that tables 4.1, 4.2, 4.3, and 4.4 show that there are fewer available H-bonds at high concentrations. Because all H-bonds are identified before pruning, this suggests that significantly more atoms are pruned at lower concentrations while more H-bonds are preserved at higher concentrations.

Figure 4.32 plots the number of floppy modes normalized per degree of freedom against the H-bond probability. Here, the floppy modes exhibit the characteristic



Figure 4.31: Data for the mean coordination number is shown as a function of H-bond probability.

curve where there are two slopes connected in a piece-wise fashion. This is in contrast to the floppy modes plotted as a function of window size which displayed a power law increase. The reason for this is that there is a linear mapping from probability to MCN. Note that low concentrations tend to deviate from this characteristic shape. This is likely because the PG is not strictly being run in the canonical ensemble (i.e., the number of particles is not preserved) due to the pruning algorithm. At



Figure 4.32: Data for the number of Floppy modes normalized per degree of freedom is shown as a function of H-bond probability.

very low concentrations, most atoms are removed from the network, so the process of normalizing by degree of freedom may give unusual results.

The RCS plotted as a function of H-bond probability is shown in figure 4.33. These are quite informative, as they indicate which concentrations are not capable of affecting a rigidity transition even when p = 1 (at 0 K) when all bonds are intact. For instance, the concentrations at 10 and 20 percent solute by weight in panel (a) never



Figure 4.33: Data for the RCS is shown as a function of H-bond probability.

show a transition. In panels (b) and (c) 40 percent solute by weight is able to undergo a rigidity transition at very high H-bond probabilities. Also note that the 40 % peak is much higher compared to other concentrations which could indicate a transition which is more first order in character. Also note that higher concentrations transition at lower probabilities which indicates again that there are more constraints left in high concentrations after pruning. Conversely, low concentrations require a higher average H-bond probability in order to experience a transition despite the fact low concentrations have more H-bonds due to the H-bond network of bulk water. This suggests that bulk water possesses a tree-like structure, and in turn, more constraints are pruned from the network at low concentrations. While it may be obvious that high concentrations should have more constraints, this is not necessarily the case if the solute is a network modifier which has the effect of breaking rigid constraints and introducing floppy modes in their place. Because there are fewer H-bonds at higher concentrations, the H-bonds which do exist must act as hard constraints in order to percolate rigidity.

Figure 4.34 plots the fraction of sites in the largest cluster as a function of the H-bond probability. Because of the high resolution afforded by a sweep over probabilities, it is possible to more clearly determine the order of the transition. A more detailed calculation will be presented later in which the β critical exponent will be calculated for CMHP systems at 300 K. Interestingly, the curve corresponding to 63.5 percent weight solute in panel (a) shows a similar anomaly at around p = 0.9 to what was observed in the TCPG in which the network appears to be frustrated.

By plotting the data as a function of MCN, we again see a spectacular collapse of the data onto a single curve in figure 4.35. As in the TCPG, the number of floppy modes begins to approach zero at around $\langle r \rangle = 2.4$. Again, the insets show a slight departure from this data collapse which is likely due to the fact that the system is not as well approximated by Maxwell's mean field approximation at high MCNs as compared to low MCNs.

Figures 4.36 and 4.37 respectively plot the RCS and fraction of sites in the largest cluster as a function of MCN. The insets show that there is a slight dependence on solute concentration, as lower (higher) concentrations experience a rigidity transition at lower (higher) MCNs. This implies the existence of spatial correlations irrespective of the observation time scale. Indeed, if a significant number of bonds are removed at low concentrations, the remaining bonds must transmit stress along structures which



Figure 4.34: Data for the fraction of sites in the largest rigid cluster is shown as a function of H-bond probability.

naturally arise due to spatial correlations.

