June 2021

POPULATION ANNEALING: ANALYSIS, OPTIMIZATION AND APPLICATION TO GLASSY SYSTEMS

Christopher A. Amey
University of Massachusetts Amherst

POPULATION ANNEALING: ANALYSIS, OPTIMIZATION AND APPLICATION TO GLASSY SYSTEMS

A Dissertation Presented

by

CHRISTOPHER AMEY

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2021

Department of Physics

POPULATION ANNEALING: ANALYSIS, OPTIMIZATION AND APPLICATION TO GLASSY SYSTEMS

A Dissertation Presented

by

CHRISTOPHER AMEY

Approved as to style and content by:

Jonathan Machta, Chair

Gregory Grason, Member

Narayanan Menon, Member

Nikolay Prokof'ev, Member

Narayanan Menon, Department Chair
Department of Physics

ACKNOWLEDGMENTS

I would like to thank my advisor and mentor Jonathan Machta for his guidance throughout my PhD and his continued support during the 2020 pandemic. I entered Jon's group with a strong interest in classical statistical physics and computational methods and gained a fascination of glass physics and Monte Carlo algorithms. I have benefited from Jon's deep knowledge in a broad range of physics and Monte Carlo topics, and his insightful comments and suggestions. I am proud of the work that we achieved together that is presented in this thesis.

I also thank my colleagues and fellow group members – Vahini Nareddy, Nathan Rose, Jared Callaham, and Tom Shneer – for regularly presenting interesting results during group meetings and for providing insightful questions into my work. Thanks to Martin Weigel for insightful discussions which have led to several new ideas. I am grateful to Olga Goulko, who deserves special praise for reading several rough drafts of my thesis and for providing much-needed feedback and support during the writing process.

Lastly, I would like to thank my family and, in particular, my parents and grandparents, without whom none of this would have been possible. They have been thoroughly supportive of my academic endeavors throughout my college and graduate careers.

ABSTRACT

POPULATION ANNEALING: ANALYSIS, OPTIMIZATION AND APPLICATION TO GLASSY SYSTEMS

MAY 2021

CHRISTOPHER AMEY

B.A., CARLETON COLLEGE

M.Sc., UNIVERSITY OF MASSACHUSETTS BOSTON

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Jonathan Machta

Glasses are physical systems that lack structural order and exhibit extremely slow dynamics, which makes them challenging to study. In this thesis we apply Monte Carlo methods to two distinct glassy systems: the 3D Edwards-Anderson spin glass and a binary hard sphere fluid. While significant progress has been made on theoretical and experimental fronts, much of our current understanding of glasses has come from numerical simulations. Standard Monte Carlo techniques cannot be used to perform equilibrium simulations due to slow dynamics in the glassy regime. As a result, several specialized techniques have been developed in order to simulate such systems, including the main topic of this thesis, population annealing Monte Carlo.

Population annealing is a sequential Monte Carlo algorithm used to perform equilibrium simulations of systems with rough free energy landscapes, such as glasses.

Unlike standard Monte Carlo algorithms, which are unable to overcome large free energy barriers, population annealing is able to sample disparate regions of configuration space in parallel. In this thesis, we discuss the optimization, analysis, and application of population annealing to model glassy systems.

The 3D Edwards-Anderson spin glass has a long history of research, however, the nature of its low-temperature phase remains unclear. We use a carefully optimized canonical ensemble version of population annealing in order to obtain new benchmark values of observables that are important to understanding its low-temperature phase. We also derive and numerically test several useful metrics that provide accurate estimates of the systematic and statistical errors of a simulation.

The binary hard sphere fluid is an example of a system that undergoes a dynamic transition from a fluid to a disordered glassy solid. Whether this system also undergoes a thermodynamic glass transition remains an open question. We use an NVT ensemble version of population annealing in order to simulate the binary fluid at high density and we present two new methods to measure the configurational entropy deep in the glass regime. Using our new numerical techniques, we are able to predict the location and existence of the thermodynamic glass transition.

CONTENTS

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

INTRODUCTION

The modern idea of a phase transition has existed since the nineteenth century, however, in the past several decades, our understanding of the underlying mechanisms has grown enormously. The ideas of symmetry and ordering have been particularly important in describing how changing a single parameter, such as temperature, can result in abrupt changes on a macroscopic scale. Two particularly well-studied systems are the liquid to crystal transitions and the paramagnet to magnet transition. These correspond, respectively, to a change from a state of random particle positions to particles ordered on a lattice, and a change from a state of randomly oriented magnetic moments to uniformly aligned magnetic moments. The basic idea of a transition from disorder to order, or from one ordering to a different ordering, has been successful in describing a large number of systems and phenomena. In this thesis, we will focus on one class of systems that cannot be described by the tools that treat transitions from order to disorder: glasses.

Many systems are described as "glassy", however, there is no clear cut definition of what constitutes a glass. Generally speaking, a glass is a system that undergoes a transition that is characterized by a slowing of dynamics with the simultaneous retention of a disordered state. From the view of phase space, a glass transition is often associated with the formation of an exponentially large (in system size) number of metastable states, and the partitioning of phase space into separate regions that are not dynamically linked, as is the case for mean field spin glasses. This phase space partitioning is also known as ergodicity breaking and is thought to be a signature of glassy behavior in a wide range of physical systems. Here we will focus on two very

different types of glasses that are still not well understood: the 3D Edwards-Anderson spin glass and the binary hard sphere configurational glass.

Unfortunately, dynamic slowing and ergodicity breaking are not simply interesting phenomena, they also make glasses difficult to study numerically and experimentally. In addition, because glassy systems do not undergo phase transitions in the typical sense, the standard tools of statistical mechanics are unsuited for their analysis. One method of study that has been particularly fruitful is numerical simulation and, in particular, the use of annealing and tempering algorithms. In this work, we make use of population annealing Monte Carlo, which is a sequential Monte Carlo algorithm [37] that was first introduced by Hukushima and Iba [54] in the context of spin glasses. Population annealing shares features with two other common algorithms, parallel tempering and simulated annealing, but has the advantage that it can be easily parallelized on a massive scale [14, 15, 34]. In addition to spin glass systems [4, 109–112], there have been several works that have utilized PA for the Ising model, the Potts models [14, 15], and recently, configurational glasses [5, 29].

We begin by briefly reviewing the history and theoretical background of spin glasses in Ch. 1 and configurational glasses in Ch. 2. We then move on to describe Markov chain Monte carlo in Ch. 3 and population annealing in Ch. 4. Our spin glass research results are presented in Ch. 5 and configurational glass results in Ch. 6 and, lastly, we suggest future work in Ch. 7.

CHAPTER 1

SPIN GLASSES

Spin glasses are disordered magnetic systems that, when cooled, undergo an unusual phase transition. They were first discovered experimentally in thin metal films that contained magnetic impurities [30, 80]. While these systems exhibit a cusp in a.c. susceptibility at a certain temperature, there is no corresponding cusp, jump, or singularity in the specific heat commonly found in phase transitions. Furthermore, magnetic neutron scattering shows no sign of long range order in the low temperature phase, and low-temperature samples exhibit aging and memory effects with properties that vary depending on the presence and strength of an external magnetic field during cooling. These characteristics are fundamentally different from most phase transitions that were typically studied at the time. To the physicists of the 1970s, the cusp in susceptibility suggested a transition, yet the specific heat was smooth and there was no apparent symmetry breaking or emergent order. The low-temperature phase appeared to be the same as a paramagnet with one exception – the individual spins remained stuck for long periods of time instead of frequently flipping. Furthermore, aging and memory effects should not exist in an equilibrium system where, according to equilibrium statistical mechanics, the current state of the system should not have a long-lasting memory of past states. These phenomena were absent when the density of magnetic impurities was decreased and so it was clear that the root cause was due to RKKY interactions between the randomly located impurities, as shown diagrammatically in Fig. 1.1. Due to the sign oscillating behavior of RKKY interactions [58, 97, 114],

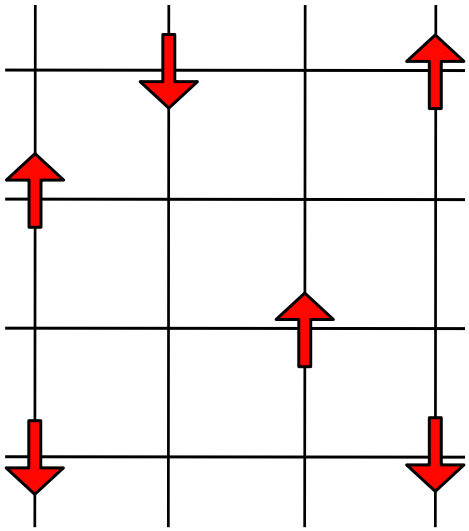

Figure 1.1. The red arrows represent randomly located magnetic impurities in a metallic crystal.

$$J^{\mathrm{3D}}_{\mathrm{RKKY}} = V_0 \frac{\cos(2k_{\mathrm{F}}\, r_{i,j} + \phi)}{r^3_{i,j}}, \tag{1.1}$$

and the random locations of the impurities, the magnetic interactions were effectively random in magnitude and sign resulting in frustration. Random positive and negative couplings make it impossible to choose a configuration of spins that simultaneously satisfies all interactions and, as a result, finding the ground state energy of the system inherently involves finding the optimal trade-off for N different spins. Knowing the mechanisms behind the interactions helps give some qualitative understanding of the underlying physics, but it was still not known whether spin glasses simply corresponded to out-of-equilibrium systems that relaxed with extremely slow dynamics or if they underwent a new sort of true thermodynamic phase transition. To better answer this question, we must turn to spin glass models and the theoretical framework which was developed to study them.

1.1 Spin glass models

The first successful model of spin glasses was invented by Edwards and Anderson in 1975 [38], with a Hamiltonian of the form

$$H = \sum_{<i,j>} J_{i,j} s_i s_j + h \sum_i s_i, \tag{1.2}$$

where the summation is over nearest neighbors on a square lattice, $J_{i,j}$ is drawn from a Gaussian distribution with mean zero and standard deviation unity, and h is an external magnetic field which will be set to zero in this work. Along with the model, Edwards and Anderson suggested an order parameter,

$$q_{\mathrm{EA}} = \frac{1}{N} \sum_{i=1}^{N} \langle s_i \rangle^2, \tag{1.3}$$

where $\langle s_i \rangle$ is the thermal average of spin i. This order parameter is, essentially, an infinite time autocorrelation function of the spins and can effectively distinguish between the paramagnet and spin glass phase despite the fact that neither have long range order. The EA order parameter is related to the probability of each local spin being up or down and, in the paramagnet phase the spins randomly flip with $q_{\mathrm{EA}} = 0$, whereas in the glass phase the spins stay fixed for long periods of time and $q_{\mathrm{EA}} > 0$.

The EA model has one noticeable difference from experimental spin glasses. The disorder of the EA model takes the form of random bonds on a regular lattice while experimental systems have randomly located magnetic impurities. Despite this rather obvious discrepancy, the EA model captures most of the salient features of experimental spin glasses. As it turns out, spin glass models are not particularly difficult to create and only really need two main ingredients: frustration and disorder[1]. It is

[1]There is a lower critical dimension.

interesting to note that frustration alone does not create a spin glass model[2], and that disorder is necessary. Although disorder is necessary, the specific distribution of bonds is not particularly important, although bond distributions that are not centered at zero can result in more complex phase diagrams with ferromagnetic or antiferromagnetic regions. What is vitally important is that the random bonds are "quenched", meaning that they do not change with time. In a spin glass model, every random set of bonds corresponds to a different glass sample, sometimes referred to as a bond or disorder "realization". Each glass sample has a partition function defined in the usual way,

$$Z_{\mathcal{J}} = \sum_{\{\sigma\}} e^{-\beta H_{\mathcal{J}}(\{\sigma\})}, \tag{1.4}$$

where the subscript $\mathcal{J}$ indicates a specific glass sample. The thermal expectation value of an observable is also defined in the standard way,

$$\langle \mathcal{O}_{\mathcal{J}} \rangle = \sum_{\{\sigma\}} \frac{e^{-\beta H_{\mathcal{J}}(\{\sigma\})}}{Z_{\mathcal{J}}} \mathcal{O}_{\mathcal{J}}(\{\sigma\}), \tag{1.5}$$

as is the free energy,

$$-\beta F_{\mathcal{J}} = \log Z_{\mathcal{J}}. \tag{1.6}$$

Since the bonds stay fixed in time, the average over different bond configurations, also known as a disorder average, must come after the thermal average, and

$$[\mathcal{O}] = \sum_{\mathcal{J}} P(\mathcal{J}) \langle \mathcal{O}_{\mathcal{J}} \rangle, \tag{1.7}$$

[2]Geometrically frustrated spin systems lacking disorder had been studied since the 1950s in the context of spin ices [7]. While low-temperature spin ice phases lack long range order and have massively degenerate sets of ground states, they do not have the rough free energy landscape or metastable states of spin glasses.

where $[\mathcal{O}]$ corresponds to the full thermal and disorder average of observable $\mathcal{O}$ and $P(\mathcal{J})$ corresponds to the probability of bond realization $\mathcal{J}$. This is problematic because in order to understand the true thermodynamic behavior of spin glasses, we need to calculate the average free energy,

$$[F_{\mathcal{J}}] = [\log(Z_{\mathcal{J}})], \tag{1.8}$$

which is not analytically tractable with normal tools.

In contrast, if the disorder was not quenched, then we would get "annealed disorder", where the bonds are in thermal equilibrium and can fluctuate with time. In this case, the bonds simply become another degree of freedom of the Hamiltonian and can be easily integrated out,

$$F^{\text{ann}} = \log\left(\sum_{\mathcal{J}} P(\mathcal{J}) \sum_{\{\sigma\}} e^{-\beta H_{\mathcal{J}}(\{\sigma\})}\right) \tag{1.9}$$

$$= \log(Z_{\text{ann}}), \tag{1.10}$$

where Z_{ann} is the annealed partition function after integrating over disorder and F^{ann} is the annealed free energy. There is a way to transform Eq. 1.8 into a form more similar to Eq. 1.9. This technique is called the "replica trick" and follows from the identity [38]

$$\log Z = \lim_{n\to 0} \frac{1}{n}(Z^n - 1). \tag{1.11}$$

The reason why this is named the replica trick is because taking the product of n partition functions, is similar to analytically creating several independent replicas of the original system. Using the replica trick, the quenched free energy simplifies to

$$[\log(Z)] = \lim_{n\to 0} \frac{1}{n}([Z^n] - 1). \tag{1.12}$$

Edwards and Anderson constructed a mean field theory approximation of the EA model and used the replica trick to calculate an approximate free energy. Although their approximation successfully reproduced the qualitative behavior of the susceptibility, it failed to properly predict the behavior of the specific heat [33]. Another approach that proved to be more successful was to design a mean field model directly, which was first done by Sherrington and Kirkpatrick. The Sherrington-Kirkpatrick model [99] is described by the Hamiltonian

$$H = \frac{1}{\sqrt{N}} \sum_{i<j} J_{i,j} s_i s_j + h \sum_i s_i, \tag{1.13}$$

which is infinite range EA model with the addition of a $1/\sqrt{N}$ term, which is necessary to keep the energy extensive in the thermodynamic limit. The Sherrington-Kirkpatrick (SK) model was not created to mimic a real physical system, but rather as a system that was amenable to theoretical calculations and, in this sense, it has been quite successful. Sherrington and Kirkpatrick's original paper purported to solve the system using the replica trick [99], although the solution was found to be unstable in the low temperature phase due to having a negative entropy. In the years that followed, there were several papers on the topic [35, 107] and, ultimately, the model was solved by Giorgio Parisi in a series of papers that fully characterized the mean field transition [89–92]. Although the full solution of the SK model is beyond the scope of this thesis, the implications of the solution are important.

1.2 Mean field solution and its implications

One of the major mysteries of spin glasses when they were first discovered was how they could undergo a phase transition without any apparent breaking of symmetry. Parisi's solution of the SK model directly addresses this mystery – the SK model does not break physical symmetries but it does break replica symmetry [78]. What this

means is not immediately obvious from the name, but is not too difficult to explain. At high temperature, the free energy landscape of a mean field spin glass has only one minimum, which corresponds to a paramagnet. In the low temperature phase, the free energy landscape becomes rough and has many local minima separated by free energy barriers. In the thermodynamic limit these barriers become infinitely high and, as a result, phase space becomes non-ergodic. Each distinct region of phase space is referred to as a thermodynamic pure state and has a weight that is proportional to the phase space region's net Gibbs weight. Although the notion of pure states was invented to describe spin glasses, they also exist in more common models. For instance, the low temperature Ising model has two pure states corresponding to up and down magnetizations and their associated excitations, each of which is separated by an infinitely large free energy barrier in the thermodynamic limit. The low temperature phase of the SK model also breaks ergodicity due to diverging free energy barriers, however, unlike the Ising model, it has an infinite number of pure states. The SK pure states correspond to spin-flip symmetric pairs of spin configurations. Aside from the global spin-flip symmetry, SK pure states are distinct from each other, with no symmetry relations between different pairs. At the same time, the SK pure states are indistinguishable from each other in every way, thermodynamically and statistically. Since there are more than a single pair of spin-flip symmetric pure states, the EA order parameter is inadequate for the SK model since, for pure states α and β, $q_{\mathrm{EA}}^{\alpha} = q_{\mathrm{EA}}^{\beta}$. Instead, define the overlap between pure states α and β as

$$q^{\alpha\beta} = \sum_{i=1}^{N} \langle s_i \rangle_\alpha \langle s_i \rangle_\beta, \tag{1.14}$$

where $\langle s_i \rangle_\alpha$ corresponds to the thermal average of spin i restricted to pure state α. The overlap function is a measure of the distance between two pure states. For example, $q^{\alpha\alpha} = q_{\mathrm{EA}}$ and, if $-\alpha$ is the spin flip symmetric state, $q^{-\alpha\alpha} = -q_{\mathrm{EA}}$ and, more generally, $q^{\alpha\beta} \in [-q_{\mathrm{EA}}, q_{\mathrm{EA}}]$. Since there are infinitely many pure states, the

order parameter that characterizes the SK transition for a glass sample is the full distribution of overlaps [102],

$$P_{\mathcal{J}}(q) = \sum_{\alpha\beta} W_\alpha W_\beta \delta(q - q^{\alpha\beta}), \tag{1.15}$$

where W_α is the Gibbs weight of pure state α and $P_{\mathcal{J}}(q)$ is the overlap distribution for an individual sample. In the high-temperature paramagnet phase, different replicas are completely random and the overlap distribution is approximately a Gaussian centered at zero. In the low-temperature glass phase, replicas become stuck in different pure states. The corresponding overlap distribution has peaks at $\pm q_{\mathrm{EA}}(T)$ and has a support composed of peaks symmetrically distributed around zero from $-q_{\mathrm{EA}}$ to q_{EA}. The two peaks at $\pm q_{\mathrm{EA}}$ correspond to spin flip symmetric pairs of states and the overlap peaks between the $\pm q_{\mathrm{EA}}$ correspond to overlaps between pairs of pure states that are not spin flip symmetric, where the value of $q_{\mathrm{EA}}(T)$ is related to thermal excitations within the pure state. As $T \to 0$, the spins become increasingly unlikely to flip and, at $T = 0$, the spins simply stay fixed in the ground state configuration resulting in $q_{\mathrm{EA}} = 1$. Taking the disorder average of the overlap,

$$P(q) = \left[\sum_{\alpha\beta} W_\alpha W_\beta \delta(q - q^{\alpha\beta}) \right], \tag{1.16}$$

gives the probability of a generic glass sample having overlap $P(q)$ between two randomly chosen pure states. A picture of overlap functions for individual glass samples and the disorder average overlap function is shown in Fig. 1.2. The peaks correspond to different delta functions and the height of each peak is proportional to the overlap weight $W_\alpha W_\beta$.

While the full replica symmetry breaking (RSB) solution proposed by Parisi is beyond the scope of this thesis, we will give a few highlights that are particularly im-

portant[3]. Parisi was able to show that in the low temperature phase of the SK model, $P(q)$ has a countably infinite number of peaks that correspond to overlaps between different pure states and that its corresponding structure is non-trivial. Parisi showed that replicas could be divided into different groups, each of which had different overlap characteristics. This initial grouping corresponds to a one-step replica symmetry breaking, however, Parisi further showed that the SK model was special and that this process could be infinitely repeated. A set of n replicas could be broken into m distinct groups and, likewise, each group of m replicas could be broken into l sub groups, and so on. The algebraic structure that Parisi found is called infinite-step replica symmetry breaking and implies a very rigid structure to the space of pure states[4]. Physically this means that at the glass transition, the free energy landscape forms a large number, say m, of minima. Each minimum is associated with a group of pure states and, in focusing on a single minimum, it becomes apparent that the basin of a single minimum actually corresponds to a large basin containing many local minima. The pure states associated with the larger basin can be further divided into l sub groups each of which correspond to states that have fall within one of the basin's local minima. When looked at closely, there are even more minima within the local minima, which further divide each of the l sub groups. The Parisi solution to the SK model suggests that this process repeats itself infinitely many times and that the relative location of the minima has a specific, rigid form which leads to a prediction of the shape of $P(q)$ that can be tested numerically.

Numerical simulations of the SK model are largely consistent with Parisi's solution and, furthermore, there has been recent mathematical work that has partially justified the mathematical tricks used by Parisi [104–106]. Despite the successes in treating

---

[3]See Ref. [78] for a more complete derivation.

[4]The space of pure states with the overlap q form a metric space that satisfies the strong triangle inequality, otherwise known as an ultrametric space.

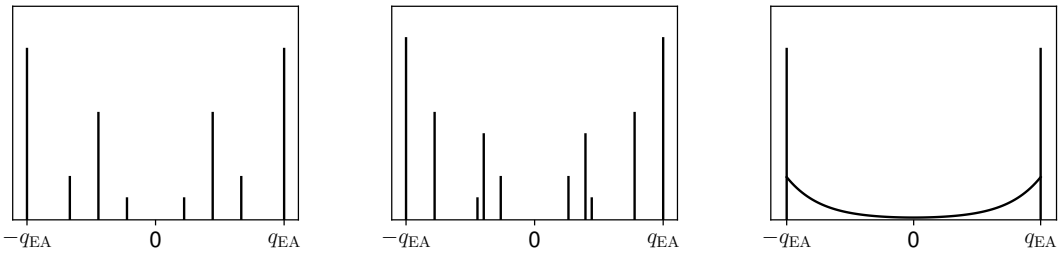

Figure 1.2. The overlap distributions for two spin glass samples at $T < T_c$ (left, middle), and the disorder average overlap function (right).

the SK model, results relating to finite-range spin glass models such as the EA model are still contentious. Proponents of the replica symmetry breaking (RSB) picture claim that finite-range spin glass models also undergo some sort of replica symmetry breaking during a glass transition and have many pure states in the low-temperature phase. There are many competing theories, the most prominent of which is probably the droplet picture. One detail that arises out of the RSB picture is that low energy excitations are extensive in size and correspond to disparate collections of spins across the whole glass sample. The droplet picture, on the other hand, is a phenomenological picture centered on the idea that there exists only a single pair of spin flip symmetric pure states, with excitations that occur in a compact region of size L with excitation energies that scale as L^{θ}, where θ is a scaling exponent. This picture was proposed by McMillan [76], Bray and Moore [28], and Fisher and Huse [40–42] and predicts a two-peaked $P(q)$ distribution. In addition to the RSB and droplet pictures, there are many other proposed scenarios. Newman and Stein, in particular, have argued against the RSB picture and have proposed the chaotic pairs picture [81–83], where a single pair of pure states exists for any finite-size sample and that these pure states change chaotically as the system size is changed. Therefore, while there is only a single pair of states for any given sample and a correspond two-peak $P(q)$, there are an infinite number of pairs of states when considering all sample sizes.

The primary way to differentiate between the different spin glass theories is by performing numerical measurements of $P(q)$ for large system sizes and analyzing the overlap near $q = 0$. The RSB picture would suggest that the distribution has weight near $q = 0$ for $T < T_c$ and the droplet and chaotic pairs pictures suggest that $P(q)$ is a two-peaked distribution. Since computer simulations are limited to finite-sized systems, no free energy barriers diverge and the concept of a pure state is nebulous. In order to overcome these limitations, many simulations are performed for different system sizes in order to measure the trend of $P(q)$ near $q = 0$. Unfortunately, the computational complexity of spin glasses has made it impossible to simulate systems that are large enough to conclusively settle the debate and whether the overlap distribution approaches a Parisi-like solution or a two-peak solution in the thermodynamic limit is still an open question. We report state-of-the-art, high precision measurements of $P(q)$ for the 3D EA model in Ch. 5.

CHAPTER 2

CONFIGURATIONAL GLASSES

The configurational glass transition corresponds to a dynamical phenomenon where, upon cooling, a fluid suddenly exhibits extremely slow dynamics and effectively behaves as an amorphous solid. The glass transition is, by convention, defined as the temperature where the viscosity reaches 10^{12} Poise [19]. This temperature does not correspond to a thermodynamic transition and was chosen from a practical standpoint: fluids with viscosities much larger than 10^{12} Poise cannot be equilibrated within reasonable lab timescales. While the conventionally defined glass transition is somewhat arbitrarily chosen, there are reasons to believe that the onset of slow dynamics may be a precursor to a true thermodynamic glass transition at some lower temperature. As a matter of notation, in this thesis we will use T_{g} to denote the conventional glass transition temperature and T_{d} to denote a dynamical transition temperature in a broader sense. For instance, while a Monte Carlo simulation is guaranteed to sample from equilibrium distributions, the dynamics of the simulation are not necessarily related to the actual physical dynamics of a glassy fluid. In this case, the Monte Carlo simulation would have a different dynamical transition at $T_{\mathrm{d}} \neq T_{\mathrm{g}}$. In a population annealing simulation, the combination of the equilibration process and sweep schedule will result in a specific value of T_{d}.

Configurational glasses have many interesting and unique features in addition to their inability to equilibrate. One particularly important characteristic of configurational glasses is "fragility", which describes a glassy system's relaxation dynamics upon approaching the dynamical transition. Fragility can be determined by plotting

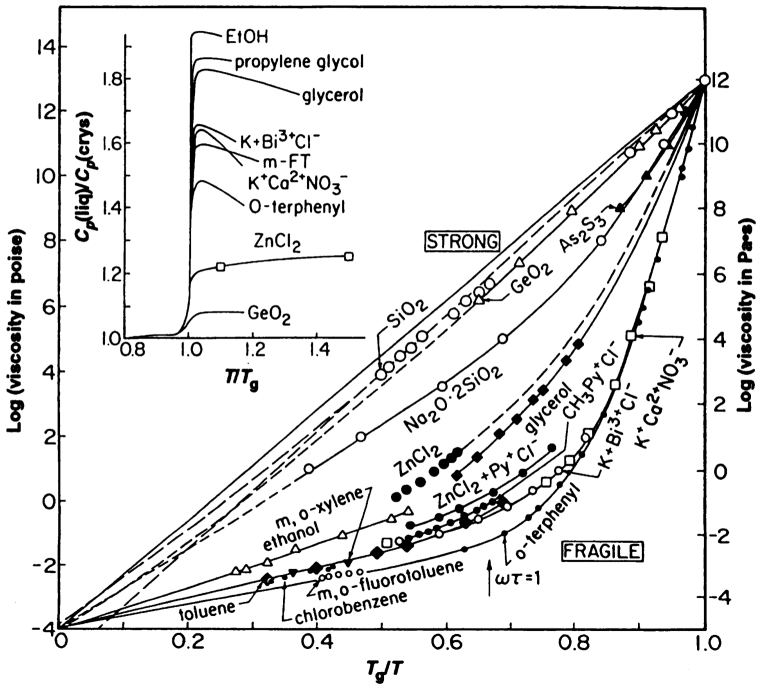

Figure 2.1. The Angell plot showing the logarithm of viscosity vs temperature for several examples of strong and fragile glasses. Here $T_{\rm g}$ corresponds to our $T_{\rm d}$. Reproduced from Ref. [9].

relaxation times or viscosities as a function of temperature for $T \lessapprox T_{\rm d}$, which can be most clearly seen in a plot that was first devised by Angell [9] and is reproduced in Fig. 2.1. A system that approaches the dynamical transition with relaxation times that increase log-linearly is called "strong" and, correspondingly, its relaxation times will obey an Arrhenius law,

$$\tau \sim e^{E_{\rm b}/k_B T}, \tag{2.1}$$

where $E_{\rm b}$ is a constant with dimensions of energy, k_B is Boltzmann's constant, and T is the temperature. In strong glass formers, configurations become stuck in local free energy minima with barrier heights that scale with a characteristic energy, $E_{\rm b}$, which is typically associated with a chemical bond. Examples of strong glasses are silica and ordinary window glass. In strong glasses, the free energy barriers do not grow as the temperature is decreased. Instead, the system becomes stuck in a local minimum due to a lack of thermal energy. The implication of these characteristics is that strong glasses do not undergo a true thermodynamic transition and, instead, simply become stuck in local minima that do not persist in the thermodynamic limit.

In contrast, the relaxation times of fragile glass formers do not scale log-linearly with temperature and as the temperature approaches $T_{\rm d}$, the slope of τ increases. This implies that as fragile glasses approach $T_{\rm d}$, free energy barriers between different minima grow and relaxation times grow super-exponentially. The exact behavior of the relaxation time near and after the fragile glass transition has not been solved for finite-dimensional models, but one widely used fit is the Vogel-Fulcher-Tammann (VFT) law,

$$\tau \sim e^{DT_0/(T-T_0)}, \tag{2.2}$$

where $T_0 < T_\mathrm{d}$ is a fitted temperature that has a non-zero, positive value. Unlike the transition at T_d, which is a dynamical phenomenon, the presence of diverging correlation times at T_0 suggests that a true finite temperature phase transition may exist for fragile glass formers [19]. Although the VFT fit is intriguing, it alone is not enough to stipulate that a thermodynamic transition exists. Several other functions with equivalent goodness of fits have been proposed, see Ref. [25], including a modified VFT law,

$$\tau \sim e^{[DT_0/(T-T_\gamma)]^\gamma}, \tag{2.3}$$

where γ is an additional fit parameter. For many systems, $\gamma = 2$ and $T_\gamma = 0$ work particularly well, which would imply that there is no finite-temperature thermodynamic glass transition. The VFT fit is particularly intriguing because the experimental value of T_0 coincides closely with another temperature that has theoretical importance, known as the Kauzmann temperature, T_K. Typically when a liquid is cooled, it will undergo an entropically driven, first order phase transition into a crystalline state. Glasses are examples of supercooled liquids that have not crystallized and, as a result, have excess entropy when compared to the thermodynamically preferred crystalline phase. Walter Kauzmann [61] experimentally measured the excess entropies of several different glasses as a function of temperature and noticed a curious trend. Some compounds appeared to have excess entropy curves that, when extrapolated, vanished at approximately the same non-zero temperature, T_K. If this was the case then it would correspond to a true thermodynamic glass transition at T_K. With this in mind, the VFT fit with $T_0 \approx T_\mathrm{K}$ takes on additional weight.

There are many different theories that attempt to describe the physics behind glassy fluids at low temperature. One explanation, originally proposed by Gibbs and DiMarzio [45, 46], is that there exists a dynamic ergodicity breaking transition where phase space separates into many dynamically distinct regions at some temperature,

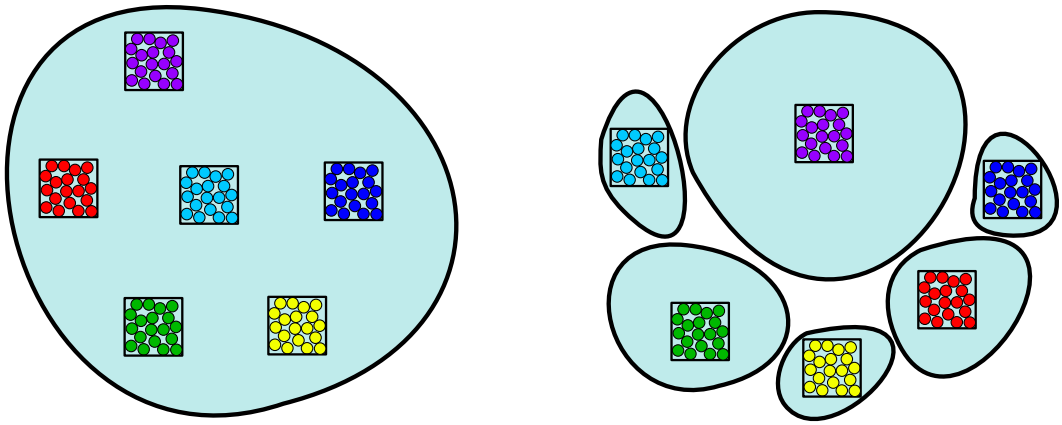

Figure 2.2. Abstract representation of an ergodic fluid phase space (left) and non-ergodic glass phase space (right).