Figure 4.38 compares the TCPG and UPPG for various concentrations. The motivation for this comparison is to test the hypothesis that temporal correlations in the network will shift the rigidity transition to lower mean coordination numbers. Temporal correlations originate from inhomogeneous clustering between different H-bond species so that some bonds are preferred over others. Various models and algorithms



Figure 4.35: Data for the number of floppy modes normalized per degree of freedom is shown as a function of mean coordination number. Each data point along each curve corresponds to a different uniform H-bond probability.

to introduce specific correlations already exist and have been shown to have an appreciable effect on the location of the rigidity transition. In particular, such algorithms may impose the requirement that all redundant bonds must be avoided when building the network or the probability of a bond existing at a site depends on the coordination number of neighboring sites. The effect of such requirements is that a transition may occur at lower MCNs (or volume fractions, probabilities, etc). Instead of constructing



Figure 4.36: Data for the RCS is shown as a function of mean coordination number. Each data point along each curve corresponds to a different uniform H-bond probability.

an algorithm or model to input these correlations, it is hypothesized that the MD simulation itself will provide correlations as it progresses towards a local energy minimum. Thus, the enthalpy-entropy competition intrinsic to soft matter systems will shape a generic graph topology which acts as a simplified, course-grain representation of the actual system.

It can be seen from figure 4.38 that there is a definite shift in the MCN between



Figure 4.37: Data for the fraction of sites in the largest cluster is shown as a function of mean coordination number. Each data point along each curve corresponds to a different uniform H-bond probability.

the TCPG and UPPG. The only difference between the TCPG and UPPG is that the former has heterogeneous H-bond probabilities which are calculated from MD simulation while the latter assigns H-bond probabilities uniformly. Therefore, it can be concluded that the presence of time correlations enables rigidity to percolate with fewer constraints. While it has been known that correlations can have this effect, it has never been confirmed in three dimensional bond bending networks consisting



Figure 4.38: Comparison between the time-correlated PG and the uniform probability PG. The fraction of sites in the largest cluster are shown for various concentrations for CMHP and CDHP at 270 K and 300 K each. There is a distinct shift in the rigidity percolation threshold with time correlations causing rigidity to percolate at lower MCNs.

of both covalent and hydrogen bonds which are constructed from MD simulation data. The shift tends to become slightly greater for increasing concentration which is clearly seen in panel (c) where 72.5 and 81.5 percent weight solute shows almost no shift between TCPG and UPPG data.

One of the difficulties in understanding the liquid-glass transition is that the prop-



Figure 4.39: Hydrogen bond probabilities calculated from simulation data. The difference between the uniform H-bond probability at the percolation threshold (as determined by the RCS peak) of the UPPG and the average H-bond probability from simulation data is plotted. When $p_{hb} > p_t$ ($p_{hb} < p_t$), the average H-bond probability is higher (lower) than the percolation threshold and the system is predicted to be rigid (floppy). The large (small) symbols refer to simulation data sampled every 6 ps (0.1 ps).

erties of a glass depend on the time scale of observation. This was incorporated into TCT as an adjustable parameter, but it is still difficult to definitively identify the rigidity transition because the time scale we choose is ultimately arbitrary. Experimentally, a similar problem exists because the chosen quench rate is also arbitrary, though there is a minimum quench rate which, if lengthened, would result in crystallization. In order to better address this problem, the average H-bond probability, p_{hb} , was calculated from MD simulations for various window sizes and subtracted from the rigidity percolation threshold probability (identified as the probability at the RCS peak), p_t , determined from the UPPG as shown in figure 4.39. When $p_t - p_{hb}$ becomes negative, the average H-bond probability is greater than the probability needed to affect a rigidity transition for that concentration. Conversely, when $p_t - p_{hb}$ becomes positive, the average H-bond probability is not high enough to cause a rigidity transition. It can be seen that there are some concentrations for which $p_t - p_{hb}$ is always positive which suggests that a transition cannot occur for any time scale. However, the MD simulation data is output with a sampling rate of 6 ps, so a smaller sampling rate may force the difference to become negative. Figure 4.39 plots both data for sampling rates of both 6 ps and 0.1 ps which are represented by large and small symbols, respectively. While the higher sampling rate does cause the systems to appear more vitreous, the effect is generally modest. Therefore, the determination of the rigidity transition is not arbitrary and very closely coincides with the volume minimums observed in the volume hysteresis data.

4.3.4 Critical Exponents

The β critical exponent has been calculated for CMHP at 300 K. Table 4.5 shows the threshold critical probability p_c , critical exponent β , and Pearson correlation coefficient for each concentration. The exponents are calculated from the UPPG because of the fine resolution in the sweep over probabilities. The values of β all tend to be around 0.5 which is in agreement with the value predicted by Landau theory. This suggests that rigidity percolation on three dimensional bond bending networks belongs to the same universality class as thermal phase transitions such as the paramagnetic to ferromagnetic transition.



Figure 4.40: Data for the fraction of sites in the percolating rigid cluster is shown as a function of $p - p_c$ on a logarithmic scale, where p is the uniform H-bond probability, and p_c is the approximate cutoff for the existence of the percolating cluster as determined from the peaks of the RCS in figure (4.33). Each data point along each curve corresponds to a different uniform H-bond probability.

Figure 4.40 plots the percolating cluster strength as a function of $p - p_c$ on a logarithmic scale. One can see a power law scaling in the vicinity of the of the transition point. Thus, we can conclude that the liquid-glass transition is driven by an underlying second order rigidity transition which belong to the same universality class as thermal phase transitions.

Table 4.5: Critical exponents for CMHP at 300 K. The critical exponent β is calculated from a linear regression of $P \sim (p - p_c)^{\beta}$ plotted on a logarithmic scale. r is the Pearson correlation coefficient for the fit. Only the first 20 data points after p_c were used in the fit.

Concentration	p_c	β	r
(% solute by weight)			
65	0.830 ± 0.005	0.434 ± 0.026	0.968
68	0.815 ± 0.005	0.431 ± 0.022	0.961
71	0.780 ± 0.005	0.507 ± 0.038	0.979
72.5	0.775 ± 0.005	0.502 ± 0.032	0.964
74	0.770 ± 0.005	0.538 ± 0.041	0.963
77	0.750 ± 0.005	0.501 ± 0.034	0.975
80	0.728 ± 0.005	0.596 ± 0.034	0.983
86	0.702 ± 0.005	0.484 ± 0.035	0.964
89	0.704 ± 0.005	0.547 ± 0.033	0.956

CHAPTER 5: Conclusion

Covalent bonded topological constraint theory has been used as a simple model for modelling glass networks for the past four decades. The notion of rigidity percolation is essential for a simple, microscopic analysis of glass networks, but its use has been confined to two dimensional central force networks or three dimensional covalent bond bending networks. Nevertheless, such networks have been the subject of vast amounts of research, and there now exists a large body of both experimental and theoretical evidence for the utility of rigidity theory. To fully appreciate any theory, however, requires knowledge of when it fails. Until the present, the question as to whether or not TCT can be applied to glasses with hydrogen bond networks has remained open. In this thesis, it has been demonstrated that TCT can successfully model glasses with H-bonds by incorporating a single adjustable parameter. This parameter is the time scale of observation which has a natural interpretation in the context of glass networks because the glass transition itself can be thought of as an experimental artifact which arises from taking measurements on time scales which are shorter than the time required to reach equilibrium. That there is a time scale dependence of the liquid-glass transition is an inescapable fact, but this is true of many other phenomena we measure. If one were to observe the evolution of earth's crust from Pangaea to present day, it would appear to be liquid. Likewise, liquid water appears to be solid on short enough time scales.

Fundamentally, the problem solved in this thesis is mapping MD trajectory data to a generic graph representation on which the pebble game algorithm can be employed for a rigidity analysis. This mapping is nontrivial and requires an analysis of the H-bond network. To this end, the H-bond network was characterized using H-bond correlation functions across CMHP and CDHP systems at 270 K and 300 K each. In order to test for any dependence on the time scale of observation, the correlation functions were calculated from simulation data sampled every 6 ps and every 0.1 ps. The optimal cutoff angle can be calculated from the correlation function using the method prescribed in this thesis, and there is no significant difference between the geometry of the H-bond network at these two time scales. Moreover, it was found that a single cutoff angle can be used in all systems and between all H-bonds, and the cutoff angle of 37 degrees generally agrees with values reported in the literature; however, unlike most analyses appearing in the literature, a rigorous rationale was given for our cutoff angle. In the process of determining this mapping, a possibly universal trait of H-bond networks has been uncovered in that a cutoff angle of approximately 37 degrees serves as a good definition for any H-bond.