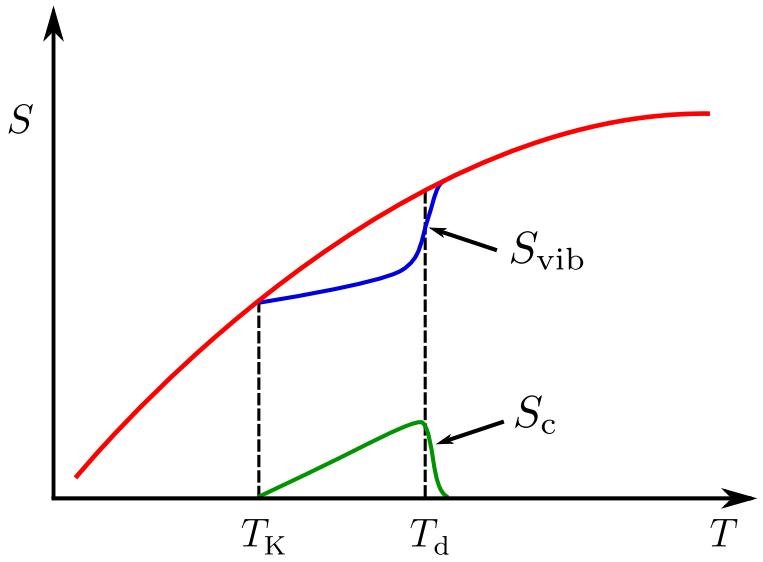

Figure 2.3. Entropies of a fluid undergoing a fragile glass transition.

$T_{\rm d}$. In this picture, when $T > T_{\rm d}$, the system behaves as a fluid that will ergodically sample all of phase space. After the dynamic transition, the mixing is no longer sufficient to equilibrate the system and phase space breaks into different regions, as seen in Fig. 2.2, each of which corresponds to a different glass state. In terms of actual degrees of freedom, particles are unable to diffuse on large length-scales and, instead, remain trapped locally in cages composed of their neighbors in which they vibrate. In this scenario, the phase space volume of each glass state corresponds to the local vibrational degrees of freedom of the particles and, furthermore, each glass state has a corresponding "glass" or "vibrational" entropy resulting from these local particle vibrations. In this thesis we use the term "glass state" to refer to a localized region in phase space and "particle configuration" or "configuration" to refer to a specific set of particle positions within a thermodynamic state. In this terminology, a glass state is made up of many different configurations which are accessible to one another on relatively short time scales. The physics of such a system is dictated by a "configurational entropy per particle", defined as

$$S_{\rm c} = \frac{1}{N} \log \mathcal{N}_g, \tag{2.4}$$

where $\mathcal{N}_g$ is the number of glass states. Understanding how $S_{\rm c}$ decreases as the temperature is decreased and, ultimately, when $S_{\rm c}$ goes to zero is paramount to understanding whether or not a thermodynamic glass transition exists. In the liquid regime, $S_{\rm c} = 0$ since there is only one free energy minimum and $S_{\rm c}$ only becomes positive at $T_{\rm d}$, when the free energy landscape divides into an exponentially large (in system size) number of local minima, as shown in Fig. 2.3. Despite this jump in configurational entropy, $T_{\rm d}$ does not correspond to a thermodynamic transition and there are no thermodynamic signatures of a transition, such as a jump in specific heat. There are, however, dynamic signatures associated with the transition at $T_{\rm d}$, such as a change in the behavior of the mean square displacement (MSD), defined as

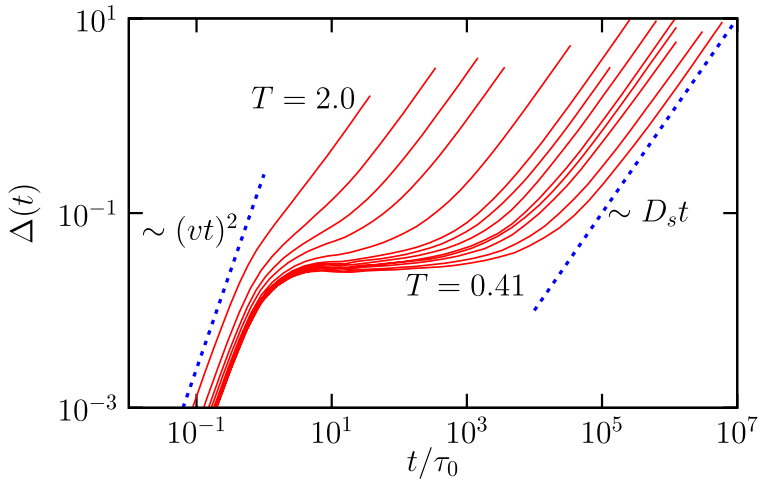

Figure 2.4. Mean square displacement of Lennard-Jones particles as a function of time for several different temperatures. At high temperatures (left), an initial ballistic regime is followed by a diffusive regime. At low temperatures (right), a plateau forms between the two regimes where the particles remain stuck. Figure reproduced from Ref. [19].

$$\Delta(t) = \left\langle \frac{1}{N} \sum_{i}^{N} |\boldsymbol{r}_i(t) - \boldsymbol{r}_i(0)| \right\rangle, \tag{2.5}$$

where $\boldsymbol{r}_i$ is the position of the i'th particle and the brackets correspond to a thermal average. In a simple fluid, particles initially move ballistically with a MSD that scales as t^2. At long times scales after many collisions occur, the particle motion becomes diffusive with a MSD that scales as t. This is true for both the fluid ($T > T_\mathrm{d}$) and the glassy ($T \approx T_\mathrm{d}$) phases, however, in the glassy phase there is an intermediate plateau between the ballistic and diffusive regimes, as seen in Fig. 2.4. This plateau corresponds to slow dynamics and is related to the emergence of a complex energy landscape that has many high-dimensional saddle points with shallow slopes [108]. Because of the complex structure of the energy landscape, particles are initially only able to probe local cages. After a sufficient amount of time, the system is able to find paths required to relax on longer length scales, resulting in the second increase in MSD. As the temperature is decreased, the relaxation times grow and the configurational entropy continually decreases. If S_c goes to zero at a "Kauzmann" temperature,

$T_K > 0$, then this would signal that a thermodynamic transition exists. Therefore, measuring S_c and extrapolating it to zero is a fundamental way to determine the high-density equilibrium behavior of supercooled fluids. In the Adam-Gibbs picture [1], the Kauzmann transition results in the structural relaxation time increasing exponentially with $1/(TS_c)$. As such, another tactic to estimating the value of T_K is to make fits to dynamical quantities, such as structural relaxation times or viscosities, as a function of temperature in order to estimate the location of a dynamic divergence corresponding to the thermodynamic transition. Another possible outcome is that the configurational entropy goes to zero at $T_K = 0$, which would imply that a finite-temperature thermodynamic transition does not occur.

2.1 Configurational glass models

Experimentally, there are many different materials that exhibit a fragile glass transition, including a wide range of molecular fluids, as seen in Fig. 2.1, and colloidal mixtures. A few examples of interactions that are frequently used in numerical simulations are soft spheres, hard spheres, and Lennard-Jones potentials. One interaction parameter that is universal to configurational fluids is the size of the particles. In short range hard or soft sphere systems, the size directly corresponds to the radius of the particle, while in systems with long-range interactions, such as the Kob-Anderson or Lennard-Jones potentials, there is an underlying characteristic length scale that plays the role of particle size. The most basic simple fluid is a system consisting of a single species of particles, which is commonly referred to as a monodisperse fluid. In principle, simple fluids can be composed of all sorts of exotic shapes, but in this thesis we will focus entirely on spherically symmetric particles. Spherical monodisperse particles can exhibit glassy behavior when cooled quickly enough, however, these glass states are metastable, supercooled fluid states, and the true thermodynamic phase is an fcc crystal. As a result, they tend to form either crystalline or polycrystalline

structures. In order to avoid crystallization it is common to introduce frustration in the form of polydispersity, that is, to introduce several species of spheres of different sizes. A frequently studied system is the "bidisperse" or "binary" fluid that contains two species of particles. For some combinations of species radii, it is possible to form uniform crystalline structures [52, 53], but when the radius-ratio of the two species is chosen carefully, the system becomes frustrated and is unable to form a single crystal. The thermodynamic state of such a frustrated binary fluid instead has a first-order transition from a disordered fluid to a phase-separated system of two fcc crystals, each composed of a single species. This phase is extremely difficult to find when starting from a homogeneous mixture and frustrated binary fluids overwhelmingly tend to freeze into disordered states, which makes them ideal systems for studying glasses. It is possible to include more sphere species in the fluid and, in the extreme limit, to have a continuous spectrum of sphere sizes. Recent work with continuously polydisperse hard sphere systems have been able to probe unprecedentedly high densities while remaining in statistical equilibrium [13, 20–22, 87, 88]. These studies suggest that a thermodynamic transition does exist in three dimensions. However, it is not clear that the physics of continuously polydisperse systems is the same as that of binary mixtures and, additionally, there are non-trivial issues with taking the thermodynamic limit of continuously polydisperse systems that need careful consideration.

The goal of this thesis is to study the glass transition using population annealing Monte Carlo. To this end, we focus on relatively well-studied binary hard sphere fluids. At low densities, hard sphere fluids behave as an ideal gas and, as ϕ is increased, the dynamics slow and the fluid eventually crystallizes or becomes glassy, as shown in Fig. 2.5. A hard sphere system can be thought of as a short-range soft sphere system in the zero-temperature limit and, as a result, temperature plays no role because there are no finite energy scales. In soft systems with $T > 0$, both temperature and packing fraction are important in non-trivial ways, and the interplay between

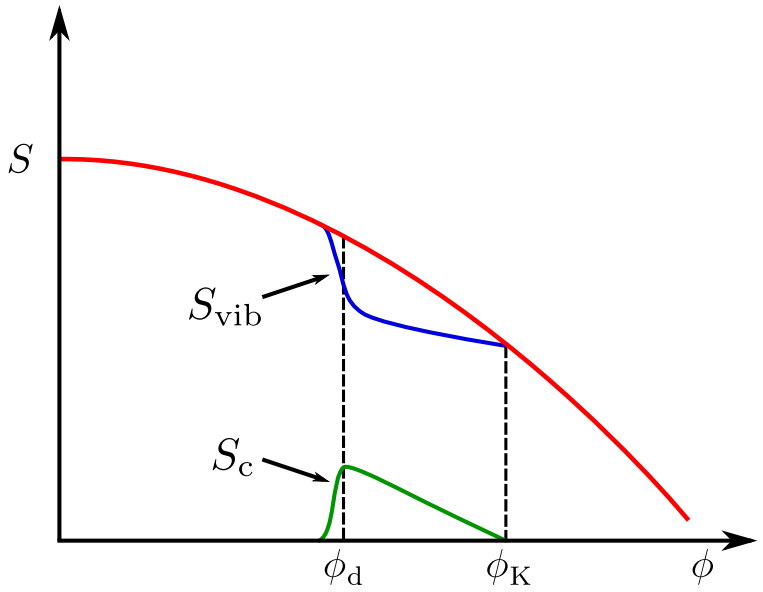

Figure 2.5. Entropy of a fluid undergoing a fragile glass transition with packing fraction as the transition parameter.

the two is an active area of research [71, 86]. Binary hard sphere fluids have been studied numerically [8, 23, 26, 32, 85, 101] and experimentally [26, 75] in the past but, despite the relatively large body of research, a detailed understanding of the physics beyond the dynamic glass transition remains elusive. The hard sphere potential is well-suited to theoretical and numerical studies due to its simple functional form and a binary mixture can be designed to prevent homogeneous crystallization so that it displays a robust glassy regime. In hard sphere systems, temperature is not an independent control parameter and instead the inverse dimensionless pressure, $1/Z = Nk_BT/PV$, is the relevant thermodynamic variable [22]. In most simulations, the dimensionless pressure is an observable and the density or packing fraction, ϕ, is the control parameter. Therefore, the picture one should have in mind for hard sphere glasses is that as the density is increased, the dimensionless pressure increases and eventually diverges once the configuration of particles can no longer be compressed at the random close packing density, ϕ_{rcp} [57]. The question of whether or not a thermodynamic glass transition exists in a hard sphere fluid is the same as the question

of whether or not S_c goes to zero at finite pressure or, equivalently, whether $\phi_K < \phi_{rcp}$ [22, 85].

2.2 Statistical mechanics of fluids

Pairwise interacting simple fluids are described by a generic Hamiltonian,

$$\mathcal{H} = \sum_i \frac{p_i^2}{2m_i} + \sum_{i \neq j} U(x_i, x_j), \tag{2.6}$$

where p_i and m_i are the momentum and mass of the i'th particle, x_i is the center of the i'th particle, and $U(x_i, x_j)$ is the pairwise interaction energy between the i'th and j'th particle. Here we will set all masses to be equal for simplicity because the physics is not significantly altered. After setting the masses to be equal and assuming monodisperse particles, the resulting canonical (NVT) partition function is calculated by integrating the Boltzmann factor over all possible particle and momentum configurations,

$$\mathcal{Z}_{NVT} = \frac{1}{N!}\frac{1}{h^{3N}} \int dx_1 ... dx_N \int dp_1 .. dp_N \, e^{-\beta\left[\sum_i \frac{p_i^2}{2m} + \sum_{i \neq j} U(x_i, x_j)\right]}, \tag{2.7}$$

where N is the total number of particles. The momentum degrees of freedom are separable and can be directly integrated, which gives

$$\mathcal{Z}_{NVT} = \frac{1}{N!}\frac{1}{\lambda_{th}^{3N}} \int dx_1 ... dx_N e^{-\beta \sum_{i \neq j} U(x_i, x_j)} \tag{2.8}$$

where $\lambda_{th} = h/\sqrt{2\pi m k_B T}$ is the thermal deBroglie wavelength and the remaining integral is the portion of the partition function that contains all of the position degrees of freedom. For polydisperse mixtures with different masses, there will be a different λ_{th} for each particle species. The $1/N!$ multiplicative factor applies to a monodisperse mixture that has N indistinguishable particles. In the case of a 50:50 binary mixture, the prefactor will be $[1/(N/2)!]^2$ because there are two species that can be

distinguished from each other. In the glassy limit, the question of whether caged particles are distinguishable will be important when calculating S_{vib} but, for the fluid case, we treat them as indistinguishable. Integrating over the position variables of Eq. 2.8 is generally analytically intractable but the integrand, which corresponds to the weight of each particle configuration, can be slightly simplified in the case of a hard sphere potential where it becomes equal to either 0 or 1 corresponding to whether or not there is a particle overlap. In the hard sphere case, the position integral can be rewritten as

$$\mathcal{Z}_{\text{NVT}} = \frac{1}{N!}\frac{1}{\lambda_{\text{th}}^{3N}}\Omega(\phi), \tag{2.9}$$

where $\Omega(\phi)$ is the volume of configuration space corresponding to non-overlapping particle configurations at packing fraction ϕ. Because each configuration of particles has either weight zero or one, the hard sphere NVT ensemble behaves very similarly to the microcanonical ensemble with packing fraction as a variable instead of energy. An important difference between the NVT and microcanonical ensembles is that in the NVT ensemble, a configuration that is legal at any given packing fraction ϕ_0 is also legal for all $\phi < \phi_0$. This is in contrast to an actual microcanonical ensemble where each configuration corresponds to a unique energy window. Regardless of this difference, as we will see in Ch. 4, the fact that all configurations have weight unity or zero makes this an ideal system to study using microcanonical population annealing.

2.3 Observables

Configurational glasses have many of the same observables as simple fluids, including the equation of state

$$Z = \frac{PV}{Nk_BT}, \tag{2.10}$$

where Z is the dimensionless pressure, P is the system pressure, V is the volume of the system, N is the total number of particles, k_B is the Boltzmann constant, and T is the temperature. In our numerical simulations, Z can be easily measured directly, and N and V are parameters; note that $Z = 1$ for an ideal gas. In a hard sphere system, there are no energy scales and k_BT is simply a constant that only changes λ_{th} and can be set arbitrarily, here we set it to unity. In a typical experiment, particle sizes are fixed and the system volume can fluctuate. Therefore, the packing fraction varies directly with volume and, for a 50:50 binary mixture,

$$\phi = N\frac{4\pi r^3}{3V}, \tag{2.11}$$

where $r^3 = (r_0^3 + r_1^3)/2$ is the average of the cubed radii of the two species. The corresponding differential relation between $d\phi$ and dV is

$$d\phi = -N\frac{4\pi r^3}{3V^2}dV, \tag{2.12}$$

which makes it possible to change variables between V and ϕ for any equation or differential relation. The standard thermodynamic relation between the entropy, $\mathcal{S}$, and pressure holds,

$$\frac{P}{T} = \left(\frac{\partial \mathcal{S}}{\partial V}\right)_{N,T}. \tag{2.13}$$

Changing variables to entropy per particle, S, and packing fraction yields

$$Z = \frac{PV}{Nk_BT} = -\phi\frac{\partial S}{\partial \phi}. \tag{2.14}$$

In the low density fluid limit a binary system is well-estimated by the phenomenological Boublík-Mansoori-Carnahan-Starling-Leland (BMCSL) equation of state [24, 73]

$$Z_{\mathrm{BMCSL}} = \frac{(1+\phi+\phi^2) - 3\phi(y_1 + y_2\phi) - y_3\phi^3}{(1-\phi)^3}, \tag{2.15}$$

where y_i are constants that depend on the ratio of polydispersity. For a 50:50 mixture of 1.4:1 size particles, $y_1 = 0.0513$, $y_2 = 0.0237$, and $y_3 = 0.9251$. This equation of state is very accurate when compared to numerically derived data [29].

A primary goal of configurational glass research is to measure S_{c} in the glassy regime, therefore, measuring entropies is of particular importance. The total entropy per particle can be obtained by integrating Eq. 2.13 from the ideal gas limit at $\phi = 0$ to the packing fraction of interest,

$$S(\phi) = -\int_0^{\phi} \frac{Z}{\tilde{\phi}} d\tilde{\phi} + K, \tag{2.16}$$

where K is a constant of integration. This integral requires some care in order to avoid a logarithmic divergence at $\phi = 0$, where $Z \to 1$, and to ensure proper normalization, see appendix B.1 for details. The vibrational and configurational entropies are significantly more difficult to obtain numerically and are unattainable analytically in finite-dimensional systems. Typically, one measures S and S_{vib} numerically in order to obtain S_{c}. The total entropy can be measured directly using thermodynamic integration and relatively standard numerical techniques. On the other hand, the vibrational entropy is actually the average entropy of the different glass states,

$$S_{\mathrm{vib}} = [S_\nu], \tag{2.17}$$

where S_ν is the entropy of glass state ν, which is the log of the glass's local phase space volume as depicted in Fig. 2.2. The average must be taken carefully, either as a weighted average of entropies of all glass configurations or as an unweighted average taken over all particle configurations, where different particle configurations can correspond to the same glass state and high-entropy glass samples will naturally

have higher weight, as will be discussed in Ch. 6. Measuring the individual glass entropies is non-trivial because of the presence of the dynamic transition and the theory and numerical methods that we have developed will be presented in Ch. 6.

2.4 Mean field connection with spin glasses

Configurational glasses have a rich history of theoretical research that is too large to cover in detail in a thesis, so here we stick to a brief overview of one of the most important frameworks called mode coupling theory (MCT) [17, 49, 70] and its connection to spin glasses. There are some clear surface-level connections between spin and configurational glasses. Both of these systems exhibit a transition from an easy-to-equilibrate regime to a disordered regime with extremely slow dynamics. Furthermore, both systems have a large number of metastable states in their respective glassy regimes and appear to break ergodicity in their glassy phases. However, there are also some clear distinctions. Neither the 3D Edwards-Anderson (EA) nor the mean field Sherrington-Kirkpatrick (SK) models undergo a dynamic phase transition. Instead, they each undergo a continuous phase transition at a well-defined transition temperature. On the other hand, configurational glasses undergo a dynamic transition at some $T_{\rm d} > T_{\rm K}$ and then, eventually, undergo a discontinuous transition at $T_{\rm K}$. In the case of a supercooled fluid, this corresponds to a dynamic transition from a fluid state with one global free energy minimum, to a supercooled state with an exponentially large number of minima to a high density disordered state with a sub-exponential number of minima. Perhaps it was because of these obvious and substantial differences that spin glasses and configurational glasses were researched independently for much of their history. Despite their differences the gap between the two fields was bridged when an unforeseen connection between certain mean field spin glasses and the mode coupling theory of configurational glasses was found in the mid 1980s [67]. The characteristics of a configurational glass transition and, as we shall see, some

mean field spin glass transitions broadly describe the random first order transition (RFOT) which was described in a series of papers by Kikrpatrick, Thirumalai, and Wolynes [64–66].

Mode coupling theory was developed to quantitatively predict the collective dynamics of supercooled liquids at moderately low temperatures. The general idea is that at low temperatures, the behavior of the density field of a fluid is dominated by a few modes that are slow to relax. Therefore, in order to understand the slow dynamics of glass samples as they relax, it is paramount to determine the properties of the slow modes. Since a glass is an out-of-equilibrium system, the relevant observables are dynamic scattering functions, $F(k,t) = \frac{1}{N}\langle \rho_k(0)\rho_{-k}(t)\rangle$, which give information about both spatial and temporal correlations. Mode coupling theory is a framework in which the dominant slow modes can be found, and a set of approximate self-consistent equations used to solve for $F(k,t)$. A detailed treatment of MCT is beyond the scope of this thesis[1], but it is sufficient to state that MCT gives a way to approximate the dynamic structure factor near T_{d} and amounts to a mean field approximation.

At approximately the same time as the creation of MCT, spin glass physicists devised the p-spin model [36, 48], which is a family of mean field spin glass models defined by the Hamiltonian

$$H = -\sum_{i_1,\ldots,i_p} J_{i_1,\ldots,i_p} S_{i_1}\cdots S_{i_p}, \tag{2.18}$$

where the sum goes over all p-tuples of spins and $J_{i_1,\ldots,i_p}$ are quenched random bonds drawn from a Gaussian distribution with zero mean and variance equal to $p!/(2N^{p-1})$. When $p = 2$, this Hamiltonian is identical to the SK Hamiltonian but for $p \geq 3$, the behavior of the p-spin model fundamentally changes. For $p \geq 3$, this model

[1]For a more thorough introduction, see Refs. [50, 94]

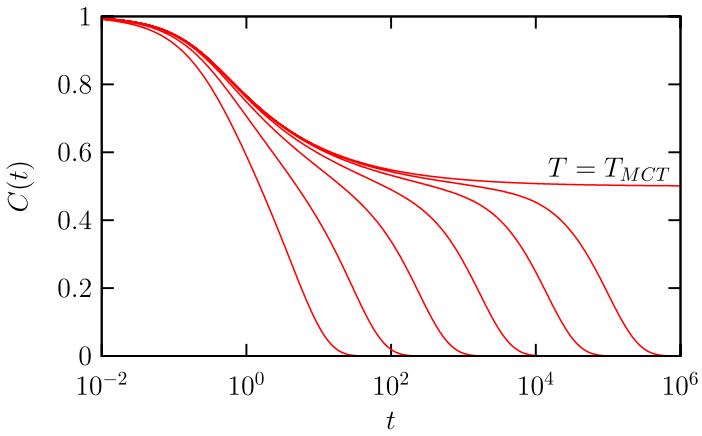

Figure 2.6. The correlation as a function of time of the p-spin model, with $p = 3$, for several temperatures approaching the dynamical transition temperature, T_{MCT}. Reproduced from Ref. [19].

undergoes a one-step replica symmetry breaking and the RSB apparatus can be used to solve for the correlation functions. Surprisingly, the resulting equations for the p-spin model correlation functions are identical to the MCT equations [31]. The solution to these equations yields familiar results including the onset of two-stage relaxation dynamics and a dynamical transition at T_{MCT} with diverging correlation times, as seen in Fig. 2.6. Furthermore, the p-spin model exhibits a roughening of the free energy landscape and the formation of an exponentially large (in system size) number of metastable states. Within the p-spin model, it is possible to calculate the configurational entropy directly and show that a thermodynamic transition occurs when when $S_{\mathrm{c}} \to 0$ at $0 < T_{\mathrm{K}} < T_{\mathrm{MCT}}$. Whether the mean field treatment accurately predicts the qualitative behavior of configurational glasses is a matter of debate, and it is still not known whether configurational glasses undergo a non-zero RFOT.

CHAPTER 3

MONTE CARLO METHODS

Most statistical physics observables are measured in the form of an expectation value with respect to a statistical ensemble. For instance, in the case of a spin system, the expectation value of an observable $\mathcal{O}$ is given by

$$\langle \mathcal{O} \rangle = \frac{\sum_{\{\sigma\}} \mathcal{O}(\sigma) W(\sigma)}{\sum_{\{\sigma\}} W(\sigma)}, \tag{3.1}$$

where the sum is over all possible configurations $\{\sigma\}$, $W(\sigma)$ is the weight of configuration σ, that depends on the statistical ensemble used, and where the normalization/partition function, is given by

$$Z = \sum_{\{\sigma\}} W(\sigma). \tag{3.2}$$

For non-interacting models like the ideal gas and for some mean field models, these calculations can be done analytically, however, for more complicated models, they tend to be impossible to carry out directly. For discrete systems, it is natural to consider direct summation, however, for even a modest sized 3D Ising model with dimensions $10 \times 10 \times 10$, there are $2^{100} \approx 1.2 \times 10^{30}$ different configurations. For a system with continuous degrees of freedom, the notion of directly calculating the corresponding integral is even more far-fetched.

Instead of attempting to calculate expectation values directly, we can estimate the expectation values from a finite subset of states. As long as the states are chosen

so that they represent the proper statistics of the ensemble, the approximation will be unbiased and, with an increasing sample size, will converge to the exact result by the law of large numbers. The basic scheme to do this is simple: randomly choose a state $\sigma_i \in \{\sigma\}$ with probability $P(\sigma_i) \propto W(\sigma_i)$ and then add $\mathcal{O}(\sigma_i)$ to a running sum,

$$\tilde{\mathcal{O}} = \frac{1}{M} \sum_{i=1}^{M} \mathcal{O}(\sigma_i), \tag{3.3}$$

where M is the number of states sampled and $\tilde{\mathcal{O}}$ is the estimator of $\langle \mathcal{O} \rangle$. By repeating this process, the running sum will become increasingly closer to $\langle \mathcal{O} \rangle$ and, in the limit of infinite samples, will converge to the exact theoretical value. Conceptually, this is similar to using a finite number of samples in a histogram to estimate a continuous probability distribution.

Although the general scheme outlined so far is fairly clear, there is still one major piece that is missing: the method by which states are generated with the proper probabilities. In this thesis we will broadly talk about two different approaches: Markov chain Monte Carlo methods (MCMC) and sequential Monte Carlo methods. Generically, a Markov chain Monte Carlo simulation will randomly create a sequence of states whereas a sequential Monte Carlo simulation will create a large ensemble of independent states in parallel.

3.1 Markov chain Monte Carlo

A Markov chain is a sequence of probabilistically chosen states where the probability of choosing a state only depends on the preceding state. In a Markov chain Monte Carlo process, a sequence of states is produced such that, in the long sequence limit, the distribution of states converges to a desired statistical ensemble. Although Markov chains are abstract mathematical constructs, here we will be discussing them in the context of physics simulations, therefore "state" is equivalent to a specific

particle configuration or point in phase space and a Markov chain is a list of different configurations. Likewise, the statistical ensemble to which the Markov chain converges is a statistical mechanics ensemble such as the canonical or microcanonical.

Here we will discuss the principles of Monte Carlo simulations in terms of discrete regions of phase space and for systems with discrete degrees of freedom, such as spin glasses, the proofs made here will be rigorous. Rigorous proofs can also be made for systems with continuous degrees of freedom, however, they do not substantively differ from the proofs for discrete systems and, for the sake of simplicity, have been omitted. In general, if the reader wants to picture a Monte Carlo "state" for a continuous phase space, they should imagine a local phase space density that, in the continuum limit, approaches a delta function. Similarly, ergodic dynamics for a continuous system means that the phase space trajectory of a state will come arbitrarily close to every other state in phase space. With these two rules, a continuous system can be effectively thought of as a discrete system.

The generic outline of a MCMC process is quite simple:

1. Initialize the simulation in random state σ
2. Transition to new state, σ' chosen with probability $P(\sigma \to \sigma')$
   (σ can be equal to σ')
3. Measure observables $\mathcal{O}(\sigma)$, add to running sum
4. Repeat steps $2, 3$

The details of how new states are proposed and then accepted or rejected are strongly related to the resulting distribution of the Markov chain. Put simply, when run for infinite time, the simulation should sample all different states and each state should be visited a number of times that is proportional to that state's statistical weight. The two critical details of a MCMC scheme are the means by which a new state is

proposed and the way that proposed states are accepted or rejected. For the method that proposes a new state, also known as the update, it is only necessary that the dynamics are ergodic given enough time, however, the details of the updating scheme can drastically change the efficiency of a simulation and designing update schemes is often more art than science. In contrast, the way that updated states are accepted or rejected is entirely determined by a single equation known as the balance relation.

Before discussing balance relations, it is helpful to discuss equilibrium simulations. Consider a long Markov chain of states produced from a simulation. If the simulation is in equilibrium, then this chain will approximate the statistical ensemble of states and, if the chain length goes to infinity, will exactly converge to the ensemble of states. More precisely, in a finite length Markov chain, each different state, σ, will be visited a number of times, $N(\sigma)$ such that

$$\frac{N(\sigma)}{\sum_{\{\sigma'\}} N(\sigma')} \approx p_\sigma, \tag{3.4}$$

where $p_\sigma = W(\sigma)/Z$ is the ensemble probability of σ and where the approximation becomes exact as the chain becomes infinitely long. In a more physical sense, the Markov chain will sample states in phase space with a density that is proportional to the ensemble weight of each state. Since a statistical ensemble is, by its very nature, an equilibrium object, it does not dynamically evolve. Therefore, a Markov chain Monte Carlo simulation that is in equilibrium must also be static. If we take the entire Markov chain and apply the probabilistic evolution to each state then, in the infinitely large chain limit, the resulting distribution must be identical to the original. Although the distribution will remain constant, the states in the chain will have changed due to the Markovian process. Looking at an individual state of an equilibrium Markov process more closely, we know that N_σ must remain constant and, therefore, the total probability of state σ updating into another state via the

Markov process must be exactly compensated with other states updating into σ. Mathematically, a Markov process being in equilibrium means that [84]

$$\sum_{\sigma} p_{\sigma} P(\sigma \to \sigma') = \sum_{\sigma} p_{\sigma'} P(\sigma' \to \sigma), \tag{3.5}$$

where the probabilities p_{σ} are given by the ensemble weight and

$$\sum_{\sigma'} P(\sigma \to \sigma') = 1, \tag{3.6}$$

meaning that all states must flow somewhere. In order to ensure that p_{σ} are equal to the ensemble probability $p_{\sigma} = W(\sigma)/Z$, the values of the transition probabilities must be chosen such that these two relations are upheld. These equations can be combined into a more compact form known as the global balance equation,

$$p_{\sigma} = \sum_{\sigma'} p_{\sigma'} P(\sigma' \to \sigma), \tag{3.7}$$

although for our immediate purposes, the previous form is more convenient to use. The intuition behind an equilibrium Markov process is identical to that behind a static fluid. A fluid being static does not mean that the fluid molecules are stationary. Instead, it means that while molecules can flow into and out of different regions, they must do so at exactly the same rate. If they did not do so at the same rate, then there would be bulk movement within the fluid. Likewise, if a Markov chain is not in equilibrium, then the probability distribution of different states will not be static because the simulation will gradually converge to the limiting distribution.

There are many ways to choose transition probabilities such that Eq. 3.7 is satisfied, the simplest of which is called detailed balance. A system satisfies detailed balance if

$$p_{\sigma} P(\sigma \to \sigma') = p_{\sigma'} P(\sigma' \to \sigma), \tag{3.8}$$

for all states σ. At first glance this may appear to be a repeat of what was stated before, but there is a subtle difference between Eq. 3.5 and Eq. 3.8. Detailed balance specifies that the net flow between any two states must be zero while the more general statement of global balance specifies that the net flow in and out of any specific state must be zero. With detailed balance, the flow from state σ to σ' must be exactly equal to the flow from state σ' to σ. One could imagine another flow that does not satisfy detailed balance but is still in equilibrium according to global balance. For instance, states could flow from σ to σ', from σ' to σ'' and so on, as long as the flow eventually closes back to σ. These kinds of situations are known as loops and, although there are cases where loops can be utilized to efficiently sample phase space, they tend to be difficult to design while ensuring that global balance is satisfied. On the other hand, designing an algorithm that satisfies detailed balance is usually significantly easier. Although there are many different MCMC algorithms that are used every day, we will focus on two of the most common, the Metropolis algorithm and event chain Monte Carlo.

3.1.1 Metropolis-Hastings algorithm

One of the simplest MCMC techniques to implement is the Metropolis-Hastings algorithm which was first stated and used by Nicholas Metropolis in 1953 [77] and later extended by Wilfred Hastings in 1970 [51]. This algorithm gives a generic way to create a sample from a desired target probability distribution, $Q(\sigma)$, provided we have a function $f(\sigma)$ that is proportional to P. For instance, in the case of the canonical ensemble, $f(\sigma) = e^{-\beta E(\sigma)}$ would be a valid choice since $Q(\sigma) = e^{-\beta E(\sigma)}/Z$. The general steps of the Metropolis-Hastings algorithm are

1. Initialize the simulation to an arbitrary state
2. Given current state σ, propose a new state σ' with probability $P_{\text{prop}}(\sigma'|\sigma)$
3. Accept new state with probability $P_{\text{acc}}(\sigma'|\sigma) = \min\left(1, \frac{P_{\text{prop}}(\sigma|\sigma')}{P_{\text{prop}}(\sigma'|\sigma)} \frac{f(\sigma')}{f(\sigma)}\right)$

4. If new state was accepted, set σ to σ'

5. Measure observables

6. Repeat steps 2–5,

where $P_{\text{prop}}(\sigma'|\sigma)$ must satisfy the ergodicity criterion. This scheme is very similar to the generic MCMC algorithm, however, there is one important difference. Before, we discussed generic probability flows between different states, whereas the Metropolis-Hastings algorithm gives a specific prescription which satisfies detailed balance. The proof can be easily seen by noting that the probability of transitioning from state σ to σ' is simply the probability of proposing an update from σ to σ', $P_{\text{prop}}(\sigma'|\sigma)$, multiplied by the probability of accepting the update, $P_{\text{acc}}(\sigma'|\sigma)$. Plugging this into the detailed balance equation gives

$$P_{\text{prop}}(\sigma'|\sigma)\, P_{\text{acc}}(\sigma'|\sigma)\, p_\sigma = P_{\text{prop}}(\sigma|\sigma')\, P_{\text{acc}}(\sigma|\sigma')\, p_{\sigma'}. \tag{3.9}$$

rearranging yields

$$\frac{P_{\text{acc}}(\sigma'|\sigma)}{P_{\text{acc}}(\sigma|\sigma')} = \frac{P_{\text{prop}}(\sigma|\sigma')\, p_{\sigma'}}{P_{\text{prop}}(\sigma'|\sigma)\, p_\sigma}. \tag{3.10}$$

Because $p_\sigma = Q(\sigma) \propto f(\sigma)$, this ratio can be further simplified to

$$\frac{P_{\text{acc}}(\sigma'|\sigma)}{P_{\text{acc}}(\sigma|\sigma')} = \frac{P_{\text{prop}}(\sigma|\sigma')}{P_{\text{prop}}(\sigma'|\sigma)} \frac{f(\sigma')}{f(\sigma)}. \tag{3.11}$$

By defining P_{acc} as

$$P_{\text{acc}}(\sigma'|\sigma) = \min\left(1, \frac{P_{\text{prop}}(\sigma|\sigma')}{P_{\text{prop}}(\sigma'|\sigma)} \frac{f(\sigma')}{f(\sigma)}\right), \tag{3.12}$$

this equality is always upheld and, therefore, detailed balance is ensured.

The Metropolis-Hastings algorithm can be used to sample generic probability distributions, but in this thesis, we are primarily interested in its use to sample ensembles related to physical systems such as the Ising model or hard sphere fluids. A simple version of the Metropolis algorithm applied to the Ising model with N spins goes as:

1. Initialize the spins randomly
2. Randomly choose one spin from σ and propose to flip it to get σ'
3. Accept the spin flip with probability $\min\left(1, e^{-\beta[E(\sigma')-E(\sigma)]}\right)$
4. If proposed flip was accepted then update σ
5. Repeat steps 2–4

Since $P_{\text{prop}}(\sigma'|\sigma) = P_{\text{prop}}(\sigma|\sigma') = 1/N$, the acceptance probability becomes the ratio of Boltzmann weights of the two configurations. For high temperatures, this algorithm works quite efficiently at sampling spin configurations. This is simply because in the high temperature limit, thermal energy dominates and the configurational energy is mostly irrelevant. Therefore, all configurations have nearly the same probability and ensuring an equilibrium sampling is nearly trivial. At lower temperature there are possible issues. For one, the existence of a phase transition is typically accompanied by a critical slowing of dynamics and a diverging of length scales. Another problem is that at temperatures below T_c, local energy barriers dominate the free energy landscape and algorithms that perform "local" updates, such as individual spin flips, become inefficient. The underlying issue is easy to see in the Ising model. The low temperature phase consists of two thermodynamic states that are separated by a diverging free energy barrier. To move across the barrier, many spins have to flip, however, individual spin flips are energetically unfavorable and are usually rejected.

3.1.2 Event chain

Event chain Monte Carlo (ECMC) is an algorithm that was developed by Krauth et al. [18] to study hard disc fluids and was later extended to general pairwise-interacting fluids [79]. Unlike the Metropolis algorithm, each ECMC move corresponds to a sequence of particle movements called "events". There are many different versions of ECMC, but we will focus on a version that uses straight event chains which, when used with hard spheres is rejection free. The outline of a single event chain is that a particle is chosen randomly and is moved in a random direction until it strikes another particle. The original particle then stops and the new particle is moved in the same direction until it strikes a third particle. This sequence is continued until the total displacement of the particles reaches an arbitrary, preset length, ℓ, as shown in Fig. 3.1. Since this algorithm moves a chain of particles instead of looking for non-overlapping displacements, it is rejection free. ECMC also satisfies detailed balance because each "forward" event chain can be reversed by selecting the lasts particle and moving it in exactly the opposite direction. This "reverse" chain exactly undoes the forward chain particle moves and brings the configuration back to its initial state. Since the initial particle and direction of the chain are chosen from uniform distributions, the probability of proposing a forward and reverse chain are exactly the same.

By virtue of being rejection free, ECMC has obvious advantages over the Metropolis algorithm, however, the full power of ECMC only comes when detailed balance is broken. The balanced version of ECMC, can still fail to sample phase space efficiently in some circumstances. Even if a forward chain is not exactly reversed, if any of the particles in the forward chain are chosen in the next step and moved backwards, then they will be moved back to their initial positions, effectively moving the configuration back near its original point in configuration space and resulting in a high autocorrelation. This may seem unlikely, but in a high-density case, where a chain may contain

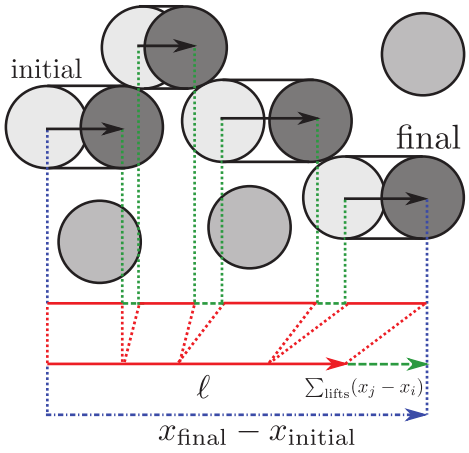

Figure 3.1. Diagram of a single event chain, reproduced from Ref. [79].

many particles, it can become an issue where after doing a long chain, another chain reverses many of the original particle moves. One remedy to this is simple: break detailed balance. In a system with periodic boundary conditions, only allowing moves in the forward direction will still result in a fully sampled configuration space and will not allow configurations to diffuse back towards their initial positions. The global balance version of ECMC has been thoroughly probed by Krauth, et al, and has been found to be extremely efficient at sampling hard particle configurations at high densities when compared to other algorithms that satisfy detailed balance.

3.1.3 Autocorrelation times

So far we have discussed the way to design a MCMC simulation that converges to the correct stationary distribution and have completely neglected the dynamics related to convergence and, more generally, how to accurately estimate the statistical and systematic errors associated with the measurement of an observable. In a MCMC simulation, the principle source of systematic errors comes from a lack of equilibration. In the infinitely long Markov chain limit, as long as Eq. 3.5 is upheld, the sample will have the correct statistics and the associated errors will go to zero. However, simulations never have infinite resources and, as a result, we need to ensure that

each simulation takes an equilibrium sample of states to make measurements. The primary challenge in doing this is that MCMC sequences nearly always have non-zero correlations. Therefore, after each measurement, it is necessary to iterate the Monte Carlo process many times in order to get a new, uncorrelated measurement. A simulation that is well-equilibrated will have made many uncorrelated measurements so that no major regions of phase space have been unsampled. It is possible to quantify the level of equilibration in a simulation and the number of Monte Carlo steps necessary to decorrelate measurements by using autocorrelation times. It is important to emphasize that we have, to some extent, abused the word "time" since we are not measuring time in seconds or minutes, but in units of Monte Carlo steps. In a Monte Carlo simulation, time is usually parameterized in "sweeps". In a system with N particles, one sweep corresponds to proposing Monte Carlo updates to N particles. It is not necessary that the updates are accepted or that each particle is offered an update, just that the Monte Carlo dynamics has worked on N different particles. This is a particularly useful parameterization because it is intrinsically intensive. In general, autocorrelation times are derived from the two-time autocorrelation function, which is defined as the correlation function of an observable with itself at a later time in the simulation,

$$C_{\mathcal{O}}(t) = \langle \mathcal{O}_{t_1} \mathcal{O}_{t_2} \rangle - \langle \mathcal{O}_{t_1} \rangle \langle \mathcal{O}_{t_2} \rangle \tag{3.13}$$

$$= \langle \mathcal{O}_{t_1} \mathcal{O}_{t_2} \rangle - \langle \mathcal{O} \rangle^2, \tag{3.14}$$

where $C_{\mathcal{O}}(t)$ is the autocorrelation function of observable $\mathcal{O}$ for time $t = t_2 - t_1$, and $t_{1,2}$ correspond to different sampling points in our Markov sequence. Note that the zero-time autocorrelation function is simply the variance, $C_{\mathcal{O}}(0) = \mathrm{var}(\mathcal{O})$. For $t \to \infty$, the asymptotic behavior of the autocorrelation function is expected to be

$$C(t) \sim e^{-t/\tau_{\exp}}, \tag{3.15}$$

where τ_{exp} is the exponential autocorrelation time and the $\mathcal{O}$ has been omitted. The number of independent measurements that can be made in a MCMC simulation scales as $t_{\text{sim}}/\tau_{\text{exp}}$, where t_{sim} is the total simulation time. Put otherwise, if optimally sampled, the systematic errors of a MCMC simulation scale as $\sqrt{\tau_{\text{exp}}/t}$. A simple way of estimating τ_{exp} involves taking many measurements over the course of a simulation and performing an exponential fit to the data, however, more advanced Fourier methods are typically more efficient. One major complication can arise if the dynamics of a MCMC simulation has many competing time scales. When several of the longest time scales are close to each other, many Monte Carlo steps will be required to resolve the true asymptotic behavior of $C(t)$.

While τ_{exp} gives information about systematic errors, a corresponding autocorrelation time related to statistical errors, τ_{int}, can also be measured. This is done by estimating the standard deviation of the mean of the relevant observable [56],

$$\sigma^2_{\overline{\mathcal{O}}} = \langle \overline{\mathcal{O}}^2 \rangle - \langle \overline{\mathcal{O}} \rangle^2 \tag{3.16}$$

$$= \frac{1}{t^2_{\text{sim}}} \sum_{i=1}^{M} \left(\langle \mathcal{O}^2_{t_i} \rangle - \langle \mathcal{O}_{t_i} \rangle^2 \right) + \frac{1}{t^2_{\text{sim}}} \sum_{i \neq j}^{M} \left(\langle \mathcal{O}_{t_i} \mathcal{O}_{t_j} \rangle - \langle \mathcal{O}_{t_i} \rangle \langle \mathcal{O}_{t_j} \rangle \right), \tag{3.17}$$

where M is the number of measurements made, t_{sim} is the total number of sweeps performed, and $t_{i,j}$ are the times of the i'th and j'th measurements. If the measurements are completely uncorrelated, then $\langle \mathcal{O}_{t_i} \mathcal{O}_{t_j} \rangle = \langle \mathcal{O}_{t_i} \rangle \langle \mathcal{O}_{t_j} \rangle$, and the second term of Eq. 3.17 cancels, yielding

$$\sigma^2_{\overline{\mathcal{O}}} = \frac{1}{t^2_{\text{sim}}} \sum_{t_i=1}^{t_{\text{sim}}} \left(\langle \mathcal{O}^2_{t_i} \rangle - \langle \mathcal{O}_{t_i} \rangle^2 \right) \tag{3.18}$$

$$= \frac{1}{t_{\text{sim}}} \left(\langle \mathcal{O}^2_{t_1} \rangle - \langle \mathcal{O}_{t_1} \rangle^2 \right) \tag{3.19}$$

$$= \frac{C_{\mathcal{O}}(0)}{t_{\text{sim}}} \tag{3.20}$$

$$= \frac{\text{var}(\mathcal{O})}{t_{\text{sim}}}, \tag{3.21}$$

which is simply the standard definition of the standard deviation of the mean for uncorrelated samples. However, unless many Monte Carlo sweeps are made between each measurement, say $100 \times \tau_{\text{exp}}$, the assumption that simulation data is uncorrelated is unrealistic. Furthermore, uncorrelated data is neither efficient nor necessary to properly estimate observables or statistical errors, provided that the full form of Eq. 3.17 is used. Expanding and reordering Eq. 3.17 gives

$$\sigma_{\bar{\mathcal{O}}}^2 = \frac{C_{\mathcal{O}}(0)}{t_{\text{sim}}} \left[1 + 2 \sum_{i=1}^{M} \left(\frac{\langle \mathcal{O}_{t_1} \mathcal{O}_{t_{i+1}} \rangle - \langle \mathcal{O}_{t_1} \rangle \langle \mathcal{O}_{t_{i+1}} \rangle}{C_{\mathcal{O}}(0)} \right) \left(1 - \frac{t_i}{t_{\text{sim}}} \right) \right], \tag{3.22}$$

where the first term in the sum can be identified as the two-time autocorrelation function. Simplified, this becomes

$$\sigma_{\bar{\mathcal{O}}}^2 = \frac{2C_{\mathcal{O}}(0)}{N} \tau_{\text{int}}, \tag{3.23}$$

where τ_{int} is the integrated autocorrelation time defined as

$$\tau_{\text{int}} = \frac{1}{2} + 1 \sum_{i=1}^{M} \frac{C_{\mathcal{O}}(t_i)}{C_{\mathcal{O}}(0)} \left(1 - \frac{t_i}{t_{\text{sim}}} \right). \tag{3.24}$$

The sum in this equation is equivalent to integrating the autocorrelation function, which is what gives τ_{int} its name. The standard error of the mean of $\mathcal{O}$ is directly related to τ_{int} by

$$\sigma_{\bar{\mathcal{O}}} = \sqrt{\frac{2C_{\mathcal{O}}(0)}{t_{\text{sim}}} \tau_{\text{int}}}. \tag{3.25}$$

With these two time scales, it is possible to determine whether a simulation has sampled enough of phase space to reduce systematic errors and to make an accurate measurement of an observable (τ_{exp}) and whether enough data has been collected to make an precise measurement of the observable (τ_{int}). Although these tools allow

make it possible to determine whether or not a simulation is in equilibrium, they do not help to speed the equilibration. For a standard statistical mechanics system, such as the Ising model or a simple fluid, the Monte Carlo algorithms presented so far, in combination with the methods to estimate errors, will be enough to properly equilibrate a simulation and measure observables. However, as discussed in the first chapter, glassy systems have a number of interesting properties that also make them particularly difficult to simulate. Extremely slow dynamics resulting from frustration, and an exponentially large number of metastable states are particularly troublesome for a Monte Carlo simulation. The methods discussed so far will not work to properly sample the phase spaces of such systems and more advanced Monte Carlo algorithms become necessary.

3.2 Tempering and annealing algorithms

3.2.1 Parallel tempering

In cases where there are multiple different thermodynamic minima, using a simple form of MCMC like the Metropolis algorithm may not be good enough, and more complex tools become necessary. For instance, the Ising model in the low temperature phase has two thermodynamic minima, however, the Metropolis algorithm discussed in the previous section is unable to move between them and the simulation will become trapped in either the up or down state. For an Ising model without a field this is not actually a problem; the two minima are equivalent in all ways and it is sufficient to only probe one minimum at a time. However, if a small field were added to the Ising model, then a simulation running a simple Metropolis algorithm runs the risk of becoming stuck in the wrong state that does not correspond to the true thermodynamic minimum. For glasses, this problem is exacerbated further because there is typically an exponentially large number of local minima, each of which is surrounded by large, or even diverging, free energy barriers. In these cases, there

is little hope of finding the true thermodynamic behavior from a simple Metropolis simulation and special techniques need to be used.

Parallel tempering (PT) is a MCMC algorithm that was designed to probe systems that have rough free energy landscapes, such as glasses. The general idea behind a PT simulation is relatively straightforward and is shown diagrammatically in Fig. 3.2 for the case of the canonical ensemble and where a transition occurs after some temperature T_c. Instead of running a single simulation at a single temperature, several independent simulations are run in parallel at different temperatures, $\beta_0 < \beta_1 < ... < \beta_N$. These simulations can take the form of a fixed-temperature Monte Carlo scheme such as the Metropolis algorithm, or even a molecular dynamics simulation. At periodic intervals, configurations at neighboring temperatures, $\beta_j < \beta_i$, will be prompted to swap positions with probability

$$p = \min\left(1,\, e^{\Delta E \Delta\beta}\right), \tag{3.26}$$

where $\Delta\beta = \beta_i - \beta_j$ and $\Delta E = E_i - E_j$. The swap probability used for parallel tempering satisfies detailed balance and ensures that the simulations at different temperatures remain in thermal equilibrium. Between swaps, the single-temperature simulations are used to further equilibrate the configurations and to accrue data for the local temperatures. Eventually, as the simulation time increases, more local minima are probed in the low temperatures and equilibrium measurements are made at all different temperatures. Like any MCMC process, the errors associated with measured observables decrease as the simulation time increases and autocorrelation times are used to estimate the equilibration of data obtained at each temperature step. Of course there are many details required to make an efficient simulation, such as optimal placement of the temperature steps and choice of equilibrating scheme, however, those are beyond the scope of this thesis. It is important to note that this technique is not specific to the canonical ensemble or with using temperature

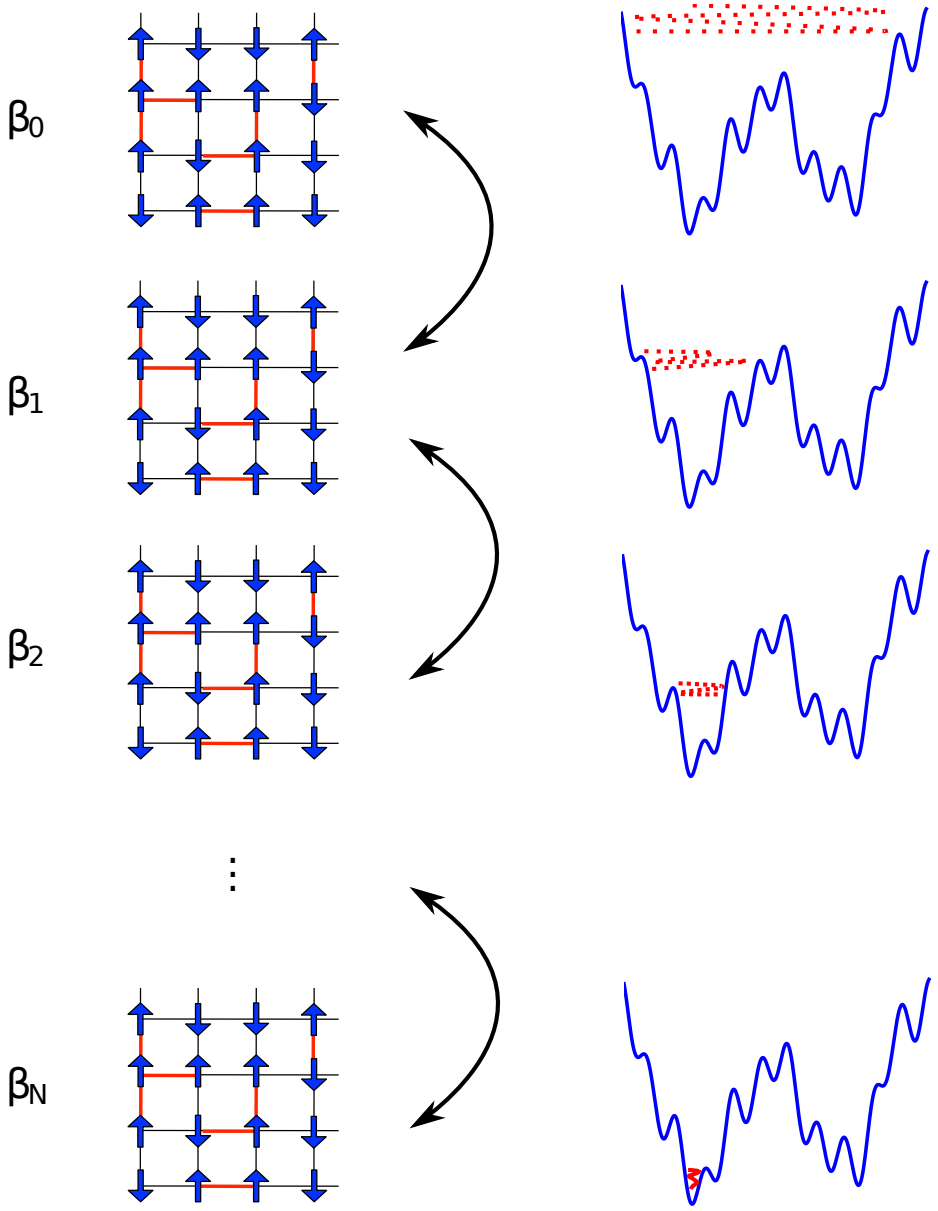

Figure 3.2. Diagram of a parallel tempering simulation. The blue curve corresponds to the free energy landscape of the system, the red dashed line corresponds to the Monte Carlo dynamics at a specific temperature, and the temperature schedule is given by $\beta_0 < \beta_1 < \beta_2 < ... < \beta_N$.

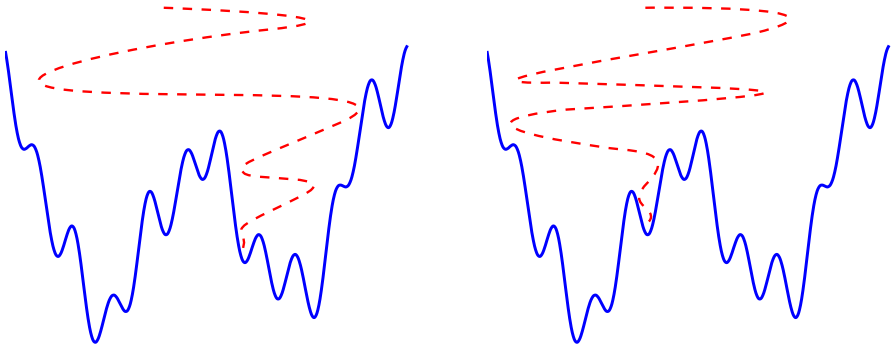

Figure 3.3. Diagram of a simulated annealing simulation. The simulation is initialized in an easy-to-equilibrate region of parameter space, for instance a high temperature, and is gradually annealed towards a difficult-to-equilibrate region of parameter space such as a low temperature phase. Eventually, each simulation gets stuck in a localized region of phase space associated with a local free energy minimum.

as the tempering parameter. The algorithm is easily generalizable and, for instance, one could temper in density or chemical potential or any number of thermodynamic parameters.

3.2.2 Simulated annealing

There are times when only the ground state of a system is needed and, in these cases, simulated annealing (SA) [63] is the simplest Monte Carlo method that can be used. In the context of the canonical ensemble, a simulated annealing algorithm is quite straightforward: initialize a simulation at high temperature and, while running MCMC or some other ergodic dynamics, gradually anneal towards the low temperature regime. Eventually each simulation will become trapped in a local minimum of the free energy which, as $T \to 0$, corresponds to a local minimum of the energy. This scheme is shown diagrammatically in Fig. 3.3.

Simulated annealing must be repeated many times and, as the number of runs goes to infinity, the probability that the true ground state and energy minimum will be found goes to unity. This process works because, by gradually changing the thermodynamic parameter while running MCMC, the simulation is able to better

probe the free energy landscape and is able to avoid getting stuck in unfavorable local minima. Throughout the annealing process, MCMC is constantly being run and, therefore, the simulation tends to sample regions with higher statistical weight more often. While simulated annealing can efficiently find ground states, because the temperature is constantly changing, the simulation is never in thermodynamic equilibrium and finite-temperature observables cannot be measured, although there are ways to take equilibrium measurements if the proper weighted average is used, see Ref. [95]. If one were to run a large number, or "population", of simulated annealing simulations in parallel and, after each annealing step, weight and resample the population so that it was approximately in thermal equilibrium, then one would be able to make equilibrium measurements of observables while retaining the benefits of annealing. An algorithm that fits this description is population annealing Monte Carlo.

CHAPTER 4

POPULATION ANNEALING

Population annealing Monte Carlo (PA) is a sequential Monte Carlo algorithm that can efficiently measure equilibrium properties and find ground states of systems with rough free energy landscapes. Population annealing was originally devised by Hukushima and Iba [54] and falls into the class of particle filters and evolutionary genetic algorithms, where an ensemble of simulations is carried out simultaneously in order to estimate a probability distribution. One noteworthy feature of population annealing is that it naturally lends itself to massive parallelization.

Population annealing is a generic technique that can be applied to any statistical ensemble, but in this chapter we will mainly focus on the canonical form of the algorithm with temperature as the annealing parameter, as described in Refs. [72,110]. Glassy systems that are particularly well-suited to the canonical ensemble are spin glasses or soft configurational glasses, where the free energy landscape undergoes a sudden increase in complexity as the temperature is decreased. An example of PA being used with another ensemble will be shown in Ch. 6, where we look at a hard sphere fluid in the NVT ensemble with density as the annealing parameter. As seen in Ch. 2.4, the hard sphere NVT ensemble is similar to the microcanonical ensemble, but there are some subtle and important differences. A true microcanonical ensemble implementation of PA that was applied to the Potts model can be found in Ref. [95] and is similar, in many respects, to another sequential Monte Carlo algorithm named nested sampling [10, 11, 74, 93, 100]. Further examples of the canonical ensemble version of PA can be found in Refs. [4, 14, 15, 34, 69, 109–112].

4.1 Algorithm outline

Population annealing is similar to a parallel version of simulated annealing in many respects: a population of R configurations, henceforth called "replicas" or "population members", is initialized in an easy-to-equilibrate region of parameter space, and the population is slowly annealed towards a more difficult target region. In addition, there is a resampling step after each annealing step where individual replicas may be copied or eliminated according to their ensemble weights. Resampling ensures that the ensemble of replicas is kept near statistical equilibrium, however, the resampled ensemble has degeneracies due to copied replicas, and does not fully sample the low-energy spectrum at the lower temperature. In order to address these issues, an equilibrating procedure such as Markov Chain Monte Carlo or molecular dynamics is used to decorrelate and additionally equilibrate the population. The generic outline of the algorithms can be summarized as

1. Initialize R independent simulations in an easy-to-equilibrate region of parameter space

2. Equilibrate and decorrelate population, typically with MCMC or MD

3. Change annealing parameter
   (Population is now out of equilibrium)

4. Weight and resample the population to better approximate ensemble
   (Population is now approximately in equilibrium)

5. repeat steps 2-4.

This process is illustrated in Fig. 4.1. Although a PA simulation can traverse through a multi-dimensional parameter space, it is more common to anneal in a single parameter. In the canonical ensemble, the annealing schedule is composed of N_T inverse

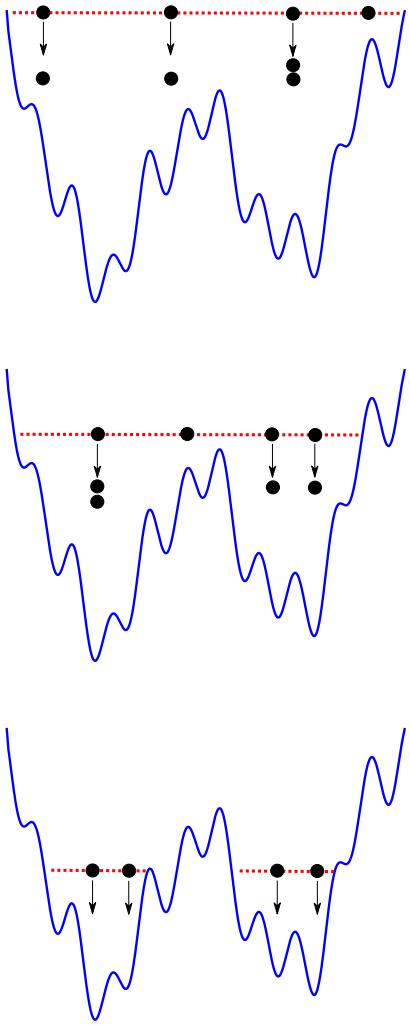

Figure 4.1. Diagram of a population annealing simulation. The simulation is initialized in an easy-to-equilibrate region of parameter space, for instance a high temperature (top), and is gradually annealed towards a difficult-to-equilibrate region of parameter space such as a low temperature phase (bottom). After each annealing step, the population is resampled according to the weight of each configuration so as to stay, approximately, in thermal equilibrium. After resampling, the population is then further equilibrated, locally in phase space, using MCMC. With this scheme, a population annealing simulation can sample many local minima even in the low-temperature, glassy phase.

temperatures $\{\beta_{N_T-1}, ..., \beta_0\}$, where $\beta_{j+1} < \beta_j$, where the initial temperature is typically infinite, $\beta_{N_T-1} = 0$, and the final β_0 is chosen as desired.

Consider the resampling step from inverse temperature β to β'. Each replica, denoted by a subscript i, is given a resampling factor τ_i that is proportional to the ratio of the Boltzmann weights of that replica at the two temperatures,

$$\tau_i = \frac{R}{\tilde{R}_\beta} \frac{e^{-(\beta'-\beta)E_i}}{Q(\beta, \beta')}, \tag{4.1}$$

where E_i is the energy of replica i, Q is a normalization factor used to control the population size,

$$Q(\beta, \beta') = \frac{1}{\tilde{R}_\beta} \sum_{i=1}^{\tilde{R}_\beta} e^{-(\beta'-\beta)E_i}, \tag{4.2}$$

and $\tilde{R}_\beta$ is the population size at step β. The actual population size $\tilde{R}_\beta$ may fluctuate around the target population size R. The resampling factor is the expected number of copies of a replica, that is, $\tau_i = \langle n_i \rangle$, where n_i is the stochastically chosen integer number of copies of a replica i. There are several ways to implement resampling in PA. In this work, when dealing with the canonical ensemble, we follow the method used in Ref. [110], which minimizes the variance of n_i by choosing $n_i(\beta, \beta') = \lfloor \tau_i(\beta, \beta') \rfloor$ with probability $\lceil \tau_i(\beta, \beta') \rceil - \tau_i(\beta, \beta')$ and $n_i(\beta, \beta') = \lceil \tau_i(\beta, \beta') \rceil$ otherwise. This method results in an average population equal to R, with fluctuations of order $\sqrt{R}$. The resampling step ensures that, for large R, if the population is an equilibrium ensemble at β, it will also be an equilibrium ensemble at β'. One important aspect of the resampling step is the fraction of the population that is eliminated and replaced via resampling, which is dubbed the "culling fraction". This fraction will depend on the size of each annealing step, with a larger step resulting in a larger fraction being eliminated.