From the H-bond analysis, it was found that there are fewer H-bonds at higher solute concentrations, yet there are more overall constraints present in the network at higher concentrations. This is due to the pruning algorithm which removes all tree-like structures leaving only rings of atoms and chains which connect those rings. The observed dynamics can be attributed to enthalpy-entropy competition. As an entropy crisis, the transition to the glassy state removes degrees of freedom so that the H-bonds which are present are unlikely to be pruned. At the same time, the increased strain also increases the enthalpy contribution. The net result is structural arrest which was quantified in this work via a rigidity analysis.

The H-bond analysis is crucial in constructing a generic graph representation on which the PG algorithm can be run. After all H-bonds have been identified, MD trajectory data is mapped to this representation and the PG algorithm is run. The topology can be constructed in one of two ways: (i) the H-bond probabilities calculated over a window in time for every H-bond in the system are input into a 5000 trial Monte Carlo simulation and (ii) a uniform H-bond probability is assigned to all identified H-bonds for 500 Monte Carlo trials. The former has time correlations because the MD simulation is heterogeneously assigning H-bond probabilities so that some bonds are more persistent than others with respect to a given window of time. The nature of these correlations ultimately depend on enthalpy-entropy compensation and the tendency of the system to settle into a local minimum. On the other hand, assigning all H-bonds an equal probability removes temporal correlations, but spatial correlations still remain due to clustering preferences of the various species.

The floppy modes, RCS, and fraction of sites in the rigid cluster were plotted as a function of window size for the TCPG and uniform H-bond probability for the UPPG. They were also plotted as a function of MCN and doing so resulted in a collapse of the data onto a "master curve". This collapse suggests that CMHP and CDHP systems at various temperatures and concentrations all belong to a universality class. The likely cause of this universality is that all systems simulated are three dimensional bond bending networks. It would therefore be interesting to test other bond bending networks for this universality.

Choline ions are not expected to facilitate a percolating rigid cluster because they can not form closed loops due to the inability of the Nitrogen in choline to form H-Bonds. Despite this, it was found that choline is present in the percolating cluster, but it does not appear to be crucial for rigidity percolation. Choline ions are dangling rigid clusters which do not seem to significantly participate in any interactions. Nevertheless, they may still play an important role in disrupting the H-Bond network, but the extent of this disruption is difficult to understand from CMHP and CDHP systems alone. This can be tested by replacing Choline ions with smaller cations such as sodium.

The largest cluster histograms show that rigidity primarily percolates within the trehalose matrix, but the next most abundant species is water. This surprising result suggests that water is being trapped in trehalose cages but this alone does not imply that water should be rigid. The respective geometries of water and trehalose must also fit together in such a way that they can form rigid clusters together. It would therefore be interesting to substitute water with a different solvent which is known to bond to trehalose but has a different geometry which may not be conducive to the formation of rigid clusters. Glycerol may be a suitable candidate for such an investigation.

Another important result of this work is that the presence of temporal correlations cause a shift in the rigidity transition to lower MCNs. This has already been shown to occur in two dimensional central force networks which are generated using self-organization or equilibration algorithms. These algorithms attempt to simulate enthalpy-entropy compensation which occurs in physical systems. It has been shown here that the correlations present in MD simulations do result in a better glassforming ability as compared to the mean field approximation (UPPG). This suggests that glasses naturally self-organize in order to percolate rigidity at low MCNs.

Overall, there appear to be few differences between CDHP and CMHP based on rigidity data alone. The fact that rigidity percolation closely coincides with volume minimums from volume hysteresis data suggests that rigidity is a mechanism in the liquid-glass transition. There are additional mechanisms, however, relating to the clustering of water and phosphate which cause the liquid-glass transition to occur at lower solute concentrations in CMHP as compared to CDHP. This thesis also presents evidence that the underlying mechanism for the liquid-glass transition is, in part, a second order rigidity transition which belongs to the same universality class as thermal phase transitions ($\beta = 1/2$).

As an exploratory investigation, this thesis posed more questions than it answered but has cleared many paths for future inquiry. Some of the most interesting questions which deserve attention are related to the role of water in rigidity percolation. It is clear that there are correlations present in the TCPG which do not exist in the UPPG.
To what extent does water introduce correlations and is the network frustrated? Rigidity can only percolate off of rings, so what are the most important rings for rigidity percolation in this system? Are the ring distributions responsible for the subtle shift in the percolation threshold seen in the UPPG data? What will be the effect of substituting trehalose with another disaccharide which is a less effective glass former?