The normalization factor $Q(\beta, \beta')$ is actually an estimator of $Z(\beta')/Z(\beta)$ [72],

$$Q(\beta, \beta') = \frac{1}{\tilde{R}_\beta} \sum_{i=1}^{\tilde{R}_\beta} e^{-(\beta'-\beta)E_i} \tag{4.3}$$

$$\approx \langle e^{-(\beta'-\beta)E_\sigma} \rangle \tag{4.4}$$

$$= \sum_\sigma e^{-(\beta'-\beta)E_\sigma} \left(\frac{e^{-\beta E_\sigma}}{Z(\beta)} \right) \tag{4.5}$$

$$= \frac{\sum_\sigma e^{-\beta' E_\sigma}}{Z(\beta)} \tag{4.6}$$

$$= Z(\beta')/Z(\beta), \tag{4.7}$$

where Eq. 4.4 is the exact thermal average and the summation over σ is over all configurations. The weight factor at each step can be thermodynamically integrated to give a free energy estimator,

$$-\beta_k \tilde{F}(\beta_k) + \beta_{N_T-1} \tilde{F}(\beta_{N_T-1}) = \log Z(\beta_k) - \log Z(\beta_{N_T-1}) \tag{4.8}$$

$$= \log \left(\frac{Z(\beta_k)}{Z(\beta_{k+1})} \frac{Z(\beta_{k+1})}{Z(\beta_{k+2})} \cdots \frac{Z(\beta_{N_T-2})}{Z(\beta_{N_T-1})} \right) \tag{4.9}$$

$$= \sum_{l=k+1}^{N_T-1} \log \left(\frac{Z(\beta_{l-1})}{Z(\beta_l)} \right) \tag{4.10}$$

$$\approx \sum_{l=k+1}^{N_T-1} \log Q(\beta_l, \beta_{l-1}). \tag{4.11}$$

In a typical PA simulation, we initialize the population in a regime where the partition function has a trivial analytic form. For an Ising spin glass, where we will use this algorithm, annealing begins at $\beta_{N_T-1} = 0$ where all configurations are equally weighted and the value of the partition function is $\Omega = 2^N$. The resulting free energy estimator is

$$-\beta_k \tilde{F}(\beta_k) = \sum_{l=k+1}^{N_T-1} \log Q(\beta_l, \beta_{l-1}) + \log \Omega, \tag{4.12}$$

where $\tilde{F}$ is the estimator of the free energy.

Observables from different independent simulations can be combined with a weighted average defined as

$$\overline{\mathcal{O}} = \frac{\sum_{m=1}^{M} \tilde{\mathcal{O}}_m R_m \exp[-\beta \tilde{F}_m]}{\sum_{m=1}^{M} R_m \exp[-\beta \tilde{F}_m]}, \tag{4.13}$$

where the index m iterates over different simulations. This method of averaging is particularly useful because as the number of runs increases, both statistical and systematic errors decrease. The intuition behind the weighted average formula is relatively simple. If one had unlimited computing resources, there would be no culling of the population and, instead, the population would grow after each annealing step. The population is kept from growing in an unbounded manner purely for the sake of computational limitations. If we were to allow a simulation's population to increase without bound then it would grow as $R \exp[-\beta \tilde{F}]$, and a large population would imply that the corresponding simulation found a particularly high-weight area of phase space. The formula for the weighted average assigns a weight to each simulation according to its projected population size. In the limit of a population of size one, this becomes identical to a weighted average of different simulated annealing simulations[1].

4.2 Error analysis

Because it is a sequential Monte Carlo algorithm, the error analysis in population annealing is fundamentally different than that in a Markov chain Monte Carlo algorithm. The equilibration of a MCMC simulation is determined by comparing run time to equilibriation simulation times, see Ch. 3.1.3, and as the simulation time goes to infinity, the measurements converge to the exact result. Population annealing, on

[1]For a more complete discussion on the comparison between weighted simulated annealing and population annealing, see Ref. [95].

the other hand, samples phase space with a population all in parallel and, therefore, converges to the exact result as $R \to \infty$. Instead of equilibration simulation times τ, the comparable quantity in a PA simulation is equilibrium population sizes ρ, and a simulation is well-equilibrated when $\rho/R \ll 1$, with the exact equilibration criteria depending on the specific application. Just as there are many different equilibration autocorrelation times in an MCMC simulation, there are also many different equilibrium population sizes in a PA simulation. Here we define the errors of a PA simulation and three associated equilibrium population sizes, ρ_f, ρ_t, and ρ_s, that were originally derived and defined in Ref. [110].

The systematic errors of a population annealing simulation arise from not having a large enough population to properly sample the phase space. In a system with several local minima, a PA simulation with too small of a population will not sample statistically important regions of phase space and the resulting observable measurements will be biased. A well-equilibrated PA simulation will sample all statistically important regions of state space and, as a result, will give consistent measurements of the free energy observable. In the large R limit, phase space is well-sampled so generic observables, $\mathcal{O}$, and the free energy estimator can be assumed to be a bivariate Gaussian. Using this assumption, the systematic error associated with measuring observable $\mathcal{O}$ is given by [110]

$$\Delta\mathcal{O} = \beta \operatorname{cov}(\tilde{\mathcal{O}}, \tilde{F}). \tag{4.14}$$

Furthermore, we can define an equilibrium population size ρ_f as

$$\rho_f = \lim_{R\to\infty} R \operatorname{var}(\beta\tilde{F}). \tag{4.15}$$

This limit is generally believed to be well-defined because of a central limit theorem type argument. Equation 4.2 can be approximated as a sum of approximately R

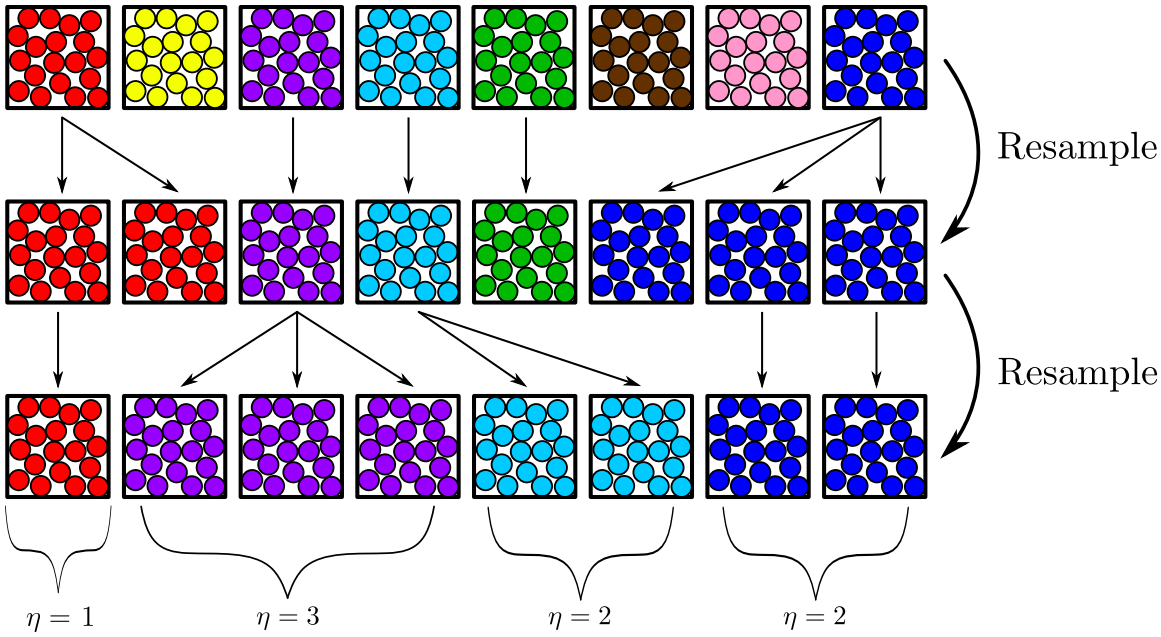

Figure 4.2. Diagram of a population annealing simulation over several annealing steps. Each family is color coded and corresponds to a statistically correlated subset of the population. As annealing and resampling take place, the families sizes fluctuate according to the internal correlations of the population.

random variables and the free energy is simply the sum of $\log Q$ for subsequent annealing steps. Altogether, the free energy estimator behaves as a sum of R random numbers, and the variance of such a quantity, by the central limit theorem, scales as $1/R$ as long as the simulation is well-equilibrated and all of the variables are independent. In general, if a region of parameter space is easy to equilibrate using MCMC, then $\rho_f \lessapprox 1$ and a simulation that cannot be equilibrated using MCMC will result in $\rho_f > 1$. With this last point it is important to emphasize that although ρ_f will converge to a constant, its specific value for a PA simulation will depend on the precise annealing protocol. A PA simulation with many MCMC sweeps will tend to equilibrate more effectively and, as a result, will have a smaller value of ρ_f.

The statistical errors in a population annealing simulation originate from a population that is finite in size, which is further exacerbated by correlations within the population that appear after resampling. In easy-to-equilibrate regions of parameter

space, copies of configurations can be easily decorrelated using few MCMC sweeps, however, in difficult-to-equilibrate regimes, where MCMC fails to fully equilibrate, the copies remain correlated to varying degrees. In order to keep track of these correlations, we introduce the notion of "families" within the population, where members of the population that originated from the same initial state are said to be in the same family, and each family, i has size η_i, as seen in Fig. 4.2. At the beginning of a PA simulation, there are R families, each of size unity. As the simulation progresses, the number of families remains R but many of the families die off entirely and have zero members. A bound on the statistical errors can be found by assuming that replicas within a family are 100% correlated,

$$\mathrm{var}(\tilde{\mathcal{O}}) \leq \mathrm{var}(\mathcal{O}) \sum_{i=1}^{R} \eta_i^2, \tag{4.16}$$

where the index runs over all of the families and $\mathrm{var}(\mathcal{O})$ refers to the variance of the observable in the Gibbs distribution. Note that at the beginning of the simulation, when all families have a single member, this becomes an equality. An equilibrium population size can be defined from the family distribution as

$$\rho_t = \lim_{R\to\infty} \frac{1}{R} \sum_{i=1}^{R} \eta_i^2 \tag{4.17}$$

or, equivalently [4], as

$$\rho_t - 1 = \lim_{R\to\infty} \mathrm{var}(\eta). \tag{4.18}$$

Note that in a population that is entirely uncorrelated, all families are size 1 and $\rho_t = 1$, while a correlated population will result in a several larger families and $\rho_t > 1$. The statistical errors associated with observable $\mathcal{O}$ are bounded by

$$\delta\tilde{\mathcal{O}} \leq \sqrt{\frac{\mathrm{var}(\mathcal{O})\rho_t}{R}}. \tag{4.19}$$

The correlations within a population and the systematic errors of a simulation are related in a non-trivial manner, and a relation between ρ_t and ρ_f will be discussed in Ch. 5. In addition to ρ_t, the distribution of family sizes can be used to define another equilibrium population size. The entropy of the family size distribution is defined as

$$S_f = -\sum_{i=1}^{R} \frac{\eta_i}{R} \log \frac{\eta_i}{R}, \tag{4.20}$$

where e^{S_f} is a measure of the effective number of independent families in a simulation. The corresponding equilibrium population size ρ_s is defined as

$$\rho_s = \lim_{R\to\infty} \frac{R}{e^{S_f}}. \tag{4.21}$$

These metrics of error are generic to all population annealing simulations, however, as we will see in Ch. 6, there may be additional measures of equilibration that are system-specific. Additionally, the specific form of ρ_f can change depending on the ensemble used; in the canonical ensemble, $\rho_f \propto \mathrm{var}\beta\tilde{F}$, while in the microcanonical ensemble, $\rho_f \propto \mathrm{var}S$.

CHAPTER 5

SPIN GLASS SIMULATIONS

5.1 Introduction

In this chapter we discuss the results of the project on the application and optimization of population annealing to spin glass systems. The work presented here is a combination of algorithmic improvements on the canonical version of PA and new computational results on the 3D Edwards-Anderson spin glass model. We begin by giving a brief recap of the EA model and its related observables that can be measured using population annealing. We then derive a fundamental relation between ρ_f and ρ_t that intrinsically relate the statistical and systematic errors in a PA simulation, following which we present a scheme to optimally simulate a large set of disorder samples. Finally, we present new numerical results for the 3D EA spin glass model and results that support our optimizations. This contents of this chapter is based on work published in Ref. [4].

5.1.1 Model and Observables

We study and apply PA in the context of the 3D Edwards-Anderson (EA) spin glass, defined by the Hamiltonian

$$H = -\sum_{\langle n,m \rangle} J_{nm} s_n s_m, \tag{5.1}$$

where the summation is over nearest neighbors on a cubic lattice with periodic boundary conditions, $s_n = \pm 1$ are Ising spins, and J_{nm} are bonds drawn from a Gaussian

distribution with zero mean and unit standard deviation. In order to understand the thermodynamic properties of the EA model, it is necessary to do many simulations with different bond configurations and to then take an average afterwards. PA gives access to the ensemble average of observables for a fixed set of bonds, therefore many PA simulations must be conducted to take a bond average. An average of an observable $\mathcal{O}$ over bond configurations will be denoted as $[\mathcal{O}]_{\mathcal{J}}$, whereas a thermal average will be denoted as $\langle\mathcal{O}\rangle$.

There are several types of observables that are theoretically interesting in spin glasses. Observables that can be measured directly in a single spin configuration are the easiest to measure since PA effectively simulates the canonical ensemble at each temperature step. Thermodynamic quantities for a single bond configuration, such as the average energy $E = \langle H\rangle$ or average magnetization, are all straightforward to measure as a simple population average at each temperature step.

The order parameter for the EA model is obtained from the overlap between two spin configurations chosen independently from the canonical ensemble at a given temperature in the same bond configuration. The spin overlap q is defined as

$$q = \frac{1}{N}\sum_{n=1}^{N} s_n^1 s_n^2, \tag{5.2}$$

where s_n^1 is the n^{th} spin of replica 1, and N is the total number of spins in the system. The thermal distribution of overlaps $P(q)$ for a given temperature and bond configuration has support on $[-1, 1]$. The Edwards-Anderson order parameter, q_{EA}, is the thermal average of the absolute value of q in (any) single pure thermodynamic state. In order to measure $P(q)$, it is necessary to measure q many times from spin configurations 1 and 2 drawn independently from the equilibrium ensemble of replicas. This process is straightforward in PA as long as the family of each replica is recorded: PA gives access to the equilibrium ensemble, and replicas are guaranteed to be independent if they are from different families.

An important observable in spin glass theory is the disorder-averaged integrated overlap weight around the origin,

$$\mathrm{I}(q_0) = \left[\int_{-q_0}^{q_0} P(q) dq \right]_{\mathcal{J}}, \tag{5.3}$$

where the value of q_0 is somewhat arbitrary. Following previous studies [60, 110], we use $q_0 = 0.2$. The quantity $\mathrm{I}(q_0)$ was introduced to distinguish between competing theories of the low temperature spin glass phase. Replica symmetry breaking theory [89, 91, 92] predicts that $\mathrm{I}(q_0)$ goes to a nonzero constant as $N \to \infty$ for any $q_0 > 0$ while the droplet [27, 40, 41, 76] and chaotic pairs pictures [81] predict that $\mathrm{I}(q_0) \to 0$ as $N \to \infty$ for any $q_0 < q_{\mathrm{EA}}$.

The link overlap q_l is another quantity that is defined from two independent spin configurations,

$$q_l = \frac{1}{N_b} \sum_{\langle n,m \rangle} s_n^1 s_m^1 s_n^2 s_m^2, \tag{5.4}$$

where $N_b = 3N$ is the number of bonds. Like the spin overlap, the link overlap is useful for distinguishing theories of the spin glass phase [60]. It is also useful as a measure of equilibration using a relation found by Katzgraber and Young [60]. They define

$$\Delta_{\mathrm{KY}} = [e_l - e]_{\mathcal{J}}, \tag{5.5}$$

where e is the thermally averaged energy per spin and e_l is an energy-like quantity defined from the link overlap,

$$e_l = -\frac{1}{T} \frac{N_b}{N} (1 - \langle q_l \rangle). \tag{5.6}$$

For an individual bond configuration, it is not the case that $e_l = e$, however, for a disorder average over Gaussian bonds, $\Delta_{\mathrm{KY}} = 0$. The requirement that $\Delta_{\mathrm{KY}} \approx 0$ is a useful indicator of equilibration.

If PA is run to a sufficiently low temperature we can obtain an estimator of the ground state energy $\tilde{E}_0$ from the lowest energy encountered during the simulation. Population annealing also gives direct access to the estimated probability of being in the ground state at each temperature, $\tilde{g}_0$, which is simply the fraction of the population at the lowest energy found during the entire simulation. Alternatively, with the aid of the free energy estimator, it is possible to obtain an indirect estimate of the probability of being in the ground state, $\bar{g}_0$, by calculating the Boltzmann weight of the ground state,

$$\bar{g}_0 = 2\, e^{-\beta E_0 + \beta \tilde{F}}. \tag{5.7}$$

By comparing these two estimators of the probability of being in the ground state, we can asses systematic errors within our simulation.

5.2 Population Annealing Theory

5.2.1 Error Estimation

As shown in [110], the systematic error of an observable $\mathcal{O}$ is given by

$$\Delta\mathcal{O} = \mathrm{var}(\beta\tilde{F}) \left[\frac{\mathrm{cov}(\tilde{\mathcal{O}}, \beta\tilde{F})}{\mathrm{var}(\beta\tilde{F})} \right], \tag{5.8}$$

where $\tilde{\mathcal{O}}$ is the PA estimator of $\mathcal{O}$ and the (co)variances are taken with respect to independent runs of PA. The bracketed quantity is expected to converge to a finite limit as $R \to \infty$, meaning that for large R, the systematic error for any observable

is proportional $\mathrm{var}(\beta\tilde{F})$. Furthermore, the quantity $\mathrm{var}(\beta\tilde{F})$ is expected to scale as $1/R$, and so it is natural to define an equilibration population size, ρ_f, as

$$\rho_f = \lim_{R\to\infty} R\,\mathrm{var}(\beta\tilde{F}). \tag{5.9}$$

The equilibration population size sets a population scale for a given bond configuration, and by choosing the population such that $R \gg \rho_f$, systematic errors behave as ρ_f/R. One complication is that in order to measure ρ_f, many simulations of the same bond configuration must be made.

An analogous quantity for statistical errors, ρ_t, can be defined that corresponds to the population size required to minimize statistical errors. In PA, if no decorrelating Markov chain Monte Carlo were done, the statistical errors would directly scale with the second moment of the family distribution, see [110]. Therefore we define ρ_t as

$$\rho_t = \lim_{R\to\infty} \frac{1}{R}\sum_{i=1}^{R} \eta_i^2, \tag{5.10}$$

where the summation is over families and η_i is the size of family i. Note that we can also express ρ_t in terms of the variance of the family size distribution. Since the average family size is one,

$$\rho_t - 1 = \lim_{R\to\infty} \mathrm{var}(\eta). \tag{5.11}$$

Because we do perform MCMC at each annealing step, the actual statistical errors will be lower and ρ_t can be used as a conservative estimate for the population size necessary to minimize statistical errors. Specifically, the statistical error $\delta\mathcal{O}$ in measuring an observable $\mathcal{O}$ in a PA run with population size R is bounded by

$$\delta\mathcal{O} \leq \sqrt{\frac{\mathrm{var}(\mathcal{O})\rho_t}{R}} \tag{5.12}$$

where var($\mathcal{O}$) here refers to the underlying variance of the observable in the Gibbs distribution.

Unlike ρ_f, ρ_t can be easily estimated from a single run and, as will be shown in the next section, there is a close relationship between ρ_f and ρ_t that can be used to our advantage.

5.2.2 Relation between ρ_t and ρ_f

In this section we provide an argument justifying the inequality

$$\rho_t - 1 > \rho_f. \tag{5.13}$$

The argument uses a modified version of PA where the exact free energy is known and is used for normalizing the weight of each spin configuration. This modified version of PA is similar to the idea of using blocks of a much larger total population to calculate ρ_f and ρ_t. In this version of PA, the weight of spin configuration i is given by

$$\tau_i = e^{-(\beta'-\beta)E_i+\beta'F'-\beta F}, \tag{5.14}$$

where F and F' are the exact free energies at β and β', respectively. The actual number of copies n_i of configuration i is a random non-negative integer, $n(\tau_i)$, whose mean is τ_i. Our implementation of $n(\tau)$ is given in Ch. 4 but the details are not important to the argument. Here we divide n_i into τ_i and a random remainder $z(\tau_i)$, whose mean is zero,

$$n_i = \tau_i + z(\tau_i). \tag{5.15}$$

The total population $\tilde{R}_{\beta'}$ at temperature step β' is given by

$$\tilde{R}_{\beta'} = \sum_{i=1}^{\tilde{R}_\beta} n_i, \tag{5.16}$$

which can also be expressed in terms of the family size distribution,

$$\tilde{R}_{\beta'} = \sum_{j=1}^{R} \eta_j, \tag{5.17}$$

where η_j is the size of family j at temperature step β'. In the exact free energy version of PA, families are independent of each other, so the variance of the sum is the sum of variances and $\mathrm{var}(\tilde{R}_{\beta'}) = R\,\mathrm{var}(\eta)$. From the definition of ρ_t, Eq. 5.11,

$$\rho_t(\beta) - 1 = \lim_{R\to\infty} \frac{1}{R} \mathrm{var}(\tilde{R}_\beta) \tag{5.18}$$

Thus, in the exact free energy version of PA, there are relatively large fluctuations in the population size that scale as $\sqrt{\rho_t R}$ in contrast to the version we implement in the simulations, where the population size is nudged toward R at every step and population fluctuations scale as $\sqrt{aR}$ with $a \lesssim 1$.

We can also derive an expression for ρ_f that is related to population size fluctuations within the exact free energy version of PA. Starting from the factor $Q(\beta, \beta')$ [see Eq. 4.2] we have,

$$\begin{aligned} Q(\beta,\beta') &= \frac{1}{\tilde{R}_\beta} \sum_{i=1}^{\tilde{R}_\beta} e^{-(\beta'-\beta)E_i} \\ &= \frac{1}{\tilde{R}_\beta} \left(\sum_{i=1}^{\tilde{R}_\beta} \tau_i \right) e^{-\beta' F' + \beta F}. \end{aligned} \tag{5.19}$$

We also know that $Q(\beta, \beta')$ is an estimator of the ratio of the partition functions of two subsequent temperatures so that,

$$Q(\beta,\beta') = e^{-\beta'\tilde{F}' + \beta\tilde{F}}, \tag{5.20}$$

where $\tilde{F}$ and $\tilde{F}'$ are the free energy estimators at β and β', respectively. Combining the above two relations gives

$$\begin{aligned} e^{-\beta'\Delta\tilde{F}'+\beta\Delta\tilde{F}} &= \frac{1}{\tilde{R}_\beta}\sum_{i=1}^{\tilde{R}_\beta}\tau_i \\ &= \frac{1}{\tilde{R}_\beta}\sum_{i=1}^{\tilde{R}_\beta}[n_i - z(\tau_i)], \end{aligned} \tag{5.21}$$

where $\Delta\tilde{F} = \tilde{F} - F$ is the deviation of the free energy estimator from the exact free energy. If the population is large, this deviation is small so we can expand the exponential of the free energy, and using Eq. 5.16, we obtain

$$1 - (\beta'\Delta\tilde{F}' - \beta\Delta\tilde{F}) = \frac{\tilde{R}_{\beta'}}{\tilde{R}_\beta} - \frac{\sum_{i=1}^{\tilde{R}_\beta} z(\tau_i)}{\tilde{R}_\beta}. \tag{5.22}$$

The population size at temperature β can be decomposed as $\tilde{R}_\beta = R + \delta\tilde{R}_\beta$, where R is the mean population and $\delta\tilde{R}_\beta$ is the deviation from the mean at temperature β. Expanding in $\delta R/R$ yields,

$$\begin{aligned} 1 - (\beta'\Delta\tilde{F}' - \beta\Delta\tilde{F}) = &\left(1 + \frac{\delta\tilde{R}_{\beta'}}{R}\right)\left(1 - \frac{\delta\tilde{R}_\beta}{R}\right) \\ &- \frac{1}{R}\sum_{i=1}^{\tilde{R}_\beta} z(\tau_i)\left(1 - \frac{\delta\tilde{R}_\beta}{R}\right). \end{aligned} \tag{5.23}$$

From this point onwards, $\tilde{R}_j$ will denote the population at annealing step j with inverse temperature β_j, and τ_i^j will denote the weight of configuration i during the resampling step from β_{j+1} to β_j. Disregarding all $(\delta R/R)^2$ terms, summing Eq. 5.23 over all steps in the annealing schedule, and taking the variance of the result yields,

$$\mathrm{var}\left(\beta_k\Delta\tilde{F}_k\right) = \mathrm{var}\left(\frac{\delta\tilde{R}_k}{R} - \frac{1}{R}\sum_{j=N_{T-1}}^{k+1}\sum_{i=1}^{\tilde{R}_{j+1}} z(\tau_i^j)\right). \tag{5.24}$$

From the definition of ρ_f, Eq. 5.9, we have,

$$\rho_f(\beta_k) = \lim_{R\to\infty} \frac{1}{R} \mathrm{var} \left(\delta\tilde{R}_k - \sum_{j=N_{T-1}}^{k+1} \sum_{i=1}^{\tilde{R}_{j+1}} z(\tau_i^j) \right). \tag{5.25}$$

Expanding the variance and using Eq. 5.18 yields a relation between ρ_f and ρ_t,

$$\begin{aligned} \rho_f(\beta_k) = \rho_t(\beta_k) - 1 \\ - \lim_{R\to\infty} \frac{1}{R} \left[2 \ \mathrm{cov} \left(\delta\tilde{R}_k, \sum_{j=N_{T-1}}^{k+1} \sum_{i=1}^{\tilde{R}_{j+1}} z(\tau_i^j) \right) \right. \\ \left. - \mathrm{var} \left(\sum_{j=N_{T-1}}^{k+1} \sum_{i=1}^{\tilde{R}_{j+1}} z(\tau_i^j) \right) \right]. \end{aligned} \tag{5.26}$$

In Appendix A.1 we argue that the quantity in square brackets is greater than zero yielding the desired inequality, $\rho_t - 1 > \rho_f$.

There are two caveats concerning this inequality. First, the argument in Appendix A.1 establishing the positivity of the term in square brackets is not rigorous. More importantly, the result applies to a version of PA that is normalized by the exact free energy and has large fluctuations in population size. We conjecture that an "equivalence of ensembles" result holds for the implemented and exact free energy version of PA so that both ρ_f and ρ_t are the same for both algorithms but this question deserves further study.

We will see in the next two sections and in Appendix A.2 that the inequality between ρ_f and $\rho_t - 1$ can be extended to an approximate equality, provided that the culling fraction is small at each step. This approximate equality and, by extension, the inequality are supported by numerical results shown in Sec. 5.4.1.

5.2.3 Temperature step size, culling fraction and energy variance

As we shall see in Sec. 5.3.2, a natural way to choose the β–schedule for population annealing is to cull a fixed fraction of the population at each resampling step. In this

section we derive a relation between the culling fraction, the variance of the energy distribution, and the size of the temperature step. To derive this relation, note that the expected number of copies of each configuration is τ_i and the actual number of copies is $\lfloor \tau_i \rfloor$ with probability $\lceil \tau_i \rceil - \tau_i$ or $\lceil \tau_i \rceil$ otherwise. Thus a configuration can be eliminated only if $\tau < 1$, and the expected number of eliminated configurations in a resampling step is

$$\epsilon R = \sum_{\tau_i < 1} (1 - \tau_i), \tag{5.27}$$

where ϵ is the (expected) culling fraction. Let $\langle E \rangle$ and σ_E^2, respectively, be the thermal average energy and variance of the energy. Consider a resampling step from β to β' with $\Delta\beta = (\beta' - \beta)$. In the regime $\Delta\beta\sigma_E \ll 1$ we can expand the definition of τ_i, Eq. 4.1, to leading order in $\Delta\beta$ to obtain,

$$\tau_i = 1 - \Delta\beta(E_i - \langle E \rangle), \tag{5.28}$$

meaning that τ is approximately a Gaussian random variable with mean one and standard deviation $\Delta\beta\sigma_E$. Within this Gaussian approximation and for large R, the sum defining ϵ in Eq. 5.27 can be replaced by an integral,

$$\begin{aligned} \epsilon &\approx \int_{-\infty}^{1} (1 - \tau) \mathcal{N}(\tau; 1, \Delta\beta\sigma_E) \, \mathrm{d}\tau \\ &= \frac{\Delta\beta\sigma_E}{\sqrt{2\pi}}, \end{aligned} \tag{5.29}$$

where $\mathcal{N}(x; \mu, \sigma)$ is the pdf of the normal distribution with mean μ and standard deviation σ. If we want to eliminate a fixed fraction of the population ϵ, then the β–schedule must be chosen such that

$$\Delta\beta \approx \frac{\epsilon\sqrt{2\pi}}{\sigma_E}. \tag{5.30}$$

5.2.4 Growth of ρ_f in the MCMC-equilibrated regime

In the spin glass phase where the MCMC procedure alone is unable to equilibrate the system, the growth of ρ_f depends, in a complicated way, on the structure and temperature-dependence of the free energy landscape. However, at high temperatures where the MCMC procedure is able to fully decorrelate replicas, which we refer to as the MCMC-equilibrated regime, we can show that ρ_f is simply proportional to the number of annealing steps times the culling fraction. To understand the behavior of ρ_f in the MCMC-equilibrated regime, note that ρ_f is defined in terms of the variance of βF, [see Eq. 5.9], and take the variance of both sides of Eq. 4.8,

$$\begin{aligned} \mathrm{var}[\beta_k \tilde{F}(\beta_k)] &= \mathrm{var}\left[\sum_{l=k+1}^{N_T-1} \ln Q(\beta_l, \beta_{l-1})\right] \\ &\approx \sum_{l=k+1}^{N_T-1} \mathrm{var}\left[\ln Q(\beta_l, \beta_{l-1})\right]. \end{aligned} \tag{5.31}$$

The second approximate equality becomes exact when the population is equilibrated by the MCMC procedure during each annealing step. Using the definition of Q, Eq. 4.2, and assuming the variation of $\tilde{R}$ is negligible, we expand $\mathrm{var}(\ln Q(\beta_l, \beta_{l-1}))$ to leading order in $(\Delta\beta_l)\sigma_E(\beta_l)$,

$$\mathrm{var}\left[\ln Q(\beta_l, \beta_{l-1})\right] \approx \frac{1}{R}(\Delta\beta_l)^2 \sigma_E(\beta_l)^2, \tag{5.32}$$

where σ_E^2 is the variance of the energy distribution. Plugging Eq. 5.32 into Eq. 5.31 yields,

$$\mathrm{var}[\beta_k \tilde{F}(\beta_k)] = \sum_{l=k+1}^{N_T-1} \frac{1}{R}(\Delta\beta_l)^2 \sigma_E(\beta_l)^2. \tag{5.33}$$

From the relation between the size of the temperature step and the culling fraction, Eq. 5.29, we find,

$$R \operatorname{var}[\beta_k \tilde{F}(\beta_k)] = \sum_{l=k+1}^{N_T-1} 2\pi\epsilon(l)^2, \tag{5.34}$$

where $\epsilon(l)$ is the culling fraction at the l^{th} annealing step. Thus, for fixed culling fraction, ϵ, we find that ρ_f grows linearly in the number k of annealing steps,

$$\rho_f = 2\pi\epsilon^2 k. \tag{5.35}$$

This relation is valid if the culling fraction is small and enough MCMC sweeps are carried out in each annealing step that the replicas remain independent.

More generally, $\rho_f \geq 2\pi\epsilon^2 k$ and the inequality holds if the MCMC procedure is not able to keep the replicas fully decorrelated.

5.2.5 Growth of ρ_t in the MCMC-equilibrated regime

Similarly to the case of ρ_f, at high temperatures where the MCMC procedure is able to fully decorrelate the energy of the replicas at every annealing step, the growth of $\rho_t - 1$ is proportional to the number of annealing steps times the culling fraction. To derive this relation we note that ρ_t is equal to the variance of the family size distribution. In the MCMC-equilibrated regime, the size of a given family is described by a birth and death process. In an approximation where the annealing step k is taken to be a continuous "time" variable, the family size distribution, $P_\eta(k)$, is described by the Master Equation,

$$\dot{P}_\eta(k) = \epsilon \left[(\eta - 1) P_{\eta-1}(k) + (\eta + 1) P_{\eta+1}(k) - 2\eta P_\eta(k) \right], \tag{5.36}$$

where η is the family size and ϵ is the culling fraction. This is the classic birth–death process (see, for example,Ref. [68]) whose solution is

$$P_\eta(k) = \left(\frac{1}{1+\epsilon k} \right)^2 \left(\frac{\epsilon k}{1+\epsilon k} \right)^{\eta-1} \text{ for } \eta \geq 1, \tag{5.37}$$

and

$$P_0(k) = \left(\frac{\epsilon k}{1 + \epsilon k} \right). \tag{5.38}$$

From this distribution it is easily seen that

$$\rho_t - 1 = \mathrm{var}(\eta) = 2\epsilon k. \tag{5.39}$$

This equation holds in the MCMC-equilibrated regime where the number sweeps in each annealing step is greater than or comparable to the integrated autocorrelation time of the energy, so that the energy of every replica is independent of its family designation.

Comparing Eqs. 5.39 and 5.35, we see that in the MCMC-equilibrated regime and for small culling fraction,

$$\rho_f = \rho_t - 1 - (2\epsilon - 2\pi\epsilon^2)k. \tag{5.40}$$

To a high level of accuracy, this relation also holds outside of the MCMC-equilibrated regime. This is shown numerically in Sec. 5.4.1, and supported analytically in Appendix A.2.

5.3 Optimization of Population Annealing

In this work we focus on three general improvements to the population annealing algorithm: a method to choose the population size for each individual bond configuration, a way to choose the β–schedule to reduce statistical and systematic errors, and an ad hoc sweep schedule that improves equilibration.

5.3.1 Hardness-dependent population size

Previous authors have shown that the computational work required to equilibrate a specific bond configuration is approximately lognormally distributed for both population annealing and parallel tempering [12, 59, 110, 115]. In population annealing, the work required to equilibrate a bond configuration is proportional to ρ_f, and so our first optimization takes advantage of the relation between ρ_f and ρ_t in order to tailor the population size necessary for each bond configuration, so we do not use too large a population on an easy system and spend resources inefficiently. In order to optimize the population for each bond configuration, it is usually necessary to do several simulations. An initial simulation is done with a small population, R_0. From this simulation, we obtain an estimate of ρ_t, called $\rho_t(R_0)$. If $R_0 > 100\rho_t(R_0)$ then we assume that $\rho_t(R_0) \approx \rho_t$ and the simulation was well-equilibrated. Otherwise we do another simulation with population size, $R_1 = 150\rho_t(R_0)$, which yields $\rho_t(R_1)$. This procedure is continued until the bond configuration is equilibrated according to the criterion of $R > 100\rho_t$. If R_0 is chosen wisely, this method converges quickly and uses far fewer resources than choosing a single R adequate for all bond configurations. The values of R_0 used in our simulations are given in Table 5.1.

5.3.2 Optimal annealing schedule

In the MCMC-equilibrated regime it is possible to derive an optimal annealing schedule, which is composed of both the β–schedule and sweep schedule. This is done by minimizing ρ_f while keeping the total amount of computational work, W, fixed. We define the total work as

$$W = \sum_{l=k+1}^{N_T-1} S(\beta_l) \tag{5.41}$$

$$\approx \int_0^\beta \frac{S(\beta)}{\Delta\beta} d\beta, \tag{5.42}$$

where $S(\beta)$ is the sweep schedule, defined as the number of MCMC sweeps carried out per annealing step at inverse temperature β. In the MCMC-equilibrated regime we have an analytic expression for ρ_f from Eq. 5.33, which can also be approximated as an integral,

$$\rho_f = \sum_{l=k+1}^{N_T-1} (\Delta\beta_l)^2 \sigma_E(\beta_l)^2 \tag{5.43}$$

$$\approx \int_0^\beta \Delta\beta \sigma_E(\beta) d\beta. \tag{5.44}$$

We can use the method of Lagrange multipliers to minimize ρ_f while holding W constant by solving,

$$0 = \frac{\delta}{\delta\Delta\beta}(\rho_f + \lambda W), \tag{5.45}$$

which yields,

$$\sigma_E^2 - \lambda \frac{S(\beta)}{\Delta\beta^2} = 0. \tag{5.46}$$

Equivalently,

$$\Delta\beta\sigma_E(\beta) \propto \sqrt{S(\beta)}, \tag{5.47}$$

and using Eq. 5.30,

$$\epsilon \propto \sqrt{S(\beta)}, \tag{5.48}$$

we see that in the MCMC-equilibrated regime, the optimal number of sweeps depends on the culling fraction at each step. In the case of a fixed culling fraction, a fixed sweep schedule is optimal.

Based on these ideas, in our simulations we employed a β–schedule that holds the culling fraction roughly constant. We observed that the resulting schedule does not depend strongly on disorder realization so we chose a schedule based on a single run and tested that the culling fraction was similar for several disorder configurations of varying difficulty. This schedule was then employed without modification for production runs. Since, according to Eq. (5.30), $\Delta\beta = \epsilon\sqrt{2\pi}/\sigma_E$, the resulting schedule has many annealing steps at high temperature, where the standard deviation of the energy is large, and few annealing steps at low temperature. A similar uniform ϵ scheme was used in PA simulations of hard spheres [29]. Since the variance of the energy scales linearly with the system size (except near the critical point), the same β–schedule can be used for many system sizes, though for larger sizes the uniform culling fraction will increase. If the range of sizes studied is too large, interpolating temperatures can be added to the schedule to reduce the culling fraction. Although it is not theoretically well-justified in the critical or the low temperature regimes, we continued to use a uniform culling fraction to determine the β–schedule over the entire temperature range of our simulations.

In accordance with Eq. 5.48, we chose a sweep schedule that was fixed in the high temperature regime. We found that focusing most of the computational work in the glass transition regime minimized ρ_t, and that in the glassy phase it was favorable to rely on PA resampling and do very little MCMC work. Our ad hoc sweep schedule had three Monte Carlo sweeps performed for $\beta < 0.5$, 22 for $0.5 \leq \beta \leq 2.5$, and a single sweep for $\beta > 2.5$. Our annealing schedule is not directly comparable with the annealing schedules studied in Ref. [16]. While our β-schedules is the same as their "std(E)" schedule, the computationally meaningful density of sweeps per unit β is not the same as for any of their annealing schedules because of our non-uniform sweep schedule. A deeper understanding of the optimal annealing schedule over the whole range of temperatures remains an open problem.

	L=6	L=8	L=10
Culling fraction, ϵ	0.103	0.168	0.272
Sweeps per replica	1005	1005	1009
Temperature steps, N_T	95	95	99
Initial population, R_0	2×10^3	2×10^3	2×10^4
Maximum population	2.64×10^5	1.5×10^7	1.5×10^7
Disorder samples	2×10^4	2×10^4	5×10^3
Unequilibrated samples	0	4	145

Table 5.1. Parameters of the simulations for each of the three system sizes.

5.4 Results

This section contains two types of numerical results. Sections 5.4.1-5.4.3 address the behavior of PA, validating theoretical predictions and testing the optimization ideas described in previous sections. Section 5.4.4 reports measurements from large-scale simulations of several observables for the 3D Edwards-Anderson spin glass in order to compare with previous work and improve the state-of-the-art. The parameters used in this work are shown in Table 5.1.

5.4.1 Relationship of ρ_f and ρ_t

In Sec. 5.2.2 we showed that $\rho_t - 1 \geq \rho_f$ and, in Eq. 5.40, gave an approximate relation between these two quantities. Here we test these relationships. Our proposal to optimize population size for each disorder realization relies on the easily measured ρ_t as a proxy for the more difficult to measure ρ_f, so it is important to determine the relationship between these two quantities.

In order to accurately measure $\rho_f = R\,\mathrm{var}(\beta\tilde{F})$, we ran population annealing 48 times for each configuration, with population sizes chosen such that $R \geq 100\rho_t$. This ensured that each simulation was well-equilibrated and that we could measure $\mathrm{var}(\beta\tilde{F})$ with reasonable accuracy. We calculated ρ_f for 2000 $L = 6$ samples and for 300 $L = 8$ samples. Calculating the error of ρ_f is equivalent to calculating the error of a sample variance. As shown in previous work [110], $\beta\tilde{F}$ taken from a well-

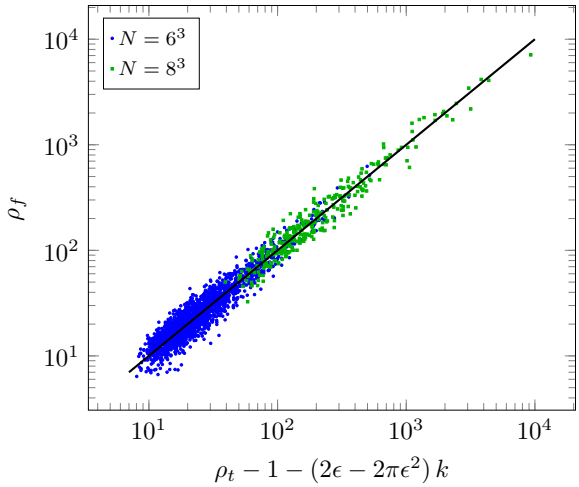

Figure 5.1. Scatter plot of $\rho_t - 1 - (2\epsilon - 2\pi\epsilon^2)k$ vs ρ_f at $\beta = 5$ ($k = 95$) for $L = 6$ and $L = 8$; each data point corresponds to a single bond configuration. The solid line corresponds to $\rho_t - 1 + (2\pi\epsilon^2 - 2\epsilon)k = \rho_f$, see Eq. 5.40.

equilibrated bond configuration is normally distributed, which makes estimation of the error of $\mathrm{var}(\beta\tilde{F})$ particularly easy [113],

$$\mathrm{var}\left[\mathrm{var}(\beta\tilde{F})\right] = \frac{2}{M-1}\mathrm{var}(\beta\tilde{F})^2, \tag{5.49}$$

where M is the number of trials. The corresponding error in ρ_f is

$$\delta\rho_f = \sqrt{\frac{2}{M-1}}\rho_f, \tag{5.50}$$

which for 48 trials gives a relative error, $\delta\rho_f/\rho_f$, of about 21%. Since ρ_t is calculated from a single disorder realization, it is expected that $\delta\rho_t \ll \delta\rho_f$. We find this to be true empirically, with $\delta\rho_f \approx 20\delta\rho_t$ for $L = 6$, and $\delta\rho_f \approx 16\delta\rho_t$ for $L = 8$ simulations.

Figure 5.1 is a scatter plot of $\rho_t - 1 - (2\epsilon - 2\pi\epsilon^2)k$ vs ρ_f at $\beta = 5$, where each point corresponds to one disorder realization and the solid line corresponds to $\rho_f = \rho_t - 1 - (2\epsilon - 2\pi\epsilon^2)k$. The results are consistent with $\rho_f \approx \rho_t - 1 - (2\epsilon - 2\pi\epsilon^2)k$ holding for all disorder realizations.

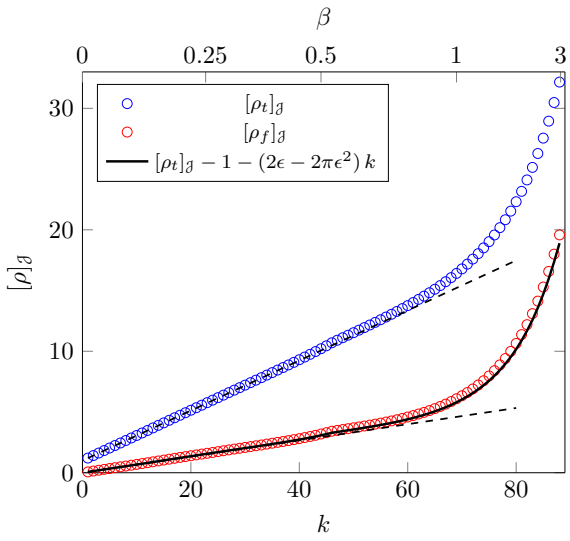

Figure 5.2. $[\rho_f]_\mathcal{J}$ and $[\rho_t]_\mathcal{J}$ of $L = 6$ as a function of annealing step k, with the nonlinear β scale on the upper x-axis. Dashed lines correspond to the theoretically predicted MCMC-equilibrated estimates, Eqs. 5.35 and 5.39. The solid line corresponds to the difference $[\rho_t]_\mathcal{J} - 1 - (2\epsilon - 2\pi\epsilon^2)\, k$, see Eq. 5.40.

In the MCMC-equilibrated regime, $\rho_f = 2\pi\epsilon^2 k$ and $\rho_t = 1 + 2\epsilon k$ as shown in Secs. 5.2.4 and 5.2.5, respectively. Figure 5.2 shows $[\rho_f]_\mathcal{J}$ and $[\rho_t]_\mathcal{J}$ for $L = 6$ with the dashed lines representing the theoretical linear dependence on number of annealing steps k. The solid line represents the estimated value of ρ_f calculated using $[\rho_t]_\mathcal{J} - 1 - (2\epsilon - 2\pi\epsilon^2)\, k$. The estimated value was found to be within 5% of the true value of ρ_f for all annealing steps for both $L = 6$ and $L = 8$ (not shown). Note that the sharp rise in both ρ_f and ρ_t occurs near the critical temperature, $\beta_c = 1.05$.

Figure 5.3 shows the disorder-averaged family size distribution for several temperatures in the MCMC-equilibrated regime for $L = 8$ and confirms the prediction of Sec. 5.2.5 of an exponential family size distribution. The straight lines in the figure are obtained from Eq. 5.37 and show that there is good quantitative agreement except for tail of the distribution. The higher values of η are underrepresented, especially for low values of β, due to the finite size of the population.

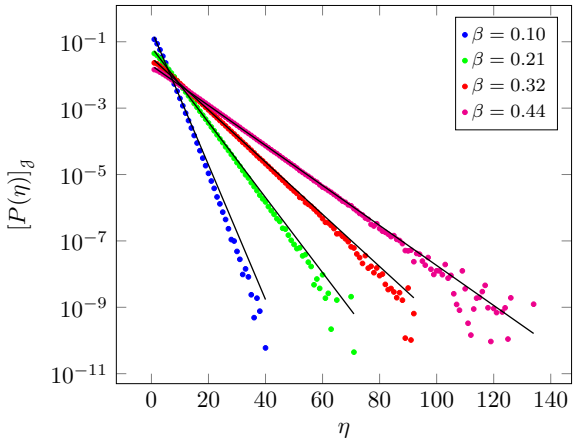

Figure 5.3. The disordered-averaged family size distribution $[P(\eta)]_J$ as a function of family size η, at several temperatures all in the high temperature regime (the β of each distribution increases from left to right). The distributions are exponential and have shape parameters that match the predictions of Eq. 5.37. The value at $\eta = 0$ is not shown in this plot.

5.4.2 Distribution of ρ_t

It is known that the computational hardness of simulating Ising spin glasses has a broad distribution with respect to disorder realizations [2, 3, 59, 110, 115]. In the context of a MCMC algorithm such as parallel tempering the computational hardness is typically measured by the exponential or integrated autocorrelation times. These quantities have been found to be approximately lognormally distributed. For population annealing, we may use ρ_f or ρ_t for the purpose of measuring computational hardness. Figure 5.4 shows histograms of $\log_{10} \rho_t$ for the three system sizes simulated. We found that a log inverse Gaussian distribution is an excellent fit to the ρ_t distributions. The fits are shown as solid lines in the figure. The three-parameter inverse Gaussian distribution is defined by,

$$P(x; \mu, \lambda, l) = \sqrt{\frac{\lambda}{2\pi(x-l)^3}} \exp\left[\frac{-\lambda(x-\mu-l)^2}{2\mu^2(x-l)}\right] \tag{5.51}$$

with $x = \ln \rho_t$. The parameters of the fits are given in Table 5.2. The shift parameter l, which shifts the support of the distribution from $(0, \infty)$ to (l, ∞), is necessary

since ρ_t is bounded away from zero. A rough estimate of shift parameter can be obtained by assuming that the easiest bond realizations are MCMC-equilibrated for all temperatures, resulting in $l = \log(\rho_t^{\min}) \approx \log(1+2\epsilon k_{\max})$. The values for l obtained from this formula are 2.97, 3.50 and 4.00 for sizes $L = 6$, 8 and 10, respectively. These values are in reasonable agreement with the fitted values shown in Table 5.2. The log inverse Gaussian also works well to fit the ρ_f distribution with a shift parameter predicted by $\rho_f^{\min} = 2\pi\epsilon^2 k$.

We can also compare the disorder averaged value of ρ_t with the predictions from the fit. The mean of $\rho_t = e^x$ is given by,

$$[\rho_t]_{\mathcal{J}} = \exp\left[l + \frac{\lambda}{\mu}\left(1 - \sqrt{1 - 2\mu^2/\lambda}\right)\right], \tag{5.52}$$

for $\lambda/2\mu^2 > 1$. The tail of the inverse Gaussian is exponential, so if $\lambda/2\mu^2 \leq 1$, the mean of e^x is infinite. The fitted values of $[\rho_t]_{\mathcal{J}}$ obtained from this equation are shown in Table 5.2. The values computed directly from the data are $[\rho_t]_{\mathcal{J}} = 46$, 385 and 12700 for $L = 6$, 8 and 10, respectively. The large discrepancy between the $L = 10$ fitted and measured values may be due to several factors. First, the number of samples for $L = 10$ is smaller than for $L = 6$ and 8, so the tail of the ρ_t distribution may not be fully sampled. Second, a significant fraction of $L = 10$ samples were not equilibrated and, for these samples, we have most likely underestimated ρ_t. If the tail of the distribution is properly described by the log inverse Gaussian, the value of $[\rho_t]_{\mathcal{J}}$ obtained from the fit may be more accurate than the average of the ρ_t data from a finite sample size. On the other hand, for $L = 10$, the ratio $\lambda/2\mu^2 = 1.04$, which is quite near the divergence at $\lambda/2\mu^2 = 1$ so results for $[\rho_t]_{\mathcal{J}}$ may be highly sensitive to errors in the fit. The near divergence of $[\rho_t]_{\mathcal{J}}$ for $L = 10$ also suggests that the annealing schedule for this size should have either more temperature steps or more sweeps per step.

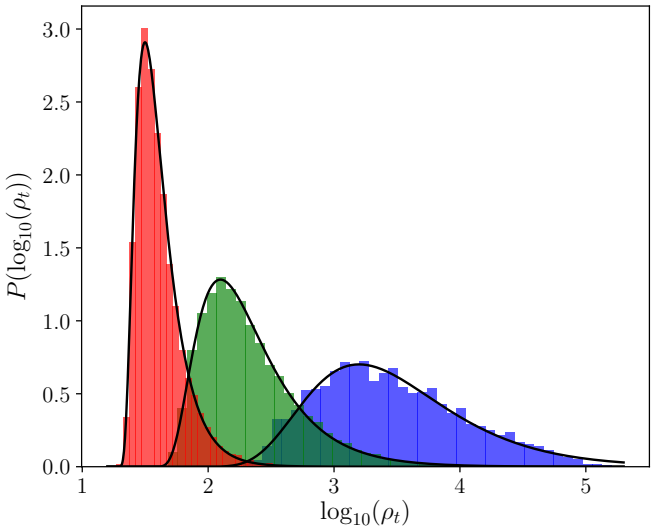

Figure 5.4. The distribution of $\log_{10}\rho_t$ for system sizes $L = 6$ (left), 8 (middle) and 10 (right). The solid lines are inverse Gaussian fits with the parameters given in Table 5.2.

	L=6	L=8	L=10
μ	0.83(1)	1.97(2)	4.53(13)
λ	3.25(16)	9.28(35)	42.7(40)
l	2.88(1)	3.39(2)	3.50(13)
$[\rho_t]_\mathcal{J}$	46.4	491	66500

Table 5.2. Fits of the ρ_t distribution to a log inverse Gaussian distribution defined in Eq. 5.51.

As seen in Table 5.2, the computational effort required to reach equilibrium scales up rapidly with system size and is broadly distributed. How does this effort translate into wall clock time on a modern CPU? For $L = 10$ the algorithm's run time on a single CPU is approximately 0.03 sec/replica. The typical value of ρ_t for $L = 10$, defined by $\exp([\log \rho_t]_{\mathcal{J}})$, is approximately 3000 and the equilibration criterion is that $R \geq 100\rho_t$ thus the typical running time is approximately 2.5 hours. This relatively benign number is, however, misleading because of the exponential tail of the $\log \rho_t$ distribution. If one instead takes the average hardness, $[\rho_t]_{\mathcal{J}} \approx 66500$, predicted from the inverse Gaussian fit (see Table 5.2), then the average running time to equilibrate every disorder realization in a very large $L = 10$ sample would be approximately 55 hours per disorder realization. This number exceeds the computing time expended on our $L = 10$ simulations since we did not equilibrate all disorder realizations. Of course, the algorithm can be efficiently parallelized so the wall clock time per replica can be made much smaller than these numbers.

5.4.3 Optimized vs. unoptimized annealing schedule

In this section we compare the performance of an optimized and unoptimized annealing schedule used in PA. Our optimized annealing schedule has a β–schedule that keeps the culling fraction ϵ nearly constant, and an ad hoc sweep schedule that concentrates sweeps over a range of temperatures around the critical point, as described in Sec. 5.3.2. The unoptimized annealing schedule has constant β steps, $\Delta\beta = 0.05$, with 10 sweeps per step, and is similar to the annealing schedule used in Ref. [110].

The figure of merit that we wish to minimize is size of the systematic errors, which scale as $\mathrm{var}(\beta F)$, times the total computational work, $W = RS$, where R is the population size and S is the total number of Monte Carlo sweeps per replica. Using the result $\rho_f \approx \rho_t$ we have that

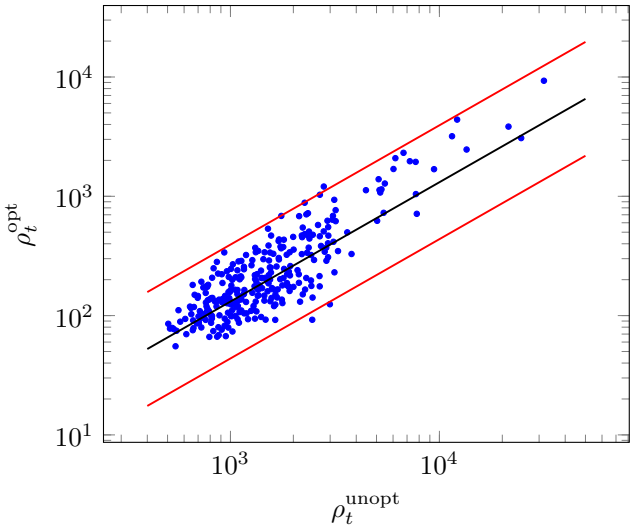

Figure 5.5. Values of ρ_t at $\beta = 5$ for the optimized and unoptimized annealing schedules. Each point corresponds to one of 300 $L = 8$ bond configurations. The horizontal coordinate of each point is the unoptimized value and the vertical coordinate the optimized value of ρ_t. The central (black) line corresponds to an improvement of the optimized relative to the unoptimized annealing schedule by a factor of 7.6, and the upper and lower (red) lines correspond to factors of $7.6/3$ and 7.6×3, respectively.

$$W \operatorname{var}(\beta F) \approx S\rho_t. \tag{5.53}$$

We have used the same number of sweeps in both the optimized and unoptimized algorithms so the comparison of the performance of the algorithms reduces to the comparison of ρ_t.

Figure 5.5 is a scatter plot of the values of ρ_t, the vertical position of each point is the optimized ρ_t^{opt}, and the horizontal position, ρ_t^{unopt}, the unoptimized value. Each point represents one of 300 disorder realizations for system size $N = 8^3$ at $\beta = 5$. The plot shows that the unoptimized algorithm is, on average, less efficient by a factor of 7.6. This means that, on average, the unoptimized algorithm requires a population 7.6 times larger to achieve the same quality of results as the optimized algorithm.

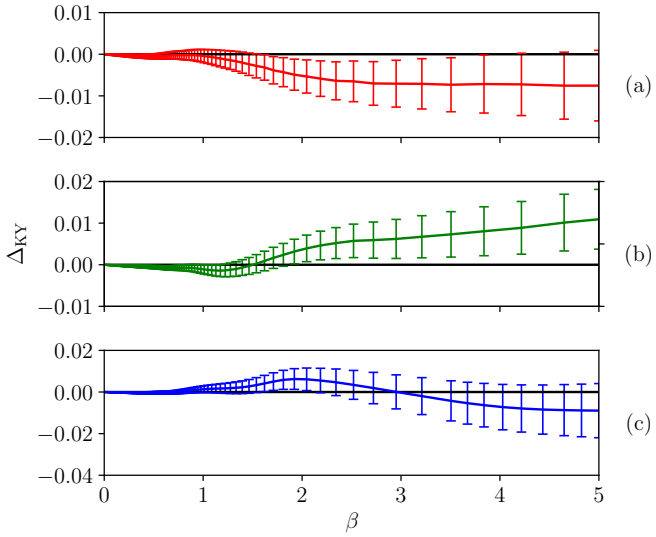

Figure 5.6. Δ_{KY} as a function of inverse temperature β for $L = 6$ (a), $L = 8$ (b), and $L = 10$ (c).

	L=6	L=8	L=10
I(0.2)	0.0188(5)	0.0185(5)	0.0185(10)
Δ_{KY}	−0.0075(85)	0.010(7)	−0.009(13)
$[E_0/N]_{\mathcal{J}}$	−1.6891(4)	−1.6951(2)	−1.6976(3)

Table 5.3. The integrated overlap I(0.2), Katzgraber-Young equilibration measure Δ_{KY}, and the bond-averaged ground state energy per spin $[E_0/N]_{\mathcal{J}}$, all measured at $\beta = 5$.

L	$[\log_{10} \tilde{g}_0]_{\mathcal{J}}$	$\log_{10} 2 - \beta[(\tilde{E}_0 - \tilde{F})]_{\mathcal{J}}/\ln(10)$
6	−0.7577(23)	−0.7549(23)
8	−1.6957(34)	−1.6900(34)
10	−3.2251(93)	−3.2104(92)

Table 5.4. Comparison of the disorder average of the logarithm of the fraction in the ground state, $[\log_{10}(\tilde{g}_0)]_{\mathcal{J}}$ at $\beta = 5$ to the indirect measure, $[\log_{10}(\bar{g}_0)]_{\mathcal{J}}$ based on the Boltzmann factor, see Eq. 5.7.

5.4.4 $L = 6, 8, 10$ results

To test the optimized algorithm and obtain state-of-the-art results for observables described in Sec. 5.1.1, we ran large-scale simulations of the 3D EA spin glass for three system sizes, with parameters provided in Table 5.1. As seen in this table, the equilibration standard, $R \geq 100\,\rho_t$, was met by nearly all configurations for $L = 6$ and $L = 8$, but approximately 3% of the configurations for $L = 10$ remained unequilibrated when the maximum population was restricted to $R = 1.5 \times 10^7$. The equilibration standard used was higher than previous papers employing PA to study spin glasses, and some systems that we rejected as unequilibrated would have been accepted previously.

It is worth emphasizing that the adaptive population scheme allowed us to sample more bond configurations than most previous studies, while ensuring that nearly all configurations were well-equilibrated. As a result, errors associated with a finite number of bond configurations are especially low for $L = 6$ and 8, where we used 2×10^4 samples. The statistical errors reported in Table 5.3 and 5.4 are obtained from the standard deviation of the observable with respect to disorder realization and do not include errors associated with individual disorder realizations. Despite the large number of disorder realizations, the error due to the finite sample size is substantially larger than the contribution from systematic and statistical errors of each disorder realization, as shown in Ref. [110].

Overall, the results in Table 5.3 are consistent with those found in previous works [110, 111]. The average ground state energy, $[E_0]_{\mathcal{J}}$, is within error bars of previous measurements and, as shown in Fig. 5.6, Δ_{KY} is close to zero for all β, which is consistent with a well-equilibrated set of samples. Our values of I(0.2) are all within two standard deviations of those found in Ref. [110], however, it is noteworthy that our values are consistently lower. Despite the slight difference in value, the trend that I(0.2) remains constant over several system sizes is evident.

Table 5.4 shows the disorder average of the logarithm of the directly measured fraction in the ground state $\tilde{g}_0$ and the indirectly measured quantity $\bar{g}_0$ calculated using the Gibbs distribution, Eq. 5.7. Although the two methods yield values that are within error bars, the computed $\bar{g}_0$ appears to be consistently larger than the measured g_0. To leading order, the free energy estimator is systematically larger than the actual free energy by $F = \tilde{F} + \rho_f/2\beta R$ [110], so it is expected that $\bar{g}_0$ would be systematically larger than g_0. It should be noted that by definition g_0 cannot be zero because the ground state energy is here defined as the lowest energy replica found, even if this is not the true ground state.

5.5 Conclusions

The work presented in this chapter makes several contributions to understanding and improving the population annealing algorithm, especially as applied to spin glasses: we have studied the behavior of two important measures of equilibration for population annealing, optimized the algorithm in several ways, and obtained state-of-the-art results for several important spin glass observables. Our results help put population annealing on a firmer footing as an effective tool for highly parallelized simulations of disordered systems such as spin glasses that have rough free energy landscapes. While this work focuses on the three-dimensional Edwards-Anderson model, many of the theoretical results and optimization methods are applicable to population annealing simulations of a much broader class of systems.

The two equilibration measures, ρ_t and ρ_f, set the population size needed to control statistical and systematic errors, respectively. The equilibrium population size ρ_f is based on the variance of the free energy and is the more fundamental measure of systematic errors but more difficult to accurately measure in a single simulation. We have demonstrated that $\rho_f \leq \rho_t - 1$, and confirmed that these two quantities are close to being equal when both are large. We have also shown that in the MCMC-

equilibrated regime, ρ_f and ρ_t each grow linearly in the number of annealing steps, a fact that can be used to design optimal annealing schedules. Finally, we have shown that for the 3D EA spin glass, the distribution of $\log \rho_t$ values is accurately described by an inverse Gaussian distribution.

We have shown that there are a number of simple modifications which improve the efficiency of population annealing. We have also shown that a β–schedule that is chosen by fixing the culling fraction is optimal in the MCMC-equilibrated regime. Lastly, we have shown that the sweep schedule can be improved by increasing the number sweeps in the critical region. Annealing schedule optimizations alone have accounted for nearly an order of magnitude improvement over previous versions of the algorithm. In the context of spin glasses, where there is a broad distribution of computational hardness, tailoring the population size to the difficulty of the disorder realization yields the single largest improvement in efficiency. In addition, the sample-dependent population size results in simulations with both less work and higher overall accuracy.

We still lack a theoretical understanding of the best sweep schedule. Intuitively, we would like the Markov chain Monte Carlo subroutine to fully equilibrate each replica in the high temperature regime. However, in the low temperature regime, this is not feasible and the goal is to equilibrate replicas only within their free energy minima, leaving re-sampling to properly redistribute replicas between distinct free energy minima. We have not yet found a principled way to achieve this goal.

CHAPTER 6

CONFIGURATIONAL GLASS SIMULATIONS

6.1 Introduction

In this chapter, we apply microcanonical (NVT) population annealing Monte Carlo and thermodynamic integration to a binary hard sphere fluid in the glassy regime. We present a new hybrid version of population annealing that is similar in many respects to that first shown in Ref. [95], and a two new ways to measure the configurational entropy of a glass. The first is a type of Frenkel-Ladd thermodynamic integration and is called the constraining shell integration method. The second is an entirely new way of integrating the configurational entropy with population annealing and is called replica thermodynamic integration. We use these methods to probe beyond the dynamic glass transition in order to make estimates of the values of ϕ_{K} and ϕ_{rcp} and to determine whether or not a thermodynamic transition occurs.

The chapter is organized as follows. We begin by briefly reviewing the binary hard sphere model and observables of interest in Sec. 6.2. We then describe the NVT version of the population annealing algorithm and introduce two new thermodynamic integration techniques to calculate the vibrational entropy of a glass state in Sec. 6.3. We present the results from large-scale simulations in Sec. 6.4 and the chapter closes with a discussion in Sec. 6.5. The work presented in this chapter has been submitted to Physical Review E and is available on arXiv, see Ref. [5].

6.2 Model and observables

The system we study is a binary hard sphere fluid with a 50:50 mixture of particles with radius ratio 1.4:1 [29, 53, 85]. This system is known to be a good glass former because although its high density equilibrium state in the thermodynamic limit is two monodisperse crystals separated by a domain wall, this state is inaccessible in simulations starting from a random mixture. One of the primary observables of interest in a simple fluid is the dimensionless pressure, Z, defined as

$$Z = \frac{PV}{Nk_BT}, \tag{6.1}$$

where P is the pressure, V is the volume, N is the total number of particles, k_B is Boltzmann's constant, and T is the temperature. It is common to use the packing fraction, ϕ, as a control parameter where

$$\phi = N\frac{4\pi r^3}{3V}, \tag{6.2}$$

and $r^3 = (r_0^3 + r_1^3)/2$ is the average of the cubed radii of the two species. Because there is no potential energy in this system, the temperature only sets the average kinetic energy of the particles and the remaining physics depends on the dimensionless ratio Z. Thus, without loss of generality, we set $k_B = T = 1$. We work in the NVT ensemble where N and ϕ are set and $Z(\phi)$ is measured. Although free energy is the thermodynamic potential for NVT, due to the triviality of the energy, all of the equilibrium physics is contained in the entropy as a function of N and ϕ.

The binary fluid equation of state is well-approximated throughout the fluid phase by the phenomenological Boublík-Mansoori-Carnahan-Starling-Leland (BMCSL) equation of state [24, 73]

$$Z_{\mathrm{BMCSL}} = \frac{(1 + \phi + \phi^2) - 3\phi(y_1 + y_2\phi) - y_3\phi^3}{(1-\phi)^3}, \tag{6.3}$$

where y_i are constants that depend on the polydispersity. For a 50:50 mixture of 1.4:1 size particles, $y_1 = 0.0513$, $y_2 = 0.0237$, and $y_3 = 0.9251$. This equation of state is very accurate when compared to numerical data [29], but it must break down at high density, as is clear since it remains finite until $\phi = 1$.