These questions are left to future work. In particular, an analysis of the ring structure may be the best route to elucidate the onset of rigidity percolation as well as the transmittance of stress above the percolation threshold.

REFERENCES

- [1] D. J. Jacobs and M. F. Thorpe *Phys. Rev. E*, vol. 53, no. 4, 1996.
- [2] D. J. Jacobs J. Phys A.: Math. and Gen., vol. 31, pp. 6653–6668, 1998.
- [3] W. Kauzmann Chem. Rev., vol. 43, p. 219, 1948.
- [4] W. H. Zachariasen J. Am. Chem. Soc., vol. 54, pp. 3841–3851, 1932.
- [5] J. C. Phillips J. Non-cryst. Solids, vol. 34, no. 2, pp. 153–181, 1979.
- [6] J. C. Phillips and M. F. Thorpe Solid State Communications, vol. 53, no. 8, pp. 699–702, 1985.
- [7] M. Thorpe, D. Jacobs, M. Chubynsky, and J. Phillips J. Non-cryst. Solids, vol. 266-269, pp. 859–866, 2000.
- [8] M. Thorpe J. Non-Cryst. Solids, vol. 76, no. 1, pp. 109–116, 1985.
- [9] A. Huerta and G. Naumis Physics Letters, Section A: General, Atomic and Solid State Physics, vol. 299, no. 5-6, pp. 660–665, 2002.
- [10] M. Chubynsky, M. Briere, and N. Mousseau Phys. Rev. E, vol. 74, no. 1, 2006.
- [11] W. G. Ellenbroek, V. F. Hagh, A. Kumar, et al. Phys. Rev. Lett., vol. 114, 2015.
- [12] S. Zhang, L. Zhang, M. Bouzid, et al. Phys. Rev. Lett., vol. 123, 2019.
- [13] S. H. Kuang Liu and J. Schwarz Phys. Rev. X, vol. 9, no. 021006, 2019.
- [14] D. Jacobs, A. Rader, L. Kuhn, et al. Proteins: Structure, Function and Genetics, vol. 44, no. 2, pp. 150–165, 2001.
- [15] *PNAS*, vol. 99, no. 6, pp. 3540–3545, 2002.
- [16] H. Flores-Ruiz and G. Naumis *Physica B: Condensed Matter*, vol. 418, 2013.
- [17] M. H. Cohen and G. S. Grest *Phys. Rev. B*, vol. 20, p. 1077, 1979.
- [18] C. Angell, K. Ngai, G. Mckenna, et al. J. Chem. Phys., vol. 88, 2000.
- [19] A. Sokolov, A. Kisliuk, D. Quitmann, et al. J.Non-Cryst. Solids, vol. 172-174, no. 1, pp. 138–153, 1994.
- [20] C. Angell J. Non-Cryst. Solids, vol. 131-133, no. 1, pp. 13–31, 1991.
- [21] A. Napolitano and P. Macedo Journal of Research of the National Bureau of Standards Section A: Physics and Chemistry, vol. 72, no. 4, p. 425, 1968.
- [22] C. Angell and L. Torell J. Chem. Phys., vol. 78, 1983.