For glassy systems, we are ultimately interested in the configurational entropy, $S_{\rm c}$. There are several ways of estimating $S_{\rm c}$ directly, including counting inherent structures [47] and using the Franz-Parisi potential [43]. However, the most common procedure, used here, is to measure the total entropy and vibrational entropy and then use the relation,

$$S = S_{\rm vib} + S_{\rm c}, \tag{6.4}$$

where S and $S_{\rm vib}$ are the total and vibrational entropies, respectively.

To better understand the definitions and relationships of these entropies, we start with the assumption that in the glassy regime configuration space is broken into ergodically disconnected regions, each of which corresponds to a different glass state. In the NVT ensemble each glass state ν appears with probability,

$$w_\nu(\phi) = \frac{\Omega_\nu}{\sum_{\tilde{\nu} \in c(\phi)} \Omega_{\tilde{\nu}}}, \tag{6.5}$$

where the sum is over the set of all glass states, $c(\phi)$, at packing fraction ϕ, and Ω_ν is the configuration space volume of glass state ν. The entropy of glass state ν is given by the standard definition [19],

$$S_\nu(\phi) = \frac{1}{N} \log \Omega_\nu. \tag{6.6}$$

Throughout this chapter all entropies are defined per particle. We define $S_{\rm vib}$ as the average over the entropies of the glass states,

$$S_{\text{vib}}(\phi) = \sum_{\nu \in c(\phi)} w_\nu(\phi) S_\nu(\phi). \tag{6.7}$$

The total entropy is given by the standard definition,

$$\Omega(\phi) = \sum_{\nu \in c(\phi)} \Omega_\nu(\phi), \tag{6.8}$$

$$S(\phi) = \frac{1}{N} \log \Omega(\phi). \tag{6.9}$$

Using these definitions and Eq. 6.4 yields an equation for the configurational entropy,

$$S_{\text{c}} = -\frac{1}{N} \sum_{\nu \in c(\phi)} w_\nu(\phi) \log w_\nu(\phi), \tag{6.10}$$

which is similar in form to the Shannon entropy. If we assume that there is a finite number of glass states, $\mathcal{N}_g$, each of which has the same statistical weight, $w_\nu = 1/\mathcal{N}_g$, then S_{c} reduces to the standard definition given in Eq. 2.4,

$$S_{\text{c}} = \frac{1}{N} \log \mathcal{N}_g. \tag{6.11}$$

The pressure and the total entropy obey the standard thermodynamic relation which, in terms of packing fraction, is given by

$$Z = -\phi \frac{\partial S}{\partial \phi}. \tag{6.12}$$

Using this relation, the dimensionless pressure can be integrated with respect to the packing fraction in order to obtain the entropy, see App. B.1 for details. Numerically measuring the vibrational entropy can also be done using thermodynamic integration but is non-trivial and is discussed below in Sec. 6.3.3.

6.3 Computational methods

6.3.1 Microcanonical population annealing

We simulate the binary hard sphere mixture at high density using an NVT ensemble version of population annealing Monte Carlo, first described in Ref. [29]. Population annealing (PA) is a sequential Monte Carlo method [4, 72, 95, 110] similar to nested sampling [74, 93, 100] that is used to simulate systems with rough free energy landscapes. The general idea of population annealing is to initialize a large ensemble of independent simulations in an easy-to-equilibrate region of parameter space and to anneal towards a difficult-to-equilibrate region. After each annealing step, the population is resampled so as to keep the distribution in equilibrium. Initially the population can be kept in equilibrium via conventional Monte Carlo schemes, however, eventually the simulation enters a regime where the system is unable to equilibrate dynamically and resampling becomes necessary to keep the population in equilibrium.

In this work, we deal with hard spheres in the NVT ensemble. All allowed configurations of hard spheres have the same energy and the roughness of the free energy landscape in the glassy regime is entirely due to a rough entropy landscape. In the NVT version of PA, each replica in the population is independently initialized as a gas of particles at low packing fraction. An equilibrating procedure such as molecular dynamics or Markov chain Monte Carlo is then applied to each member of the population. Here we use event chain Monte Carlo (ECMC) [18, 79]. This step is performed to equilibrate and decorrelate the population, however, in the glass regime, it only manages to move particles within local cages. After running Monte Carlo, the population is annealed by increasing the packing fraction, ϕ, following an annealing schedule, $\{\phi_0, \phi_1, \ldots, \phi_f\}$, where ϕ_0 is in the low density fluid phase and ϕ_f is the highest packing fraction simulated in the glassy regime. The physical process of annealing corresponds to decreasing the box volume but, in simulations, it is compu-

tationally simpler to increase the sphere radii while keeping the volume fixed. As we will see below, these two annealing processes result in different changes in entropy so a correction must be made when performing thermodynamic integration.

After increasing the sphere radii, a fraction of the population's configurations have overlaps and are illegal at the new density. Replicas with illegal configurations are erased or "culled" and are replaced by randomly resampling the remaining legal replicas with equal weight. The fraction of the population that is culled, ϵ, is called the "culling fraction" and is an important quantity for setting the annealing schedule and integrating the entropy. After culling and resampling, this process is then repeated at the new packing fraction and annealing continues until ϕ_f is reached.

In this work we use a "hybrid" resampling method that is similar to that in Ref. [95] in order to reduce the statistical errors associated with the resampling process. In the PA scheme of Ref. [29], described above, the new population is resampled from the final population at the end of the annealing step. The hybrid method decreases the statistical errors associated with resampling by increasing the frequency of sampling during a single annealing step. Instead of resampling the population once at the end of the equilibration process, the population is sampled several times and additional sweeps are performed before each sampling step. As a result, the entropy of each member of the population is taken into account several different times. This procedure is shown in Fig. 6.1. The scheme used in this work is to perform a total of c_k sweeps of ECMC on each member of the population at annealing step k, where we define the number of sweeps as the number of particle movements divided by the number of particles. The sweeps are broken into an initial burn in of $c_k/2$ sweeps. After each subsequent sweep, a fraction of the population is subsampled and configurations that are legal at the next packing fraction are saved into a reservoir. In this work, the subsampled population consists of $R/(c_k/2) * 1.5$ randomly chosen replicas.

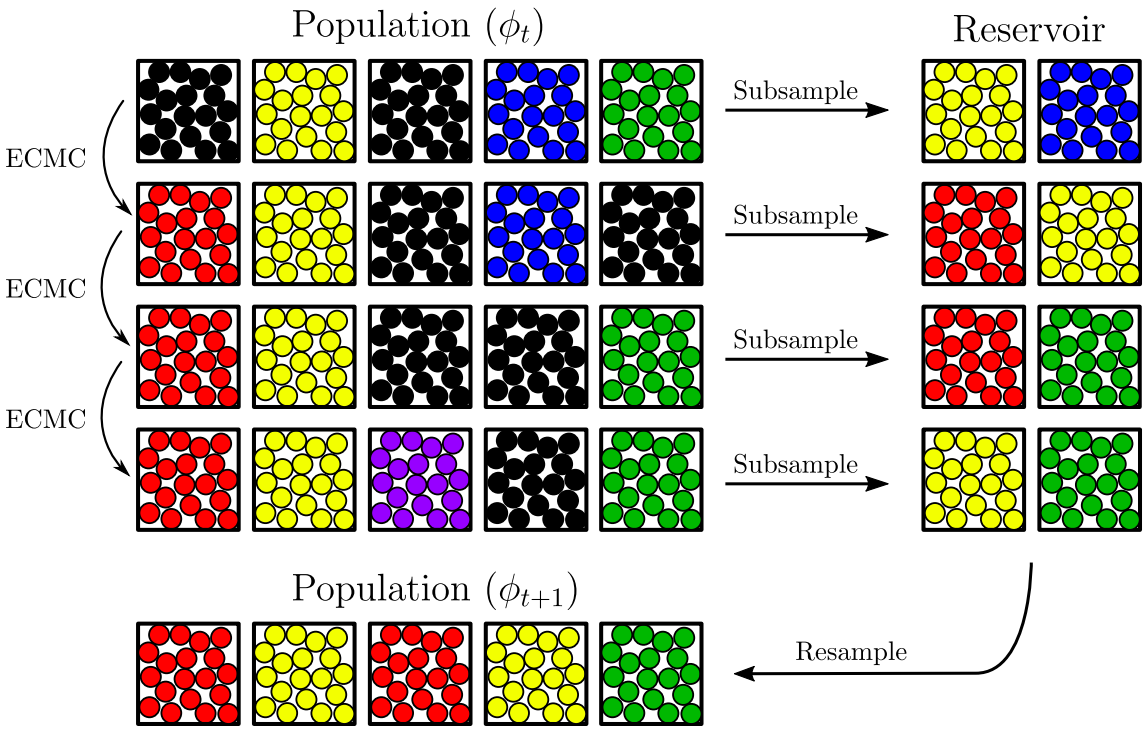

Figure 6.1. Diagram representing one hybrid microcanonical population annealing step. The packing fraction is fixed at ϕ_t and the population is equilibrated using ECMC for several intervals. After each interval, a fraction of the population is subsampled and configurations that are legal at the next packing fraction in the annealing schedule, ϕ_{t+1}, are saved in a reservoir. Illegal configurations (marked black) are not saved. After equilibration is completed, the reservoir is randomly sampled to produce a new population at packing fraction ϕ_{t+1}.

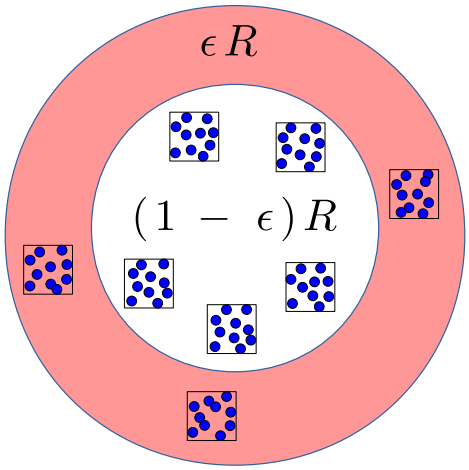

Figure 6.2. The outer circle corresponds to configuration space at the initial packing fraction ϕ_t and the inner circle corresponds to configuration space after annealing to the new packing fraction ϕ_{t+1}. The fraction of configurations that are eliminated after annealing is an estimator for the fraction of configuration space volume that has been eliminated.

Finally, at the end of the annealing schedule, R replicas are chosen at random from the reservoir to represent the population at the beginning of the next annealing step.

The subsample size was chosen with consideration of the culling fraction in order to ensure that the total number of legal configurations placed in the reservoir at the end of the annealing step would be larger than the population size, R. The factor of 1.5 in the subsampling step was added as an additional safety to ensure that the reservoir is always larger than the total population. This factor acts as a tunable parameter that determines the sampling rate of the entire population. If we choose our factor so that the reservoir is nearly exactly R, then the weight of each member of the population is effectively measured once and there will be no benefit in comparison with standard PA. If we choose a large reservoir, many times the size of R, then the weight of each replica will effectively be measured many times resulting in reduced statistical errors, but at the expense of using more computer memory.

Both the hybrid and the standard versions of microcanonical PA give access to the culling fraction, ϵ_i, at annealing step i, which is an estimator of how much con-

figuration space volume contracts after an annealing step. In particular, the volume of configuration space decreases by a factor of $1-\epsilon_i$ each annealing step, as shown in Fig. 6.2, and the corresponding change in entropy is given by

$$\Delta S_i = \frac{1}{N}\log(1-\epsilon_i) - \log(\phi_i/\phi_{i-1}), \tag{6.13}$$

where the ratio of packing fractions corrects for the fact that we keep the system volume fixed during annealing. By summing the changes in entropy over the entire simulation, it is possible to numerically integrate the total entropy,

$$S(\phi_k) = S(\phi_0) + \frac{1}{N}\sum_{i=0}^{k-1}\left[\log(1-\epsilon_i) - \log(\phi_i/\phi_{i-1})\right], \tag{6.14}$$

where $S(\phi_0)$ is the entropy at the initial packing fraction, see App. B.1 for details.

Due to limitations in computational resources, it was necessary to carry out multiple independent runs of PA rather than one run with a very large population. Results from independent simulations can be combined using weighted averaging [29, 110] to reduce both statistical and systematic errors and also to estimate the magnitude of these errors. Given M independent runs of PA, each with population size $R^{(m)}$, the weighted average $\overline{\mathcal{O}}$ of an observable $\mathcal{O}$, such as the pressure, is given by,

$$\overline{\mathcal{O}} = \frac{\sum_{m=1}^{M}\tilde{\mathcal{O}}^{(m)}R^{(m)}\exp[NS^{(m)}]}{\sum_{m=1}^{M}R^{(m)}\exp[NS^{(m)}]}, \tag{6.15}$$

where $\tilde{\mathcal{O}}^{(m)}$ and $S^{(m)}$ are the estimators of the observable and the entropy, respectively, in run m. The weighted average of the entropy itself depends on a summation over the annealing schedule and is given by a different formula,

$$\overline{S} = \frac{1}{N}\log\frac{\sum_{m=1}^{M}R^{(m)}\exp[NS^{(m)}]}{\sum_{m=1}^{M}R^{(m)}}. \tag{6.16}$$

For fixed population size, $R^{(m)} = R$, the weighted average of an observable becomes exact in the limit of infinitely many runs, $M \to \infty$.

6.3.2 Event chain Monte Carlo

The population annealing equilibrating procedure that we use in this work is event chain Monte Carlo (ECMC), which is particularly efficient at sampling 2D and 3D hard sphere configurations [18, 39, 55, 79]. In the version of ECMC used here, a particle is randomly chosen and is translated in a random direction until it collides with another particle. When a collision occurs, the moving particle is stopped and the struck particle is moved in the same direction until it collides with another particle. This process is repeated until the total distance travelled by the particles, called the chain length, is equal to a predetermined length, ℓ. When this distance is reached, the current moving particle is immediately stopped. This process is then repeated by randomly choosing a new starting particle to move in a new direction. We simulate a system with periodic boundary conditions, so it is sufficient to move particles in only the positive x, y, or z directions, which violates detailed balance but preserves global balance [18].

Event chain Monte Carlo provides a way of measuring the dimensionless pressure, as shown in Ref. [79]. Consider a single chain in the x direction. When two particles, j and k, collide, the distance between their centers projected in the x-direction is $x_k - x_j$. The "lifted" distance of an event chain, $x_{\text{final}} - x_{\text{initial}}$ is defined as

$$x_{\text{final}} - x_{\text{initial}} = \ell + \sum_{k,j} (x_k - x_j) \,, \tag{6.17}$$

where the sum takes place over all collisions in a single chain. If this process is repeated, then the dimensionless pressure is given by an average of the lifted distance over all of the chains,

$$Z = \left\langle \frac{x_{\text{final}} - x_{\text{initial}}}{\ell} \right\rangle_{\text{chains}}. \tag{6.18}$$

The total entropy of the fluid is then given by the thermodynamic integral of the population-averaged dimensionless pressure,

$$S(\phi) = S(\phi_0) - \int_{\phi_0}^{\phi} \frac{1}{R} \sum_{r=1}^{R} \frac{Z_r}{\phi'} d\phi', \tag{6.19}$$

where Z_r is the dimensionless pressure of replica r at packing fraction ϕ'.

6.3.3 Vibrational entropy

There are many ways to measure configurational and vibrational entropy, see Ref. [22] for a review. In this section, we focus on measuring the vibrational entropies of glass states directly using two new techniques. We call the first technique the "shell" method, which is similar to Frenkel-Ladd thermodynamic integration [44] that has been previously used to measure vibrational entropies of glassy systems [20]. We call the second method replica thermodynamic integration (RTI), which is a new technique that integrates the entropy of individual replicas from the fluid state into the glass state.

6.3.3.1 Constraining shell integration method

For the shell method, we measure the vibrational entropy of a glass state by taking the initial position of each particle as a reference. A spherical hard shell centered at each particle's reference position constrains that particle to remain within the shell. Particles are unable to penetrate their own shells, but are able to freely penetrate the shells of other particles. Initially the shells are much larger than the system size and, as the simulation progresses, they are gradually shrunk. Throughout this process, the entropy lost during each decrement in shell size is summed. Eventually all of the shells become sufficiently small that they no longer overlap with each other and only

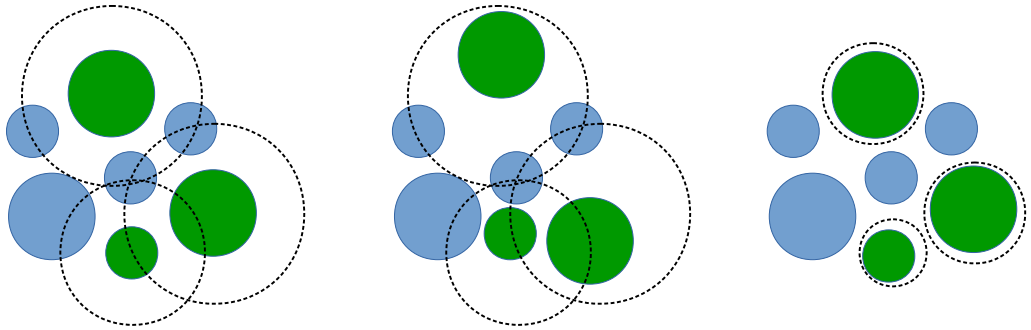

Figure 6.3. Diagram representing the shell integration method. For clarity, the shells are shown only for the three green spheres. Each sphere is contained within a hard shell that it cannot penetrate. Spheres only interact with their own shells or with other spheres that enter their shell. The shells begin very large (left) and, as the spheres are dynamically evolved the Metropolis algorithm (middle), the shells are decreased in size. Each decrement in size results in a culling in the population which corresponds to a loss of vibrational entropy. The shells are contracted and the vibrational entropy is numerically integrated via population annealing until shells no longer overlap (right). When the shells no longer overlap, the remaining entropy can be calculated analytically as an ideal gas.

contain their own particles. At this point, the particles can no longer interact with each other and the remaining entropy is simply that of each particle within its own shell, as seen in Fig. 6.3. A single particle constrained within a hard shell is simply a particle in a box or, equivalently, an ideal gas.

The shell vibrational entropy for a given configuration of particles, $S_{\text{shell}}(\vec{\boldsymbol{x}})$, where $\vec{\boldsymbol{x}}$ is the list of initial particle position vectors, is given by

$$S_{\text{shell}}(\vec{\boldsymbol{x}}) = \int_0^{\eta_f} \frac{dS[\boldsymbol{r}(\eta)]}{d\eta} d\eta + K[\boldsymbol{r}(\eta_f)], \tag{6.20}$$

where η is a parameter that controls the shell sizes, η_f is the parameter value where no shells overlap, $\boldsymbol{r}(\eta)$ is a list of the radii of all of the shells, and $K[\boldsymbol{r}(\eta_f)]$ is the sum of the ideal gas entropies of each particle in its shell. During the integration, all shells shrink at the same rate. When a shell no longer overlaps with other shells, then its integration stops, its shell radius no longer shrinks, and the sphere contained within the shell no longer contributes to the numerical integral. The remaining vibrational

entropy for the sphere/shell pair is calculated analytically and contributes to the constant K. The integration continues with the remaining shells that have overlaps. When no more shells overlap, the integration is complete and

$$K[\boldsymbol{r}(\eta_f)] = \frac{1}{N}\sum_i \log\left[\frac{4\pi}{3}(r^i_{\text{shell}} - r^i_{\text{sphere}})^3\right] + \frac{3}{2}, \tag{6.21}$$

where i enumerates the particles and r^i_{shell} is the final radius of the i^{th} shell. This formula is straightforward to understand, $\frac{4\pi}{3}(r^i_{\text{shell}} - r^i_{\text{sphere}})^3$ is simply the free volume of each shell and $3/2$ is the kinetic contribution to the entropy of the ideal gas. In general, an additional factor of $-3\log(\lambda_{\text{th}})$ is present, where λ_{th} is the thermal deBroglie wavelength, but, as before, we set $\lambda_{\text{th}} = 1$.

This method of calculating S_{vib} constrains the system so that particles are unable to switch places and, therefore, are distinguishable. This approximation becomes more accurate as structural relaxation times become very large deep within the glassy regime. This method may also be used to measure the total entropy in the fluid regime, however, since particles are no longer localized in cages, an additional term of $1 - \log(N/2)$ must be included in order to account for indistinguishability of the particles.

The numerical integration of $S_{\text{shell}}(\phi)$ is performed using population annealing as follows:

1. Choose a random glass sample from an equilibrium distribution at fixed ϕ
2. Make R_{shell} identical copies of the glass sample
3. Initialize shells so they are infinitely large
4. Evolve population using MD or MC
5. Anneal in shell size and integrate entropy.

The value of S_{shell} will depend on the number of sweeps per annealing step (see Fig. 6.9) leading to some ambiguity in the definition of S_{vib} and S_{c}. The details of the parameters used for the S_{shell} integration will be described in detail in Sec. 6.3.5.

6.3.3.2 Replica thermodynamic integration method

The second method of measuring the vibrational entropy is called replica thermodynamic integration (RTI). In abstract terms, the RTI method can be described by recasting Eq. 6.7 into a thermodynamic integral,

$$S_{\text{vib}}(\phi) = S(\phi_0) - \sum_{\nu \in c(\phi)} w_\nu(\phi) \int_{\phi_0}^{\phi} \frac{Z_\nu(\phi')}{\phi'} \, d\phi' + C. \tag{6.22}$$

where $Z_\nu(\phi')$ is an equilibrium trajectory of pressures of increasing density that ends in glass state ν at packing fraction ϕ, $w_\nu(\phi)$ the weight of ν at packing fraction ϕ, and C is a normalizing constant. This expression is different from the thermodynamic integral for the total entropy which is given by

$$S(\phi) = S(\phi_0) - \int_{\phi_0}^{\phi} \sum_{\nu \in c(\phi')} w_\nu(\phi') \frac{Z_\nu(\phi')}{\phi'} \, d\phi'. \tag{6.23}$$

In the fluid phase, for $\phi < \phi_{\text{d}}$, there is only one thermodynamic state and the two expressions are identical, however, in the glassy phase, the two become distinct due to the formation of ergodically separate glass states.

The RTI method can be realized in a PA simulation by integrating the entropy along the history of replica r at packing fraction ϕ, denoted as $S_r(\phi)$. Using the ECMC dimensionless pressure along the history,

$$S_r(\phi) = S(\phi_0) - \int_{\phi_0}^{\phi} \frac{Z_{\kappa(\phi'|r,\phi)}}{\phi'} d\phi', \tag{6.24}$$

where $\kappa(\phi'|r,\phi)$ is the replica index of the ancestor at packing fraction ϕ' of replica r at packing fraction ϕ and $S(\phi_0)$ is the initial entropy at the low packing fraction ϕ_0.

In PA, the population is an equilibrium sample of the glass states that automatically takes into account weights during resampling. This means that we can simply average S_r over the population to get the average glass entropy $S_{\rm RTI}(\phi)$, albeit with an incorrect constant of integration set in the low density limit,

$$S_{\rm RTI}(\phi) = \frac{1}{R}\sum_{r=1}^{R} S_r(\phi). \tag{6.25}$$

In the fluid regime, where replicas are individually in equilibrium, the two quantities are equal. In the glassy regime, $S_{\rm RTI} > S$ (see Fig. 6.6) because $S_{\rm RTI}$ only averages over the surviving, high-entropy replicas whereas S averages over all replicas.

To identify the correct constant of integration requires us to use the shell method or an equivalent approach. We define $S_{\rm RTI}^{(\phi')}(\phi)$ with an additive constant so that it is equal to $S_{\rm shell}(\phi')$ at packing fraction ϕ',

$$S_{\rm RTI}^{(\phi')}(\phi) = S_{\rm RTI}(\phi) + \left[S_{\rm shell}(\phi') - S_{\rm RTI}(\phi')\right]. \tag{6.26}$$

$S_{\rm RTI}^{(\phi')}(\phi)$ provides quasi-continuous estimate of $S_{\rm vib}(\phi)$ if the constant of integration is set at an appropriate point $\phi' > \phi_{\rm d}$ where $S_{\rm shell}$ can be reliably measured. It is worth mentioning that $S_{\rm RTI}$ measures entropy in a way where particle swaps are allowed and therefore naturally includes mixing entropy.

6.3.4 Equilibration

The equilibration of our simulations can be estimated in several ways. The first and perhaps most obvious is to use one of the intrinsic equilibration metric associated with population annealing [110]. A second method, discussed at the end of this section, is specific to glasses and depends on the configurational entropy.

It can be shown [29, 110] that systematic errors in PA scale as $1/R$ and that a prefactor of this scaling is an "equilibration population size", ρ_f,

$$\rho_f = \lim_{R\to\infty} R \operatorname{var}(NS), \tag{6.27}$$

where R is the population size of the simulation and S is the total entropy estimator, and the variance is measured with respect to independent runs of PA. In practice, ρ_f must be estimated from multiple runs with finite but sufficiently large population size. The intuition behind relating the variance of S to systematic errors is that when the entropy estimator has large fluctuations then independent simulations sample distinct regions of configuration space and produce different results. Furthermore, these results are, on average, biased since the correct value of an observable would be obtained from an entropically weighted average over a very large number of runs.

Although results from several independent simulations can be combined using weighted averaging, ρ_f is not a suitable measure of equilibration of the weighted average of many simulations. The extension to multiple runs was introduced in Ref. [29], using the quantity ρ_f^*,

$$\rho_f^* = \lim_{M\to\infty} R_{\text{tot}} \operatorname{var}(N\overline{S}), \tag{6.28}$$

where M is the number of simulations, $\overline{S}$ is defined in Eq. 6.16 and

$$R_{\text{tot}} = \sum_{m=1}^{M} R^{(m)}, \tag{6.29}$$

is the total population size of all the simulations, with $R^{(m)}$ the population size of the m^{th} simulation. It can be shown that ρ_f^*/R_{tot} is proportional to the systematic errors in observables obtained from the weighted average of multiple runs, all carried out with the same annealing schedule but with, perhaps, different population sizes.

Although ρ_f^* was originally calculated using bootstrap resampling [29], we found that it is preferable to estimate $\mathrm{var}(\overline{S})$ using a weighted variance of the entropies of the runs,

$$\mathrm{var}(N\overline{S}) \approx \frac{N^2 \sum_{m=1}^{M} R^{(m)} e^{NS^{(m)}} \left(S^{(m)} - \overline{S}\right)^2}{\sum_{m=1}^{M} R^{(m)} e^{NS^{(m)}}}. \tag{6.30}$$

This method gives nearly identical results to the bootstraps method used in Ref. [29], but is easier to calculate.

We can also use the configurational entropy to estimate whether the population size of the simulations is sufficient. The change in configurational entropy determines how the number of glass states decreases as the density is increased and can be used to estimate the rate of die-off of independent population members in PA. If the glass transition is suitably sharp, then it is safe to assume that the population is in equilibrium before the dynamic transition and therefore, at ϕ_d, the number of independent glass states is approximately equal to the population size, R. Since the transition is relatively sharp, it is also safe to assume that no new glass states will be discovered after the dynamic transition and that population members remain stuck in the same glass state as their ancestor at ϕ_d. With this in mind, the number of independent glass states in our population for $\phi > \phi_\mathrm{d}$, N_g, is

$$N_\mathrm{g}(\phi) = R e^{N[S_\mathrm{c}(\phi) - S_\mathrm{c}(\phi_\mathrm{d})]}. \tag{6.31}$$

This means that in order to have multiple independent glass states at packing fraction ϕ, the population size must satisfy

$$R \gg e^{-N\Delta S_\mathrm{c}}, \tag{6.32}$$

N	R	Sweeps	M	ϵ	ϕ_f
30	10^6 10^5	2.3×10^5 2.2×10^6	60 10	0.074	0.63 0.625
60	3×10^6 5×10^6	1.4×10^5	8 14	0.15	0.625
100	10^6	1.0×10^5	10	0.12	0.625

Table 6.1. Parameters for equilibrium population annealing runs: N is the number of particles, R is the population size of each run, Sweeps is the total number of ECMC sweeps per replica per run, M is the number of independent simulations, ϵ is the average culling fraction, and ϕ_f is the highest packing fraction in the runs. For $N = 60$, there were two population sizes, 3×10^6 and 5×10^6, whose simulations were combined using weighted averaging. For $N = 30$, the two sets of simulations were analyzed separately.

where $\Delta S_c = S_c(\phi) - S_c(\phi_d)$. The implication of this is that PA can only go a short way beyond the dynamic transition specific to the MCMC used to equilibrate the population.

6.3.5 Simulation details

We ran two different sets of large-scale simulations. The main set of simulations produced ensembles of glass states and measured values of S, $S_{\rm RTI}$, and Z for a quasicontinuous set of packing fractions up to packing fraction ϕ_f in the glassy regime. The second set of simulations measured $S_{\rm shell}$ for many glass samples in order to obtain an equilibrium measure of the vibrational entropy at different packing fractions and to normalize $S_{\rm RTI}$.

The parameters used in the first simulation are shown in Table 6.1. The main simulations began at $\phi_0 = 0.3$, where it is still possible to efficiently sample configurations by randomly placing spheres in a box using a rejection method. The entropy is normalized relative to the ideal gas at $\phi = 0$ with $T = 1$ and $\lambda_{\rm th} = 1$, see App. B.1. To get equilibrium samples at higher densities, we ran population annealing Monte Carlo (PA) with Event Chain Monte Carlo (ECMC) as the equilibrating dynamics. Our ECMC simuluations had chain length equal to a fixed fraction, 0.618, of the

box size. This choice of chain length is significantly longer than that of Ref. [29]. In retrospect, the choice in Ref. [29] would have been preferable. For most of the simulations, the chain schedule is given by,

$$\#\text{ event chains} = \begin{cases} 20 & \phi \leq 0.54 \\ 60 & 0.54 < \phi \leq 0.59 \\ 10 & \phi > 0.59. \end{cases} \tag{6.33}$$

The idea is to do enough chains to equilibrate in the fluid regime and partially equilibrate near the dynamic transition. At $\phi > 0.59$, when the simulation is beyond the dynamic transition, we no longer depend on using ECMC to equilibrate the entire population. Instead, we only perform a few ECMC sweeps in an attempt to equilibrate each population member locally within its own glass state. For $N = 100$, the number of event chains at each annealing step was halved, but the number of annealing steps was doubled so that the number of chains per change in packing fraction was kept constant. This was done so that the $N = 100$ culling fraction remained moderate. We also performed 10 simulations with $N = 30$, $R = 10^5$ and $\times 10$ the number of chains as in Eq. 6.33, see Table 6.1 for details.

The parameters used in the shell vibrational entropy simulations are presented in Table 6.2. We are able to measure the vibrational entropy of individual glass states with relatively small population sizes, R_{shell}, because the motion in configuration space is confined within a single glass state. Larger population sizes were tried as well and were found to produce numerically identical results which are not presented here. In order to equilibrate after each annealing step, we use the Metropolis algorithm with a fixed step size proposal chosen at each annealing step such that the acceptance ratio is always between 40% and 45%. The constraining shells are chosen such that their initial size is larger than the box containing the particles and are shrunk so that 10%

N	R_{shell}	Steps	Sweeps/step	Samples
30	10^4	3018	1600	300
60	3×10^4	6056	800	110

Table 6.2. Simulation parameters for vibrational entropy measurements using the shell method. The number of steps was different for $N = 30$ and $N = 60$ in order to keep the rate of shell contraction constant.

of the population is culled at each annealing step. Error bars for the shell method are obtained by bootstrapping over all of the shell integrated vibrational entropies.

6.4 Results

In this section we present results from a large-scale computational study of the binary hard sphere system using population annealing Monte Carlo to sample equilibrium glass states at high density. We begin by presenting equilibrium measurements of the dimensionless pressure and the entropy as a function of packing fraction in Sec. 6.4.1. Following this, we present our measurements of the configurational entropy and a detailed comparison between the two methods used to estimate the vibrational entropy in Sec. 6.4.2. We then compare the locations of the estimated jamming density, ϕ_j, and Kauzmann transition location, ϕ_{K}, in Sec. 6.4.3. Finally, we present several metrics to assess the equilibration of the simulations in Sec. 6.4.4.

6.4.1 Pressure and total entropy

The dimensionless pressure, Z, is shown in Fig. 6.4, where the solid lines correspond to simulations and the dashed black line corresponds to the BMCSL equation of state. Our simulation results deviate from the phenomenological equation of state at high packing fractions after the dynamic glass transition. Given that our simulations for sizes $N = 30$ and $N = 60$ are believed to be in statistical equilibrium for values of ϕ past ϕ_{d} and that the BMCSL equation of state is not correct for high density, as is clear by its divergence at $\phi = 1$, the deviation is at least partially due

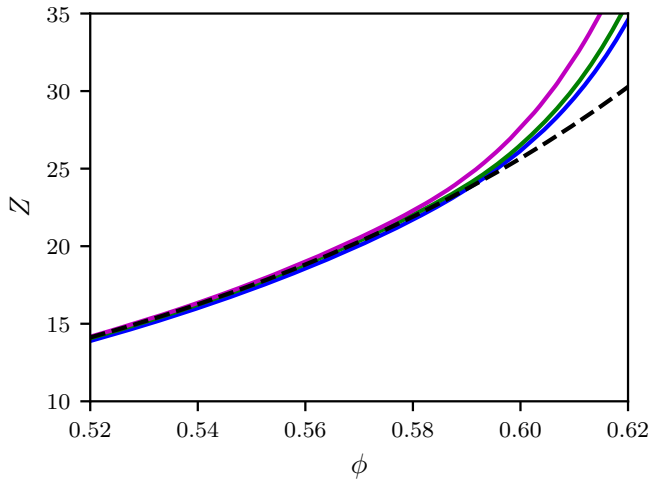

Figure 6.4. The dimensionless pressure, Z, as a function of packing fraction, ϕ for system sizes 30 (blue, bottom), 60 (green, middle), and 100 (purple, top). The dashed line is the phenomenological BMCSL equation of state (Eq. 6.3).

to the proximity to the true divergence of the pressure at random close packing, see Sec. 6.4.3.

In addition to the pressure, we have also measured the total entropy per particle, S, using ECMC, Eq. 6.19. For $N = 60$, these results can be seen in comparison to measurements from previous works in Table. 6.3.

	Z			S		
ϕ	this work	Ref. [29]	BMCSL	this work	Ref. [29]	BMCSL
0.58	22.01(1)	22.04(3)	21.90	-1.230	-1.210	-1.240
0.59	23.90(6)	23.99(14)	23.69	-1.621	-1.603	-1.629
0.60	26.4(1)	26.5	25.66	-2.044	-2.026	-2.043

Table 6.3. The dimensionless pressure and total entropy for $N = 60$ at several values of ϕ from this work compared to values obtained from Ref. [29] and to the phenomenological BMCSL equation of state (Eq. 6.3). Reference [29] entropies were modified so as to be consistent with the normalization used in the present work and to correct for the log volume term in Eq. 6.14, missing from that reference.

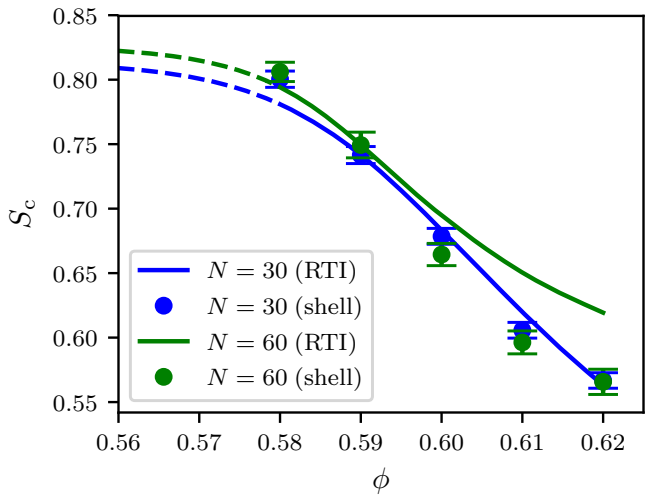

Figure 6.5. The configurational entropy measured via the replica thermodynamic integration method (lines) and the shell integration method (dots) for $N = 30$ (blue, lower) and $N = 60$ (green, upper lines). The dashed line corresponds to values of ϕ that are below the dynamic transition where the system behaves as a fluid. The RTI method produces a value of S_c at every annealing step, while the shell method was performed at $\phi = 0.58$, 0.59, 0.60, 0.61, and 0.62.

6.4.2 Vibrational and configurational entropies

We measured the configurational entropy using two different methods to obtain the vibrational entropy: replica thermodynamic integration (RTI) and shell integration. The results for these two methods are shown in Fig. 6.5, where the continuous curves correspond to the RTI method and the points correspond to the shell method. The constant of integration for RTI is obtained from the shell method at $\phi = 0.59$, that is, we use $S_{\mathrm{RTI}}^{(0.59)}$.

For the $N = 30$ system, the two methods are consistent with each other for the entire range of the simulation while for $N = 60$, the two methods agree for $\phi \lesssim 0.595$. The difference between the two methods for $N = 60$ and $\phi > 0.595$ may be due to an inadequate number of sweeps for the RTI method to estimate S_{vib}.

Although equilibrium quantities such as Z and S are largely independent of the annealing schedule provided sufficiently large population sizes are used, the values of

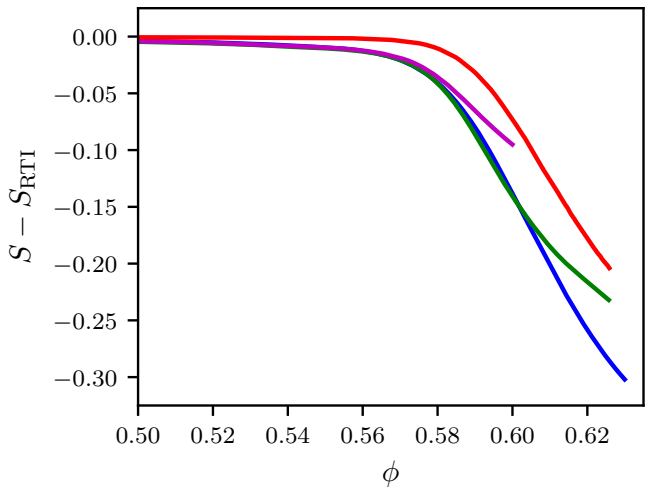

Figure 6.6. $S - S_{\mathrm{RTI}}$ as a function of ϕ for $N = 30$ (blue, bottom), $N = 60$ (green, second from bottom), and $N = 100$ (purple, second from top) with the simulation parameters of Table 6.1. The red curve (top) corresponds to an $N = 30$ simulation with ten times the number of Monte Carlo sweeps and exhibits a sharper and slightly later dynamic transition than the other simulations.

S_{RTI} and ϕ_{d}, defined roughly as the location of the shoulder of the $S - S_{\mathrm{RTI}}$ curve, depend on the annealing schedule. This can be seen explicitly in Fig. 6.6, where the blue, green, and purple curves correspond to $S - S_{\mathrm{RTI}}$ measurements using the parameters of Table 6.1 for simulations of $N = 30$, 60, and 100 particles, respectively. The blue, green and purple curves come from simulations with the same number of ECMC chains per unit change in packing fraction. The outlying red curve uppermost in the plot corresponds to an $N = 30$ simulation with $\times 10$ the number of ECMC chains as that of the standard $N = 30$ simulation (see Eq. 6.33). Here ϕ_{d} is significantly increased due to the increased ECMC equilibration. For $\phi > \phi_{\mathrm{d}}$ the curve is nearly parallel to the other curves and, after setting the constant of integration using the shell method, the configurational entropy obtained from this high sweep number simulation is nearly the same as obtained in the main, lower sweep number simulations.

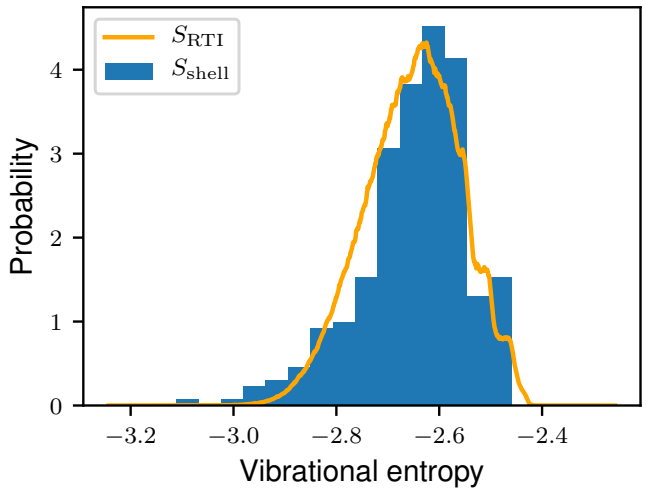

Figure 6.7. Probability density function (PDF) for S_{vib} calculated with the shell method (histogram) and RTI method (curve) for $N = 30$, $\phi = 0.60$.

S_{vib}	μ	σ	skewness	kurtosis
shell	-2.65	0.11	-0.90	1.29
RTI	-2.65	0.09	-0.25	2.76

Table 6.4. Statistics of the probability distribution functions of the vibrational entropy measured by the shell and RTI methods for $N = 30$, $\phi = 0.60$.

Finally, Fig. 6.6 illustrates an important feature of population annealing. For $\phi < \phi_{\mathrm{d}}$, we see that $S - S_{\mathrm{RTI}}$ is nearly zero. This is because each replica is equilibrated by ECMC and has the same properties as the ensemble of replicas. On the other hand, for $\phi > \phi_{\mathrm{d}}$ ergodicity is broken and each replica is confined to a single glass state so that the ensemble average no longer equals an average obtained from a single replica using ECMC. Nonetheless, the resampling step in PA ensures that ensemble averages represent equilibrium properties well beyond ϕ_{d}, as discussed in more detail in Sec. 6.4.4.

Figure 6.5 shows that values of S_{vib} obtained from RTI and the shell method give consistent results when averaged over many glass states. However, looking at the distribution of values from the two methods we can see some interesting and, as yet,

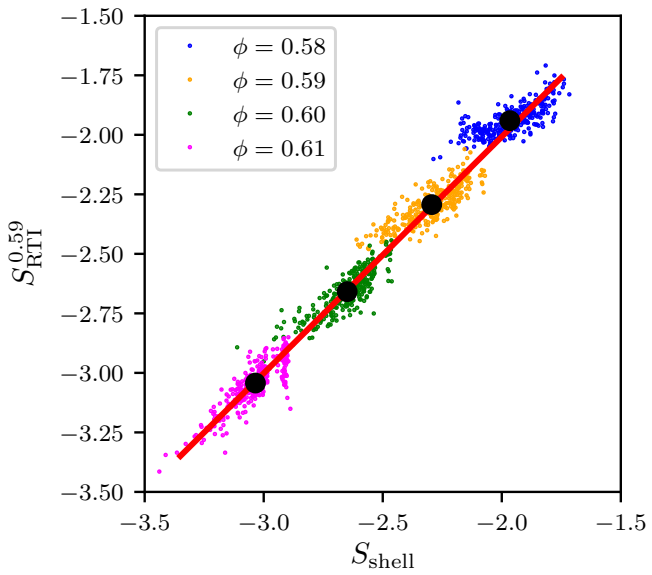

Figure 6.8. Scatter plot of $S_{\mathrm{RTI}}^{(0.59)}$ versus S_{shell} for 300 $N = 30$ glass samples at $\phi = 0.58$, 0.59, 0.60, and 0.61 (from right to left). The line of best fit for $\phi = 0.59$, 0.60, and 0.61 has a slope of 0.99 and intercept -0.02, showing excellent correspondence. The vertical sets of high-entropy values that can be seen at $\phi = 0.59$ through 0.61 correspond to the high-entropy plateaus in the vibrational entropy histograms.

not fully explained differences. Figure 6.7 shows the measured probability distribution functions of S_{shell} and S_{RTI} for $N = 30$ at $\phi = 0.60$. The histogram corresponds to S_{shell} values from 300 samples chosen randomly from the PA simulations and the curve corresponds to S_{RTI} values from 6×10^7 samples. The data was normalized so that the averages of S_{RTI} and S_{shell} are equal at $\phi = 0.59$, which is the same as the normalization used for S_{c}. As seen in the plot and Table 6.4, both distributions exhibit the same general features, but there is a significant discrepancy between the two in the low entropy tail where S_{shell} has significantly more weight than S_{RTI}. Another interesting characteristic of the distributions is the presence of flat steps in the RTI distribution and a second maximum in the shell histogram at high entropies.

A different perspective on the relationship between S_{shell} and S_{RTI} is seen in their joint probability distribution, shown in Fig. 6.8. The different sets of colored points correspond to glass states at different packing fractions, the large black points corre-

spond to the average entropies at each packing fraction, and the red line is a line of best fit for the $\phi = 0.59$, 0.60, and 0.61 data. At high packing fractions, S_{shell} and S_{RTI} are strongly correlated and the resulting best fit line has slope 0.99 and intercept -0.02. At these densities, the joint distributions are approximately symmetric about the linear fit, but there are two distinct features of the joint distribution that require explanation. The first is that there are clusters of points at high entropy that have a narrow distribution of S_{shell} values and a wide distribution of S_{RTI} values. These clusters are most pronounced for high packing fraction. The second feature is that the low entropy tail of the distribution is skewed toward higher values of S_{RTI}. This feature is most pronounced for low packing fraction. The high entropy clusters correspond to the rightmost plateau in the histogram shown in Fig. 6.7. These high entropy clusters require more investigation.

The points in the low entropy skewed part of the joint distribution that are apparent for lower packing fractions in Fig. 6.8 may represent configurations of particles that are at least partially fluid-like having diffusive rather than caged particle motions. While the RTI method should capture the full entropy of these configurations, the limited number of Metropolis sweeps used in the shell method may fail to fully explore the configuration space of these fluid-like configurations. On the other hand, for high packing fraction, the shell method finds high entropy glass states for which the RTI method finds substantially lower entropies. This may be due to the small number of ECMC chains per annealing step used in the RTI method.

One possible issue with S_{shell} is that it is numerically poorly behaved at high density. When initializing the shells for a glass state, there is a possibility that two spheres will be nearly touching. The resulting shells will be only slightly larger than the sphere sizes and, accordingly, the ideal gas entropies associated with those shells will be large due to the logarithm of the volume present in Eq. 6.21. The numerical integration of S_{shell} will compensate for the logarithmic term by having many more

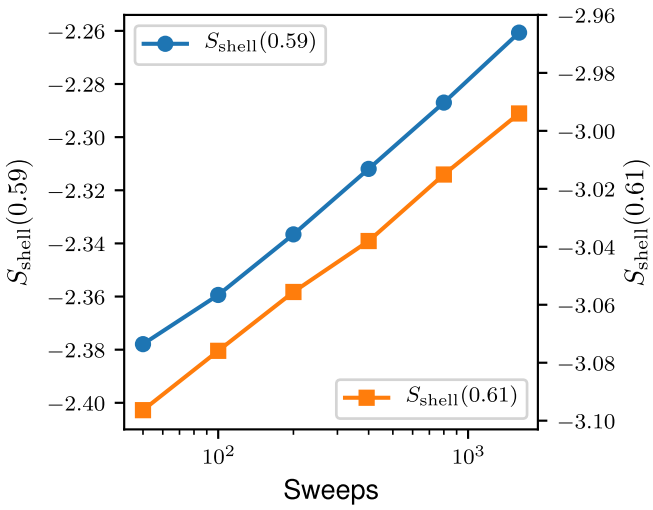

Figure 6.9. Shell vibrational entropy, $S_{\text{shell}}(\phi)$, as function of sweeps per annealing step at packing fractions 0.59 (circles, left axis) and 0.61 (squares, right axis) for $N = 30$. Both curves show logarithmic growth with the number of sweeps.

annealing steps, however, this amounts to subtracting one large number from another and is, therefore, error-prone. This numerical instability is inherent to Frenkel-Ladd techniques and becomes more problematic as the density approaches jamming.

Vibrational entropy, and thus also configurational entropy, is not uniquely defined because glass states are metastable and vibrational entropy slowly increases as the time allowed to explore configuration space increases. Figure 6.9 is a plot of the shell vibrational entropy as a function of sweeps for glass samples at $\phi = 0.59$ and $\phi = 0.60$, where each data point corresponds to an average over twenty different samples. The plot exhibits a linear-log behavior, which is consistent with the known logarithmic relaxation dynamics of configurational glasses.

6.4.3 Transition locations

The jamming density, ϕ_j, can be estimated by making a free volume fit [98] to the dimensionless pressure at densities beyond equilibrium,

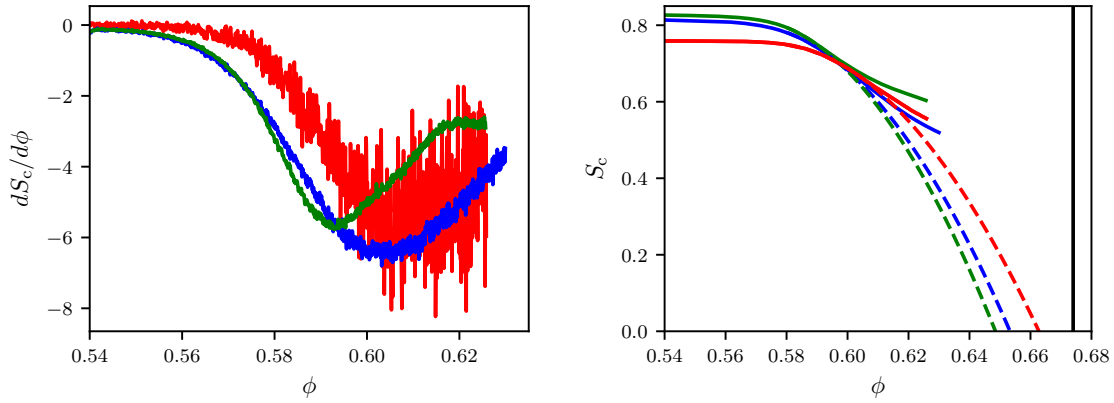

Figure 6.10. (left panel) The derivative of the configurational entropy with respect to packing fraction for $N = 30$, (simulations with 2×10^5 sweeps, blue; simulations with 2×10^6 sweeps, red) and $N = 60$ (green). We use the $dS_c/d\phi$ plot to determine the fitting range for our data. (right panel) Solid lines are S_c curves from the RTI data and dashed lines are best fit quadratic extrapolation of these curves. For $N = 30$ with 2×10^5 and 2×10^6 sweeps, the fit is over the range $\phi = 0.58$ to 0.595 and $\phi = 0.585$ to 0.595, respectively, and for $N = 60$ the fit is over the range $\phi = 0.58$ to 0.59. The vertical solid line is the jamming point as determined by the free volume fit.

$$Z = \frac{d'\phi_j}{\phi_j - \phi}. \tag{6.34}$$

We make this fit for $\phi > 0.61$ and obtain values of d' equal to 2.83 and 2.85 and ϕ_j equal to 0.676 and 0.673 for $N = 30$ and $N = 60$, respectively. We also performed this for $N = 30$ with ten times the number of sweeps and obtained an estimated ϕ_j of 0.677 with $d' = 2.91$. Because our simulations have fallen out of equilibrium over most of the range of the fits, the measured dimensionless pressures are higher than their equilibrium values and the estimated values of ϕ_j act as lower bounds to the random close packed density, ϕ_{rcp}. Our measured values of $\phi_j = 0.673$ and 0.676 are slightly larger than those found in Ref. [85] and within error bars of those found in Ref. [29].

We can estimate the location of the Kauzmann transition by extrapolating $S_c \to 0$ with our curves obtained with the RTI method. As shown in the left panel of Fig. 6.10, the $dS_c/d\phi$ curves are linear for $\phi > \phi_d$ as long as the simulation remains in

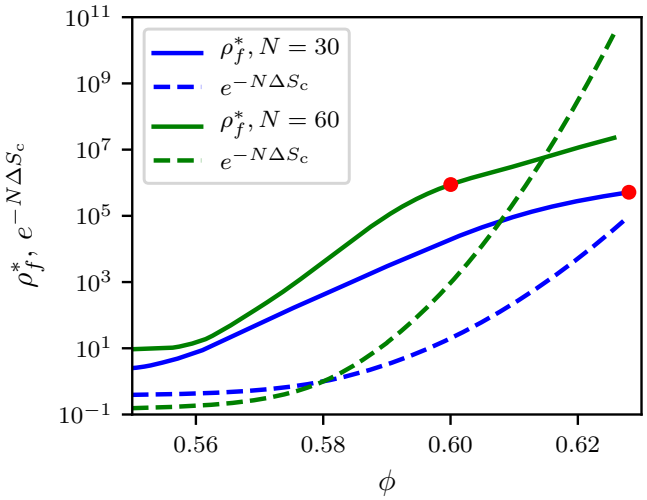

Figure 6.11. Plot of ρ_f^* (solid) and $e^{-N\Delta S_c}$ (dashed) vs ϕ for $N = 30$ (blue, lower) and $N = 60$ (green, upper). Extrapolated values of S_c were used for $\phi > 0.59$. The red dots on the ρ_f^* curves correspond to the point where the simulation falls out of equilibrium according to $\rho_f^*/R_{\text{tot}} = 0.01$, where R_{tot} is the population summed over all simulations.

equilibrium. We make a quadratic fit to S_c for $N = 30$ with $0.58 < \phi < 0.595$ and $N = 60$ with $0.58 < \phi < 0.59$ and obtain ϕ_K estimates of 0.653 and 0.649, respectively. We also performed this for $N = 30$ with ten times the number of event chains over the range of $0.585 < \phi < 0.595$ and obtained an estimated ϕ_K of 0.663. These fits are shown as dotted lines in the right panel of Fig. 6.10. The fit ranges were chosen to be after the dynamic glass transition but within the equilibrated regime. The resulting extrapolations are generally consistent and support the inequality $\phi_K < \phi_j \leq \phi_{rcp}$. Note that if the RTI data is normalized using a shell simulation with more sweeps, see Fig. 6.9, then ϕ_K will shift to a yet lower value. Thus the simulations suggest that a thermodynamic glass transition occurs at finite pressure. This is the central physics result of this work.

6.4.4 Equilibration

As discussed in Sec. 6.3.4, one can estimate the systematic errors of a weighted average of many PA simulations using the quantity ρ_f^* (see Eq. 6.28). Following Ref. [29], we set $\rho_f^* < 0.01 R_{\rm tot}$ as the threshold for equilibration of the weighted average of the simulations. Here $R_{\rm tot}$ is the total population of the combined simulations. The solid lines in Fig. 6.11 show ρ_f^* as a function of packing fraction for $N = 30$ (blue, lower) and $N = 60$ (green, upper). The equilibration threshold is $\rho_f^* = 6 \times 10^5$, 9.4×10^5 for $N = 30, 60$, respectively. These thresholds are shown by red dots in the figure and suggest that the $N = 30$ simulations are equilibrated over the entire range up to $\phi = 0.63$ while the $N = 60$ results fall out of equilibrium at $\phi = 0.60$.

Figure 6.11 shows a scatterplot of the joint distribution of the dimensionless pressure, Z, and total entropy, S, at $\phi = 0.60$ for the three system sizes. Each point represents a single simulation included in the weighted average of observables and in the computation of ρ_f^*. The negative slope of the joint distribution shows that runs with higher values of S have lower values of Z, as expected. In weighted averaging, these high entropy runs are more heavily weighted so that the high entropy tail of the distribution must be well-sampled to accurately estimate Z, ρ_f^* and S. Furthermore, for insufficient sampling, ρ_f^* and S will tend to be underestimated and Z overestimated. It is clear that the high entropy tail is poorly sampled for the $N = 60$ and $N = 100$ runs but perhaps adequately sampled for $N = 30$. Thus, it seems likely that the range of equilibration suggested by the red dots in Fig. 6.11 is overly optimistic due to insufficiently many runs used in estimating ρ_f^*. A more conservative approach would be to extrapolate ρ_f^* using its nearly pure exponential behavior before the knee of the curve. An exponential fit in the range before the knee of curve is insensitive to the precise fitting range. Setting the fitted function to the equilibration threshold yields more conservative estimates that equilibration is achieved for $\phi \lesssim 0.618$ for $N = 30$ and $\phi \lesssim 0.598$ for $N = 60$.

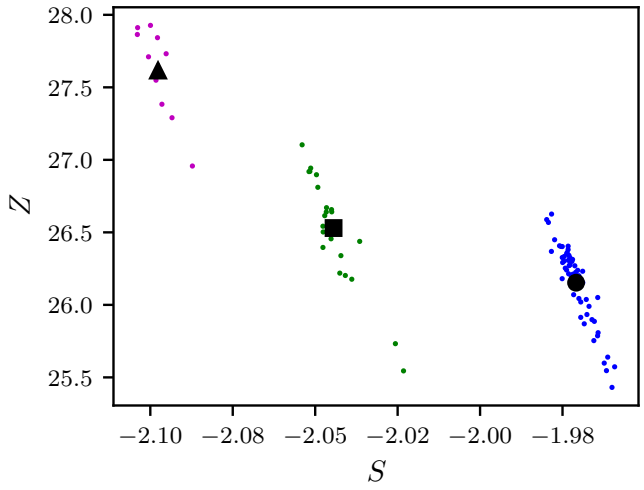

Figure 6.12. Scatter plot of dimensionless pressure Z vs the total entropy per particle S at $\phi = 0.60$ for $N = 30$ (bottom, circle), $N = 60$ (middle, square), and $N = 100$ (top, triangle). Small symbols correspond to independent runs and large black symbols correspond to the weighted average of all runs.

In Fig. 6.11, we also plot $e^{-N\Delta S_c}$ vs ϕ, where $\Delta S_c(\phi) = S_c(\phi) - S_c(\phi^*)$. In this expression we used the fitted value of S_c for $\phi > 0.59$, described in the previous section instead of the measured value since we believe the extrapolation of S_c is more accurate than the measured values deep in the glassy regime and leads to a more conservative criterion for equilibration. The initial packing fraction $\phi^* = 0.58$ is chosen to be close to the dynamic transition, $\phi^* \approx \phi_d$, where each replica in the population is assumed to become trapped in a distinct glass state. For values of ϕ such that $e^{-N\Delta S_c} < 0.01 R_{tot}$ (shown in the plot by the height of the red dot), the combined simulations have sampled more than 100 equilibrium glass states. We see that this occurs over the whole range for which the $\rho_f^* < 0.01 R_{tot}$ criteria is satisfied.

6.5 Discussion

In this work we have introduced a new method to compute the configurational entropy of structural glasses and presented several new results relating to the glassy regime of binary hard sphere mixtures. Most importantly, we found new estimates of

the Kauzmann density, ϕ_{K}, and bounds to the random close packing density, ϕ_{rcp}. The jamming density was obtained by fitting the pressure divergence and provides a lower bound on ϕ_{rcp} that is slightly larger than previous estimates. The Kauzmann density was extrapolated from the configurational entropy measured in the equilibrium fluid regime beyond the dynamic glass transition, $\phi > \phi_{\mathrm{d}}$. Our results suggest that the Kauzmann transition occurs prior to random close packing, $\phi_{\mathrm{K}} < \phi_{\mathrm{rcp}}$, so that a thermodynamic glass transition does indeed exist in this system.

To obtain these results we introduced two new computational methods for calculating the vibrational entropy of a configurational glass: the shell method, which is a variant of the Frenkel-Ladd method, and replica thermodynamic integration, which relies on integrating the entropy of individual replicas in population annealing starting from low density. Replica thermodynamic integration produces a continuous curve of vibrational entropy but requires an alternative method to set an additive constant.

Population annealing has been shown to be capable of equilibrating configurational glasses beyond the dynamic glass transition [29]. Here the dynamic transition is associated with the computational method, either Markov chain Monte Carlo or molecular dynamics, that drives equilibration in the algorithm at each packing fraction. The limitation in probing equilibrium properties with population annealing is related to the precipitous decline in configurational entropy as the density is increased. In population annealing, each member of the population is frozen in a glass state at ϕ_{d} so that upon entering the glass regime there are R glass states, where R is the population size. As the packing fraction increases, the number or distinct glass states in the population decreases exponentially as $R\, e^{N[S_{\mathrm{c}}(\phi)-S_{\mathrm{c}}(\phi_{\mathrm{d}})]}$ and collapses into a single glass state when this number reaches one. The exponential dependence on N explains why the method is restricted either to small systems or packing fractions only slightly above ϕ_{d}. On the other hand, population annealing and closely related techniques such as parallel tempering are the only methods known to us for going

beyond the dynamic transition of a Monte Carlo or molecular dynamics scheme acting at a fixed packing fraction. It should be noted that the decrease in glass configurations with density imposes the same limitations on parallel tempering in going beyond the dynamic transition. However, in continuously polydisperse systems for which the swap algorithm is effective, population annealing (or parallel tempering) combined with the swap algorithm would permit direct measurements of equilibrium properties somewhat beyond the dynamic transition for the swap algorithm, which is already deep within the glassy regime.

CHAPTER 7

OUTLOOK

The work presented in this thesis falls into three categories: improvements to the population annealing algorithm, numerical approaches to glassy systems, and the application of population annealing to glassy systems. Going forward, there are many different possible avenues of research in each of these directions. Here we discuss a few ideas that follow directly from the work presented so far, and a few ideas that are entirely new.

Perhaps most obvious would be to extend the configurational glass work that was discussed in Ch. 6. There are still several open questions pertaining to the replica thermodynamic integration method and the distinct features present in the S_{RTI} histogram. A first step would be to collect S_{RTI} data for larger system sizes such as $N = 60$ and $N = 100$. This would help to determine if the features are a result of finite-size effects specific to the $N = 30$ system. Another possible, numerically cheap direction would be to simulate systems with 31 or 32 particles. If the histogram features are a result of finite-size effects, then it is quite possible that simply adding one or two particles will result in a measurable difference in $P(S_{\mathrm{RTI}})$.

The replica thermodynamic integration method is not limited to configurational glasses. Configurational entropy is also a quantity of interest in the field of spin glasses, where it is referred to as "complexity". Applying the RTI method to a spin glass may reveal new strengths, or weaknesses, to the RTI approach. Systems where the behavior of the configurational entropy is understood, such as the p-spin model, would be ideal test candidates. Finite-dimensional spin glasses, while not as

well understood, would provide a more interesting environment in which RTI could be tested. Apart from glasses, any system that has a rough free energy landscape could, in principle, be a candidate for measuring S_{RTI}. A non-zero value of S_{RTI} is a direct measure of a lack of dynamic equilibration which, in and of itself, is a possible use-case. In order to measure S_{RTI}, one would either have to measure entropies or, similarly, free energies of individual population members. Integrating the entropy of individual population members has been performed in the microcanonical ensemble, e.g. Ref. [95], and could easily be performed in the canonical ensemble using the thermodynamic relation

$$T = \left(\frac{\partial E}{\partial S} \right)_{N,V} . \tag{7.1}$$

This equation can be rearranged and numerically integrated to obtain the entropy of each replica. Normalization of S_{RTI} is specific to the model and would have to be considered carefully.

While running large-scale simulations of hard sphere configurational glasses, we have occasionally found partially ordered states. These are collections of particles that have some clear spatial symmetries but do not form a normal ordered state. While the thermodynamically preferred state of our binary mixture is a phase separated fcc crystal, this is only true in the thermodynamic limit and it is possible that our small systems prefer more exotic particle configurations. It is worth mentioning that we did not find any evidence that partially ordered states were the cause of the plateaus found in the S_{RTI} histograms.

Another direction would be to focus on the development of PA. Population annealing Monte Carlo falls into a large family of genetic algorithms and particle filters that overlap to varying degrees. For instance, both the microcanonical version of PA and nested sampling [10,11,74,93,100], while developed independently, are essentially the same algorithm. Despite their overwhelming similarities, there are substantial dif-

ferences in the way the two algorithms resample the population and estimate errors. In addition, there is a substantial difference in the backgrounds of the people working on the two algorithms. One obvious direction for future PA research would be to consolidate the ideas from both algorithms and, potentially, reach out for collaboration with Skilling, et al. It would also be prudent to better understand the place that population annealing has in the broader family of sequential Monte Carlo algorithms.

A second research topic specific to the development of population annealing is the implementation and optimization of multi-parameter simulations. Systems which contain several control parameters are common in physics and in many cases, reaching a target phase while in equilibrium is most efficiently done via an annealing path that includes several parameters. The notions of multiparameter annealing and optimal paths to equilibrium are not new [96], however, they have not been implemented in a population annealing simulation. Since population annealing is a fundamentally different scheme than parallel tempering or simulated annealing, it is worth performing exploratory research. A topic that is related to multi-parameter annealing is the use of auxiliary fields that are eventually annealed to zero. Physical fields that are annealed to zero are common, for instance annealing a magnetic field towards a zero-field target, however, auxiliary fields do not have to be physical and can potentially speed equilibration. The goal of such research would be to identify an optimal annealing path, possibly including non-physical fields, that results in a shortcut to equilibration at the target.