- [23] M. Micoulaut Advances in Physics: X, vol. 1, no. 2, pp. 147–175, 2016.
- [24] B. Wunderlich, "The nature of the glass transition and its determination by thermal analysis," in Assignment of the Glass Transition (R. Seyler, ed.), pp. 17– 31, ATSM International, 1994.
- [25] N. Nikulsin, E. A. Singam, G. Elliot, et al. Phys. Chem. Chem. Phys., vol. 20, no. 32, pp. 20899–20909, 2018.
- [26] L. Weng and G. D. Elliot *Pharm. Res.*, vol. 32, pp. 2217–2218, 2014.
- [27] S. Magazu, G. Maisano, F. Migliardo, and C. Mondelli *Biophysical Journal*, vol. 86, no. 5, pp. 3241–3249, 2004.
- [28] E. Cornicchi, G. Onori, and A. Paciaroni Phys. Rev. Lett., vol. 95, no. 15, 2005.
- [29] J. Barre *Phys. Rev. E*, vol. 80, 2013.
- [30] D. J. Koeze and B. P. Tighe *Phys. Rev. Lett*, vol. 121, 2018.
- [31] M. Micoulaut, J.-Y. Raty, C. Otjacque, and C. Bichara Phys. Rev. B, vol. 81, no. 174206, 2010.
- [32] Y. Takeuchi J. Phys.: Conf. Ser, vol. 454, 2013.
- [33] D. Thirumalai, R. D. Mountain, and T. R. Kirkpatrick Phys. Rev. A, vol. 39, no. 7, 1988.
- [34] N. J. Giordano and H. Nakanishi, Computational Physics. Pearson Prentice Hall, 2006.
- [35] E. Paquet and H. L. Viktor *BioMed Research International*, vol. 2015, 2015.
- [36] G. Jaeger Arch. Hist. Exact Sci., vol. 53, pp. 51–81, 1998.
- [37] N. J. Giordano and H. Nakanishi. Pearson Prentice Hall, 2006.
- [38] B. Servatius and H. Servatius, *Rigidity Theory and Applications*. Kluwer Academic/Plenum Publishers, 1999.
- [39] L. Gonzalez, H. Wang, D. Livesay, et al. Algorithms for Molecular Biology, vol. 10, no. 1, 2015.
- [40] P. Lucas, E. A. King, O. Gulbiten, J. L. Yarger, E. Soignard, , and B. Bureau Phys. Rev. B, vol. 80, no. 214114, 2009.
- [41] M. Bauchy Computational Materials Science, no. 159, pp. 95–102, 2019.
- [42] Q. Zheng and H. Zeng Int. J. Appl. Glass Sci., vol. 11, pp. 432–441, 2020.
- [43] F. Weinhold and R. A. Klein *Mol. Phys.*, vol. 110, no. 9-10, 2012.

- [44] R. Kumar, J. Schmidt, and J. Skinner J. Chem. Phys., vol. 126, no. 20, 2007.
- [45] G. Hura, J. M. Sorenson, R. M. Glaeser, and T. Head-Gordon J. Chem. Phys., vol. 113, p. 9140, 2000.
- [46] A. K. Soper *Chemical Physics*, vol. 258, pp. 121–137, 2000.
- [47] S. Pagnotta, M. Ricci, F. Bruni, et al. Chem. Phys., vol. 345, no. 2-3, pp. 159– 163, 2008.
- [48] S. Gehrke and B. Kirchner J. Chem. Eng. Data, vol. 65, pp. 1146–1158, 2020.
- [49] J. Teixeira and M. C. Bellissent-Funel Phys. Rev. A, vol. 31, no. 3, 1984.
- [50] M. Kropman and H. Bakker J. Chem. Phys., vol. 115, p. 8942, 2001.
- [51] D. Prada-Gracia, R. Shevchuk, and F. Rao J. Chem. Phys., vol. 139, no. 8, 2013.
- [52] R. H. Henchman and S. J. Irudayam J. Phys. Chem., vol. 114, pp. 16792–16810, 2010.
- [53] D. Laage and J. Hynes J. Phys. Chem. B, vol. 112, no. 45, pp. 14230–14242, 2008.
- [54] F. W. Starr, J. K. Nielsen, and H. E. Stanley Phys. Rev. Lett., vol. 82, no. 11, pp. 2294–2297, 1999.
- [55] M. Hakala, K. Nygard, S. Manninen, et al. J. Chem. Phys., vol. 125, p. 084504, 2006.
- [56] A. Luzar and D. Chandler *Nature*, vol. 379, pp. 55–57, 1996.
- [57] A. Luzar J. Chem. Phys., vol. 113, p. 10663, 2000.
- [58] J. Farmer and D. J. Jacobs *PLOS ONE*, vol. 13, no. 5, 2018.
- [59] T. S. Pandian, S. J. Cho, and J. Kang J. Org. Chem., vol. 78, no. 23, pp. 12121– 12127, 2013.
- [60] G. Biroli1 and J. P. Garrahan J. Chem. Phys., vol. 138, no. 12A301, 2013.