Stable high-density glass samples have been produced in a lab setting using vapor deposition techniques [62, 103] that reach regimes of the glass transition that had not been previously accessible experimentally. It is believed that vapor deposited glass samples are able to relax to equilibrium configurations because of additional mobility at the surface of the deposition. A numerical scheme to emulate a vapor deposition would be easy to design using population annealing, parallel tempering,

or even simulated annealing. An example layout of a simulation would be to have a rectangular box that is square in two dimensions but with an lengthened height. The square directions would have periodic boundary conditions while the long direction would have walls. In addition to the standard pairwise interactions, there would be a potential that pulls the spheres towards the bottom of the box. There are many possible potentials to choose from but, in general, a potential would be chosen to result in a relatively tight packing near the bottom of the box and a nearly unconstrained gas at the top. As the simulation progressed, either the range of the potential or the strength of the potential would be gradually increased via annealing so as to pull more particles into a growing dense region. The end goal would be to build up a 3D glass sample at the bottom of the box. Equilibration between annealing steps would take the form of Monte Carlo or molecular dynamics moves. The Metropolis algorithm would be trivial to implement but tends to be inefficient in configurational glass simulations. Event chain Monte Carlo has a general form that works with long range interactions and likely can be modified to work within an external field. Molecular dynamics simulations can definitely describe the motion of interacting particles in fields with the downside of being numerically expensive. The general outline of a simulated annealing simulation would be:

1. Randomly initialize particles in an elongated box with zero field
2. Increase the strength of the field
3. Equilibrate using MC or MD
4. Repeat steps 2,3.

A population annealing or parallel tempering scheme could also be performed in the canonical ensemble or even in the grand canonical ensemble if a chemical potential was included. A proof of principle simulated annealing simulation would be particularly easy to implement using HOOMD-blue [6] with molecular dynamics as the

equilibrating method. There are a few places where such a simulation may run into trouble. The bottom of the box being a flat wall may result in the formation of non-trivial structures – a possible way around this would be to implement a disordered hard wall so as to minimize any tendency to crystallize or form 2D glassy layers.

A related simulation direction would be to anneal in dimension. The conceptually simplest implementation would be to start with a periodic 4D cube and to then gradually decrease the length of one side of the cube until the aspect ratio becomes large enough to approximate a 3D system. Another implementation would be to anneal from 4D to 3D using a potential in one dimension that would be gradually increased until the particles formed a near-3D shape. Both of these methods would introduce additional degrees of freedom through which the system could relax, but they also introduce the complication of analyzing a near-3D glass instead of an actual 3D glass. It would be necessary and interesting to compare the physical properties of such a system to the equilibrium properties of a truly 3D system.

APPENDIX A

SPIN GLASS APPENDIX

A.1 Covariance inequality

In Sec. 5.2.2 we showed that $\rho_t - 1 > \rho_f$, at least for the exact free energy version of PA, if the following inequality holds,

$$2 \text{ cov} \left(\delta\tilde{R}_k, \sum_{j=N_{T-1}}^{k+1} \sum_{i=1}^{\tilde{R}_{j+1}} z(\tau_i^j) \right) - \text{var} \left(\sum_{j=N_{T-1}}^{k+1} \sum_{i=1}^{\tilde{R}_{j+1}} z(\tau_i^j) \right) > 0. \tag{A.1}$$

To establish this inequality, we begin by noting that $\delta\tilde{R}_k$ is the sum of deviations from the initial population size R that have accumulated during each resampling step,

$$\begin{aligned} \delta\tilde{R}_k &= \sum_{j=N_{T-1}}^{k+1} \sum_{i=1}^{\tilde{R}_{j+1}} \left(n(\tau_i^j) - 1 \right) \\ &= \sum_{j=N_{T-1}}^{k+1} \sum_{i=1}^{\tilde{R}_{j+1}} \left(\tau_i^j + z(\tau_i^j) - 1 \right). \end{aligned} \tag{A.2}$$

Using this result to expand the covariance term shows that the desired inequality, Eq. A.1 can be re-written as

$$2 \text{ cov} \left(\sum_{j=N_{T-1}}^{k+1} \sum_{i=1}^{\tilde{R}_{j+1}} \tau_i^j, \sum_{j=N_{T-1}}^{k+1} \sum_{i=1}^{\tilde{R}_{j+1}} z(\tau_i^j) \right) + \text{var} \left(\sum_{j=N_{T-1}}^{k+1} \sum_{i=1}^{\tilde{R}_{j+1}} z(\tau_i^j) \right) > 0. \tag{A.3}$$

The variance is obviously non-negative but we do not have a proof that the covariance term is also non-negative. However, we can motivate this assertion by noting that

if the population at an earlier resampling step is stochastically increased ($z > 0$) then later population sizes will tend to be increased, i.e. $\mathrm{cov}[\tilde{R}_{j'}, z(\tau_i^j)] \geq 0$ for all i and all $j' < j$. Furthermore, if the population is stochastically increased at an early resampling step, since it is now larger, it will better explore the low energy tail of the Gibbs distribution so that $\mathrm{cov}[\tau_{i'}^{j'}, z(\tau_i^j)] \geq 0$ for all i and i', and all $k < j$. These two mechanisms both cause the covariance term in Eq. A.3 to be positive. It is worth noting that $\mathrm{cov}[\tau_{i'}^{j'}, z(\tau_i^j)] = 0$ for all i and i', and all $j' < j$.

A.2 Variance expansion

The variance term from Eq. A.3 can be further expanded to get to a form similar to that of Eq. (5.40). To do this, we make the approximation

$$\mathrm{var}\left(\sum_{j=N_{T-1}}^{k+1}\sum_{i=1}^{\tilde{R}_{j+1}} z(\tau_i^j)\right) \approx \sum_{j=N_{T-1}}^{k+1}\sum_{i=1}^{\tilde{R}_{j+1}} \mathrm{var}[z(\tau_i^j)]. \tag{A.4}$$

It is possible to do this step exactly by noting that only z_i from the same family can be correlated and by including covariance terms between intra-family $z(\tau_i)$, however, these terms have been found to be numerically insignificant.

The sum of the variances of $z(\tau_i)$ during an annealing step can be calculated using a method similar to that in Sec. 5.2.3. We begin by writing the values and probabilities of $z(\tau)$,

$$z(\tau) = \begin{cases} \tau - \lfloor\tau\rfloor & \text{w/ prob. } \lceil\tau\rceil - \tau \\ \tau - \lceil\tau\rceil & \text{w/ prob. } \tau - \lfloor\tau\rfloor, \end{cases} \tag{A.5}$$

which means that we can write the variance explicitly,

$$\mathrm{var}[z(\tau)] = (\tau - \lfloor\tau\rfloor)^2(\lceil\tau\rceil - \tau) + (\tau - \lceil\tau\rceil)^2(\tau - \lfloor\tau\rfloor). \tag{A.6}$$

For culling fraction small, we can make the approximation that within a single annealing step, τ is a Gaussian random variable. This allows us to replace the sum of variances with an integral that can be calculated explicitly. Using Eq. A.6,

$$\sum_{i=1}^{\tilde{R}_{j+1}} \mathrm{var}[z(\tau_i^j)] \approx \int_{-\infty}^{\infty} \mathrm{var}(z(\tau))\mathcal{N}(\tau; 1, \Delta\beta\sigma_E)\,\mathrm{d}\tau. \tag{A.7}$$

$$= 2\epsilon - 2\pi\epsilon^2. \tag{A.8}$$

Summing over all annealing steps and simplifying the expression assuming a constant culling fraction gives the desired result

$$\sum_{j=N_{T-1}}^{k+1} \sum_{i=1}^{\tilde{R}_{j+1}} \mathrm{var}[z(\tau_i^j)] \approx \sum_{j=N_{T-1}}^{k+1} \left[2\epsilon(l) - 2\pi\epsilon(l)^2\right] \tag{A.9}$$

$$= \left(2\epsilon - 2\pi\epsilon^2\right) k. \tag{A.10}$$

APPENDIX B

CONFIGURATIONAL GLASS APPENDIX

B.1 Entropy normalization

To find the normalization of the entropy per particle, we start with the thermodynamic definition,

$$\frac{\partial S}{\partial \mathcal{V}} = \frac{P}{T}, \tag{B.1}$$

where the Boltzmann constant has been set to unity and $\mathcal{V}$ is the volume per particle. This equation can be integrated in volume to get

$$S(\mathcal{V}) = \int_{\mathcal{V}_\mathrm{I}}^{\mathcal{V}} \frac{P}{T} d\mathcal{V}' + S(\mathcal{V}_\mathrm{I}), \tag{B.2}$$

where $\mathcal{V}_\mathrm{I}$ is the ideal gas limit of $\mathcal{V}$ with the property that $\mathcal{V}_\mathrm{I} \gg 1$. In this limit, we can make the approximation that the system begins as a binary mixture of two ideal gasses with entropy

$$S(\mathcal{V}_\mathrm{I}) = \log\left(\frac{\mathcal{V}_\mathrm{I}}{\lambda_\mathrm{th}^3/2}\right) + \frac{5}{2}, \tag{B.3}$$

where λ_th is the thermal deBroglie wavelength. By setting $\lambda_\mathrm{th} = 1$, the resulting entropy at $\mathcal{V}$ becomes

$$S(\mathcal{V}) = \lim_{\mathcal{V}_\mathrm{I}\to\infty}\left[\int_{\mathcal{V}_\mathrm{I}}^{\mathcal{V}} \frac{P}{T} d\mathcal{V}' + \log(\mathcal{V}_\mathrm{I})\right] + \frac{5}{2} + \log(2). \tag{B.4}$$

We can rewrite this in terms of packing fraction by using the relationship between ϕ and $\mathcal{V}$,

$$\phi = \frac{4\pi}{3\mathcal{V}} \frac{r_0^3 + r_1^3}{2}, \tag{B.5}$$

where r_0 and r_1 are the radii of the small and large particles which, in our system, have values $r_0 = 1$ and $r_1 = 1.4$. Changing variables from $\mathcal{V}$ to ϕ gives

$$S(\phi) = \lim_{\phi_\mathrm{I} \to 0} - \int_{\phi_\mathrm{I}}^{\phi} \frac{Z(\phi')}{\phi'} d\phi' - \log(\phi_\mathrm{I}) + \log\left(\frac{4\pi}{3} \frac{r_0^3 + r_1^3}{2}\right) + \frac{5}{2} + \log 2, \tag{B.6}$$

where the $\log(\phi_\mathrm{I})$ term cancels with the logarithmically diverging pressure in the integral. By rearranging, we get a form without explicit divergences,

$$S(\phi) = \lim_{\phi_\mathrm{I} \to 0} \int_{\phi_\mathrm{I}}^{\phi} \frac{1 - Z(\phi')}{\phi'} d\phi' - \log(\phi) + \log\left(\frac{4\pi}{3} \frac{r_0^3 + r_1^3}{2}\right) + \frac{5}{2} + \log 2. \tag{B.7}$$

Typically we do not integrate from $\phi = 0$ and instead start at a non-zero initial packing fraction. Starting at ϕ_0, the total entropy in a population annealing simulation is given by

$$S(\phi) = S(\phi_0) - \int_{\phi_0}^{\phi} \frac{\tilde{Z}(\phi')}{\phi'} d\phi', \tag{B.8}$$

where $\tilde{Z}$ is the equilibrium pressure obtained during the population annealing simulation. The normalization entropy is given by Eq. B.7,

$$S(\phi_0) = \int_0^{\phi_0} \frac{1 - Z_{\mathrm{BMCSL}}(\phi')}{\phi'} d\phi' - \log(\phi_0) + \log\left(\frac{4\pi}{3} \frac{r_0^3 + r_1^3}{2}\right) + \frac{5}{2} + \log 2, \tag{B.9}$$

where we have set ϕ_I to zero because the BMCSL equation of state, defined in Eq. 6.3, can be explicitly integrated at $\phi = 0$.

BIBLIOGRAPHY

[1] Adam, Gerold, and Gibbs, Julian H. On the temperature dependence of cooperative relaxation properties in glass-forming liquids. *The Journal of Chemical Physics 43*, 1 (1965), 139–146.

[2] Alder, S., Trebst, S., Hartmann, A. K., and Troyer, M. Dynamics of the Wang Landau algorithm and complexity of rare events for the three-dimensional bimodal Ising spin glass. *J. Stat. Mech.* P07008 (2004).

[3] Alvarez Baños *et al.*, R. Nature of the spin-glass phase at experimental length scales. *J. Stat. Mech.* P06026 (2010).

[4] Amey, Christopher, and Machta, Jonathan. Analysis and optimization of population annealing. *Phys. Rev. E 97* (Mar 2018), 033301.

[5] Amey, Christopher, and Machta, Jonathan. Measuring glass entropies with population annealing. *arXiv e-prints* (Mar. 2021), arXiv:2103.13837.

[6] Anderson, Joshua A., Glaser, Jens, and Glotzer, Sharon C. Hoomd-blue: A python package for high-performance molecular dynamics and hard particle monte carlo simulations. *Computational Materials Science 173* (2020), 109363.

[7] Anderson, P. W. Ordering and antiferromagnetism in ferrites. *Phys. Rev. 102* (May 1956), 1008–1013.

[8] Angelani, Luca, and Foffi, Giuseppe. Configurational entropy of hard spheres. *Journal of Physics: Condensed Matter 19*, 25 (jun 2007), 256207.

[9] Angell, C. A. Formation of Glasses from Liquids and Biopolymers. *Science 267* (1995), 1924.

[10] Baldock, Robert J. N., Bernstein, Noam, Salerno, K. Michael, Pártay, Lívia B., and Csányi, Gábor. Constant-pressure nested sampling with atomistic dynamics. *Phys. Rev. E 96* (Oct 2017), 043311.

[11] Baldock, Robert J. N., Pártay, Lívia B., Bartók, Albert P., Payne, Michael C., and Csányi, Gábor. Determining pressure-temperature phase diagrams of materials. *Phys. Rev. B 93* (May 2016), 174108.

[12] Banos, R Alvarez, Cruz, A, Fernandez, L A, Gil-Narvion, J M, Gordillo-Guerrero, A, Guidetti, M, Maiorano, A, Mantovani, F, Marinari, E, Martin-Mayor, V, Monforte-Garcia, J, Sudupe, A Munoz, Navarro, D, Parisi, G, Perez-Gaviro, S, Ruiz-Lorenzo, J J, Schifano, S F, Seoane, B, Tarancon, A, Tripiccione, R, and Yllanes, D. Nature of the spin-glass phase at experimental length scales. *Journal of Statistical Mechanics: Theory and Experiment* (2010), P06026.

[13] Baranau, Vasili, and Tallarek, Ulrich. Relaxation times, jamming densities, and ideal glass transition densities for hard spheres in a wide range of polydispersities. *AIP Advances 10*, 3 (2020), 035212.

[14] Barash, L. Y., Weigel, M., Borovský, M., Janke, W., and Shchur, L. N. GPU accelerated population annealing algorithm. *Computer Physics Communications 220* (2017), 341 – 350.

[15] Barash, L. Y., Weigel, M., Shchur, L. N., and Janke, W. Exploring first-order phase transitions with population annealing. *The European Physical Journal Special Topics 226* (2017), 595–604.

[16] Barzegar, Amin, Pattison, Christopher, Wang, Wenlong, and Katzgraber, Helmut G. Optimization of population annealing monte carlo for large-scale spin-glass simulations. *Phys. Rev. E 98* (Nov 2018), 053308.

[17] Bengtzelius, U, Gotze, W, and Sjolander, A. Dynamics of supercooled liquids and the glass transition. *Journal of Physics C: Solid State Physics 17*, 33 (nov 1984), 5915–5934.

[18] Bernard, Etienne P., Krauth, Werner, and Wilson, David B. Event-chain monte carlo algorithms for hard-sphere systems. *Phys. Rev. E 80* (Nov 2009), 056704.

[19] Berthier, Ludovic, and Biroli, Giulio. Theoretical perspective on the glass transition and amorphous materials. *Rev. Mod. Phys. 83* (Jun 2011), 587–645.

[20] Berthier, Ludovic, Charbonneau, Patrick, Coslovich, Daniele, Ninarello, Andrea, Ozawa, Misaki, and Yaida, Sho. Configurational entropy measurements in extremely supercooled liquids that break the glass ceiling. *Proceedings of the National Academy of Sciences 114*, 43 (2017), 11356–11361.

[21] Berthier, Ludovic, Coslovich, Daniele, Ninarello, Andrea, and Ozawa, Misaki. Equilibrium sampling of hard spheres up to the jamming density and beyond. *Phys. Rev. Lett. 116* (Jun 2016), 238002.

[22] Berthier, Ludovic, Ozawa, Misaki, and Scalliet, Camille. Configurational entropy of glass-forming liquids. *The Journal of Chemical Physics 150*, 16 (2019), 160902.

[23] Berthier, Ludovic, and Witten, Thomas A. Glass transition of dense fluids of hard and compressible spheres. *Phys. Rev. E 80* (Aug 2009), 021502.

[24] Boublík, Tomáš. Hard-sphere equation of state. *The Journal of Chemical Physics 53*, 1 (1970), 471–472.

[25] Bouchaud, Jean-Philippe, and Biroli, Giulio. On the adam-gibbs-kirkpatrick-thirumalai-wolynes scenario for the viscosity increase in glasses. *The Journal of Chemical Physics 121*, 15 (2004), 7347–7354.

[26] Brambilla, G., El Masri, D., Pierno, M., Berthier, L., Cipelletti, L., Petekidis, G., and Schofield, A. B. Probing the equilibrium dynamics of colloidal hard spheres above the mode-coupling glass transition. *Phys. Rev. Lett. 102* (2009), 085703.

[27] Bray, A. J., and Moore, M. A. Critical behavior of the three-dimensional Ising spin glass. *Phys. Rev. B 31* (1985), 631.

[28] Bray, A. J., and Moore, M. A. Scaling theory of the ordered phase of spin glasses. In *Heidelberg Colloquium on Glassy Dynamics and Optimization*, L. Van Hemmen and I. Morgenstern, Eds. Springer, New York, 1986, p. 121.

[29] Callaham, J., and Machta, J. Population annealing simulations of a binary hard-sphere mixture. *Phys. Rev. E 95* (2017), 063315.

[30] Cannella, Y., and Mydosh, J. A. Magnetic ordering in gold-iron alloys (susceptibility and thermopower studies). *Phys. Rev. B 6* (1972), 4220.

[31] Cavagna, Andrea. Supercooled liquids for pedestrians. *Physics Reports 476*, 4 (2009), 51 – 124.

[32] Chaudhuri, Pinaki, Berthier, Ludovic, and Sastry, Srikanth. Jamming transitions in amorphous packings of frictionless spheres occur over a continuous range of volume fractions. *Phys. Rev. Lett. 104* (Apr 2010), 165701.

[33] Chowdhury, Debashish. *Spin Glasses and Other Frustrated Systems*. World Scientific, Singapore, 1986.

[34] Christiansen, Henrik, Weigel, Martin, and Janke, Wolfhard. Accelerating molecular dynamics simulations with population annealing. *Phys. Rev. Lett. 122* (Feb 2019), 060602.

[35] de Almeida, J. R. L., and Thouless, D. J. Stability of the Sherrington-Kirkpatrick solution of a spin glass model. *J. Phys. A 11* (1978), 983.

[36] Derrida, B. Random-Energy Model: Limit of a Family of Disordered Models. *Phys. Rev. Lett. 45*, 2 (July 1980), 79–82.

[37] Doucet, A., de Freitas, N., and Gordon, N., Eds. *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, New York, 2001.

[38] Edwards, S. F., and Anderson, P. W. Theory of spin glasses. *J. Phys. F: Met. Phys. 5* (1975), 965.

[39] Engel, Michael, Anderson, Joshua A., Glotzer, Sharon C., Isobe, Masaharu, Bernard, Etienne P., and Krauth, Werner. Hard-disk equation of state: First-order liquid-hexatic transition in two dimensions with three simulation methods. *Phys. Rev. E 87* (Apr 2013), 042134.

[40] Fisher, D. S., and Huse, D. A. Ordered phase of short-range Ising spin-glasses. *Phys. Rev. Lett. 56* (1986), 1601.

[41] Fisher, D. S., and Huse, D. A. Absence of many states in realistic spin glasses. *J. Phys. A 20* (1987), L1005.

[42] Fisher, D. S., and Huse, D. A. Equilibrium behavior of the spin-glass ordered phase. *Phys. Rev. B 38* (1988), 386.

[43] Franz, Silvio, and Parisi, Giorgio. Phase diagram of coupled glassy systems: A mean-field study. *Phys. Rev. Lett. 79* (Sep 1997), 2486–2489.

[44] Frenkel, Daan, and Ladd, Anthony J. C. New monte carlo method to compute the free energy of arbitrary solids. application to the fcc and hcp phases of hard spheres. *The Journal of Chemical Physics 81*, 7 (1984), 3188–3193.

[45] Gibbs, J. H. Nature of the glass transition in polymers. *The Journal of Chemical Physics 25*, 1 (1956), 185–186.

[46] Gibbs, J. H., and Di Marzio, E. A. Nature of the Glass Transition and the Glassy State. *J. Chem. Phys. 28* (1957), 373.

[47] Goldstein, Martin. Viscous liquids and the glass transition: A potential energy barrier picture. *The Journal of Chemical Physics 51*, 9 (1969), 3728–3739.

[48] Gross, D.J., and Mezard, M. The simplest spin glass. *Nuclear Physics B 240*, 4 (1984), 431 – 452.

[49] Götze, W. Properties of the glass instability treated within a mode coupling theory. *Zeitschrift für Physik B Condensed Matter 60*, 2 (June 1985), 195–203.

[50] Hansen, J.P., and McDonald, I.R. *Theory of Simple Liquids (Third Edition)*, third edition ed. Academic Press, Burlington, 2006.

[51] Hastings, W. K. Monte carlo sampling methods using markov chains and their applications. *Biometrika 57*, 1 (1970), 97–109.

[52] Hopkins, Adam B., Jiao, Yang, Stillinger, Frank H., and Torquato, Salvatore. Phase diagram and structural diversity of the densest binary sphere packings. *Phys. Rev. Lett. 107* (Sep 2011), 125501.

[53] Hopkins, Adam B., Stillinger, Frank H., and Torquato, Salvatore. Densest binary sphere packings. *Phys. Rev. E 85* (Feb 2012), 021130.

[54] Hukushima, K., and Iba, Y. Population annealing and its application to a spin glass. In *The Monte Carlo Method In The Physical Sciences: Celebrating the 50th Anniversary of the Metropolis Algorithm* (2003), James E. Gubernatis, Ed., vol. 690, AIP, pp. 200–206.

[55] Isobe, Masaharu, and Krauth, Werner. Hard-sphere melting and crystallization with event-chain monte carlo. *The Journal of Chemical Physics 143*, 8 (2015), 084509.

[56] Janke, Wolfhard. Monte carlo methods in classical statistical physics. In *Computational Many-Particle Physics*, H. Fehske, R. Schneider, and A. Weiße, Eds. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008, pp. 79–140.

[57] Kamien, Randall D., and Liu, Andrea J. Why is random close packing reproducible? *Phys. Rev. Lett. 99* (Oct 2007), 155501.

[58] Kasuya, Tadao. A Theory of Metallic Ferro- and Antiferromagnetism on Zener's Model. *Progress of Theoretical Physics 16*, 1 (07 1956), 45–57.

[59] Katzgraber, H. G., Körner, M., and Young, A. P. Universality in three-dimensional Ising spin glasses: A Monte Carlo study. *Phys. Rev. B 73* (2006), 224432.

[60] Katzgraber, H. G., Palassini, M., and Young, A. P. Monte Carlo simulations of spin glasses at low temperatures. *Phys. Rev. B 63* (2001), 184422.

[61] Kauzmann, W. The nature of the glass state and the behavior of liquids at low temperatures. *Chem. Rev. 43* (1948), 219.

[62] Kearns, Kenneth L., Swallen, Stephen F., Ediger, M. D., Wu, Tian, Sun, Ye, and Yu, Lian. Hiking down the energy landscape: Progress toward the kauzmann temperature via vapor deposition. *The Journal of Physical Chemistry B 112*, 16 (2008), 4934–4942.

[63] Kirkpatrick, S., Gelatt, Jr., C. D., and Vecchi, M. P. Optimization by simulated annealing. *Science 220* (1983), 671.

[64] Kirkpatrick, T. R., and Thirumalai, D. p-spin-interaction spin-glass models: Connections with the structural glass problem. *Phys. Rev. B 36* (1987), 5388.

[65] Kirkpatrick, T. R., and Thirumalai, D. Mean-field soft-spin Potts glass model: Statics and dynamics. *Phys. Rev. B 37* (1988), 5342.

[66] Kirkpatrick, T. R., and Wolynes, P. G. Connections between some kinetic and equilibrium theories of the glass transition. *Phys. Rev. A 35* (1987), 3072.

[67] Kirkpatrick, T. R., and Wolynes, P. G. Stable and metastable states in mean-field Potts and structural glasses. *Phys. Rev. B 36* (1987), 8552.

[68] Krapivsky, P.L., Redner, S., and Ben-Naim, E. *A Kinetic View of Statistical Physics*. Cambridge University Press, 2010.

[69] Kumar, Ravinder, Gross, Jonathan, Janke, Wolfhard, and Weigel, Martin. Massively parallel simulations for disordered systems. *The European Physical Journal B 93*, 5 (May 2020), 79.

[70] Leutheusser, E. Dynamical model of the liquid-glass transition. *Phys. Rev. A 29* (May 1984), 2765–2773.

[71] Liu, Andrea J., and Nagel, Sidney R. Jamming is not just cool any more. *Nature 396*, 6706 (Nov. 1998), 21–22.

[72] Machta, J. Population annealing with weighted averages: A Monte Carlo method for rough free-energy landscapes. *Phys. Rev. E 82* (2010), 026704.

[73] Mansoori, G. A., Carnahan, N. F., Starling, K. E., and Leland, T. W. Equilibrium thermodynamic properties of the mixture of hard spheres. *The Journal of Chemical Physics 54*, 4 (1971), 1523–1525.

[74] Martiniani, Stefano, Stevenson, Jacob D., Wales, David J., and Frenkel, Daan. Superposition enhanced nested sampling. *Phys. Rev. X 4* (Aug 2014), 031034.

[75] Masri, D El, Brambilla, G, Pierno, M, Petekidis, G, Schofield, A B, Berthier, L, and Cipelletti, L. Dynamic light scattering measurements in the activated regime of dense colloidal hard spheres. *Journal of Statistical Mechanics: Theory and Experiment 2009*, 07 (jul 2009), P07015.

[76] McMillan, W. L. Domain-wall renormalization-group study of the three-dimensional random Ising model. *Phys. Rev. B 30* (1984), R476.

[77] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. Equation of State Calculations by Fast Computing Machines. *J. Chem. Phys. 21* (1953), 1087.

[78] Mezard, M., Parisi, G., and Virasoro, M. *Spin Glass Theory and Beyond*, vol. 9 of *World Scientific Lecture Notes in Physics*. World Scientific Publishing Company, 1987.

[79] Michel, Manon, Kapfer, Sebastian C., and Krauth, Werner. Generalized event-chain monte carlo: Constructing rejection-free global-balance algorithms from infinitesimal steps. *The Journal of Chemical Physics 140*, 5 (2014), 054116.

[80] Mulder, C. A. M., van Duyneveldt, A. J., and Mydosh, J. A. Susceptibility of the CuMn spin-glass: Frequency and field dependences. *Phys. Rev. B 23* (Feb 1981), 1384–1396.

[81] Newman, C. M., and Stein, D. L. Multiple states and thermodynamic limits in short-ranged Ising spin-glass models. *Phys. Rev. B 46* (1992), 973.

[82] Newman, C. M., and Stein, D. L. Simplicity of state and overlap structure in finite-volume realistic spin glasses. *Phys. Rev. E 57* (1998), 1356–1366.

[83] Newman, C. M., and Stein, D. L. Ordering and broken symmetry in short-ranged spin glasses. *J. Phys.: Condens. Matter 15* (2003), R1319–R1364.

[84] Newman, M. E. J., and Barkema, G. T. *Monte Carlo Methods in Statistical Physics*. Oxford, Oxford, 1999.

[85] Odriozola, G., and Berthier, L. Equilibrium equation of state of a hard sphere binary mixture at very large densities using replica exchange monte carlo simulations. *J. Chem. Phys. 134*, 5 (2011), 054504.

[86] O'Hern, Corey S., Silbert, Leonardo E., Liu, Andrea J., and Nagel, Sidney R. Jamming at zero temperature and zero applied stress: The epitome of disorder. *Phys. Rev. E 68* (Jul 2003), 011306.

[87] Ozawa, Misaki, Parisi, Giorgio, and Berthier, Ludovic. Configurational entropy of polydisperse supercooled liquids. *The Journal of Chemical Physics 149*, 15 (2018), 154501.

[88] Ozawa, Misaki, Scalliet, Camille, Ninarello, Andrea, and Berthier, Ludovic. Does the adam-gibbs relation hold in simulated supercooled liquids? *The Journal of Chemical Physics 151*, 8 (2019), 084504.

[89] Parisi, G. Infinite number of order parameters for spin-glasses. *Phys. Rev. Lett. 43* (1979), 1754.

[90] Parisi, G. A sequence of approximated solutions to the S-K model for spin glasses. *J. Phys. A 13* (1980), L115.

[91] Parisi, G. The order parameter for spin glasses: a function on the interval 0–1. *J. Phys. A 13* (1980), 1101.

[92] Parisi, G. Order parameter for spin-glasses. *Phys. Rev. Lett. 50* (1983), 1946.

[93] Pártay, Lívia B., Bartók, Albert P., and Csányi, Gábor. Nested sampling for materials: The case of hard spheres. *Phys. Rev. E 89* (Feb 2014), 022302.

[94] Reichman, David R, and Charbonneau, Patrick. Mode-coupling theory. *Journal of Statistical Mechanics: Theory and Experiment 2005*, 05 (may 2005), P05013.

[95] Rose, Nathan, and Machta, Jonathan. Equilibrium microcanonical annealing for first-order phase transitions. *Phys. Rev. E 100* (Dec 2019), 063304.

[96] Rotskoff, Grant M., and Crooks, Gavin E. Optimal control in nonequilibrium systems: Dynamic riemannian geometry of the ising model. *Phys. Rev. E 92* (Dec 2015), 060102.

[97] Ruderman, M. A., and Kittel, C. Indirect exchange coupling of nuclear magnetic moments by conduction electrons. *Phys. Rev. 96* (Oct 1954), 99–102.

[98] Salsburg, Z. W., and Wood, W. W. Equation of state of classical hard spheres at high density. *The Journal of Chemical Physics 37*, 4 (1962), 798–804.

[99] Sherrington, D., and Kirkpatrick, S. Solvable model of a spin glass. *Phys. Rev. Lett. 35* (1975), 1792.

[100] Skilling, John. Nested sampling for general bayesian computation. *Bayesian Anal. 1*, 4 (12 2006), 833–859.

[101] Speedy, Robin J. The hard sphere glass transition. *Molecular Physics 95*, 2 (1998), 169–178.

[102] Stein, D. L., and Newman, C. M. *Spin Glasses and Complexity*. Primers in Complex Systems. Princeton University Press, 2013.

[103] Swallen, Stephen F., Kearns, Kenneth L., Mapes, Marie K., Kim, Yong Seol, McMahon, Robert J., Ediger, M. D., Wu, Tian, Yu, Lian, and Satija, Sushil. Organic glasses with exceptional thermodynamic and kinetic stability. *Science 315*, 5810 (2007), 353–356.

[104] Talagrand, M. *Spin glasses: a Challenge for Mathematicians*. Springer, Berlin, 2003.

[105] Talagrand, M. The generalized Parisi formula. *Comptes Rendus Mathematique 337* (2003), 111.

[106] Talagrand, M. The Parisi formula. *Ann. of Math. 163* (2006), 221.

[107] Thouless, D. J., Anderson, P. W., and Palmer, R. G. Solution of 'Solvable model of a spin glass'. *Phil. Mag. 35* (1977), 593.

[108] Wang, Chengju, and Stratt, Richard M. Global perspectives on the energy landscapes of liquids, supercooled liquids, and glassy systems: Geodesic pathways through the potential energy landscape. *The Journal of Chemical Physics 127*, 22 (2007), 224504.

[109] Wang, W. Numerical simulations of Ising spin glasses with free boundary conditions: The role of droplet excitations and domain walls. *Phys. Rev. E 95* (2017), 032143.

[110] Wang, W., Machta, J., and Katzgraber, H. Population annealing: Theory and application in spin glasses. *Phys. Rev. E 92* (2015), 063307.

[111] Wang, W., Machta, J., and Katzgraber, H. G. Comparing Monte Carlo methods for finding ground states of Ising spin glasses: Population annealing, simulated annealing, and parallel tempering. *Phys. Rev. E 92* (2015), 013303.

[112] Wang, W., Machta, J., and Katzgraber, H. G. Bond chaos in spin glasses revealed through thermal boundary conditions. *Phys. Rev. B 93* (2016), 224414.

[113] Wilks, S. S. *Mathematical Statistics*. John Wiley & Sons, Inc., New York, 1962.

[114] Yosida, Kei. Magnetic properties of cu-mn alloys. *Phys. Rev. 106* (Jun 1957), 893–898.

[115] Yucesoy, B., Machta, J., and Katzgraber, H. G. Correlations between the dynamics of parallel tempering and the free-energy landscape in spin glasses. *Phys. Rev. E 87* (2013), 012104.