Direct computation of the packing entropy of granular materials

Daniel A. Asenjo-Andrews

St. Edmund's College

University of Cambridge

A Thesis submitted for the degree of

Doctor of Philosophy

September 2013

Direct computation of the packing entropy of granular materials

Daniel A. Asenjo-Andrews

Summary

Granular materials are the second most manipulated material in industry after water and their properties are of great importance for the pharmaceutical, food, mechanosynthesis and semiconductor industries. Up to 60% of the capacity of some industrial plants is used to process them.

This thesis describes computer simulations that aim to evaluate the number of distinct packings of a granular material or, more generally, the so-called ‘granular entropy’. Monte Carlo simulations are used to probe the energy landscape of jammed systems of disks interacting via a repulsive, finite-range potential. In these simulations we make use of a soft-sphere model with a hard core that approaches the hard-sphere model as the width of the soft shell is decreased.

To compute the packing entropy, we use and develop Monte Carlo techniques to determine the volumes of the basins of attraction of the potential energy minima at different system sizes. Such Monte Carlo simulations require energy minimisation after every trial move to make sure that all accepted moves keep the system within the same basin of attraction. Hence efficient energy minimisation is a point of paramount significance in this work. A first objective was to find a suitable minimisation algorithm.

We report a study of the basins of attraction for potential energy minima defined by different minimisation algorithms for an atomic system. The findings indicate that whereas some minimisation algorithms produce compact basins, others produce basins with complex boundaries or basins consisting of disconnected parts. For the remainder of our work, the FIRE algorithm was chosen because it produces compact basins at a reasonable computational cost.

Once the minimisation algorithm is chosen a numerical approach is used to compute the number of ways in which N particles can pack into a given volume V. This technique extends the existing methods in such a way that it can be applied to much larger systems than before (over 100 particles instead of 16). Many of the caveats of previous methods are addressed. Using this novel approach, the system size dependence of the number of distinct packings of a system of poly-disperse soft disks is studied. Our simulations enable us to validate a more than 20 years old conjecture due to Edwards.

The distribution of jamming densities produced by different protocols has been studied. We found that the distribution of jamming volumes that are generated by starting from different initial densities cannot be characterised by an Edwards’-style compactivity although it is possible to construct an ensemble where compactivity is well defined.

Declaration

I hereby confirm that this is my own work except where indicated. I have referenced all sources, indicated any work which is the outcome of a collaboration, referenced and put in inverted commas all quoted text and given the sources of all pictures, data, etc. that are not my own. I have not incorporated any unattributed material from reports or essays of any other students either past or present. I have not sought or used the help of any external professional agencies.

Parts of this thesis are included in work that has either been published or is to be published:

- D. Asenjo, F. Paillusson and D. Frenkel, *Numerical Calculation of Granular Entropy*, Phys. Rev. Lett. **112**, 098002, (2014).
- D. Frenkel, D. Asenjo and F. Paillusson, *The other entropy*, Mol. Phys., **111**, 3641, (2013).
- D. Asenjo, J. Stevenson, D. J. Wales, and D. Frenkel, *Visualizing Basins of Attraction for Different Minimization Algorithms*, J. Phys. Chem. B, **117**, 12717, (2013).

This dissertation has not been submitted in whole or in part for any other degree or diploma at this or any other university and does not exceed 60,000 words in length.

Signature: ....
Name: Daniel A. Asenjo-Andrews.
Date: April 15, 2014.

Acknowledgements

I would like to thank my supervisor Daan Frenkel for his advice and support without which this project would not have been possible. Also I would like to acknowledge the assistance of Alberto Sicilia and Fabien Paillusson who have worked on this project with me. Many thanks to David Wales, Jake Stevenson, Victor Rühle, Thomas Stecher and the Frenkel Group (plus extras) for their suggestions and collaboration. I would like to express my gratitude towards the people in the Theory Sector (RIG?) for their good vibrations. I am especially grateful to Sue Harding for her kind and constant support, for always "making things happen".

None of this would have been possible without the never-ending help from my family, in Chile and the UK, especially my wife Johanna. Finally I would like to acknowledge the generous financial support from Becas Chile CONICYT, Gobierno de Chile and St. Edmund's College, Cambridge.

Contents

Summary iii

Declaration v

Acknowledgements vii

List of acronyms and abbreviatons xiii

1 Introduction and motivation 1

2 The state of the art: literature review 7

2.1 Experimental work 7
2.2 Computer simulations 10
2.3 Edwards' theory 11
2.4 Defining random close packing 14
2.5 The jamming transition and phase diagram 15
2.6 Mean-field theory and glasses 21
2.7 Final remarks 22

3 Computer simulations and methods 23

3.1 Basins of attraction 23
3.2 Monte Carlo 24
3.3 Periodic boundary conditions 26
3.4 Parallel tempering 26
3.5 Thermodynamic integration 28

4 The basin volume method 31

4.1 Introduction 31
4.2 The volume calculation 32

4.3 Details of the model 36
4.3.1 Soft spheres 36
4.3.2 Gauss-Lobatto integration 37
4.3.3 Volume calculation: the procedure 38
4.3.4 Other technical details 40
4.4 Results and discussion 40
4.5 Conclusions 44

5 Defining basins with different minimisers 47
5.1 Introduction 48
5.2 Methods 50
5.3 Results and Discussion 54
5.4 Conclusions 59

6 Calculating the entropy 63
6.1 Details of the model 64
6.1.1 Hard cores with strongly repulsive shell 65
6.1.2 Floaters or rattlers 68
6.2 Results 69
6.2.1 Behaviour of the mean square displacement 69
6.2.2 Behaviour of $\langle u^2 \rangle$ vs k 72
6.2.3 Free energy distributions 72
6.2.4 The bias problem 74
6.2.5 Fitting the free energy distribution 78
6.2.6 Bootstrapping 80
6.2.7 Configurational entropy 80
6.3 Conclusions 87

7 Jamming density distribution and compactivity 89
7.1 Introduction 89
7.2 Finding the jamming density 93
7.3 Results and discussion 95
7.4 Conclusions 99

8 General conclusions and outlook 101

A Minimisation algorithms **105**
A.1 Line search 106
A.2 Steepest descent 106
A.3 Conjugate gradient 107
A.4 BFGS and L-BFGS 108
A.5 FIRE 109
A.5.1 Original FIRE 109
A.5.2 Modified FIRE 110

B Centre of mass constraint in free energy calculation. **113**

Bibliography **117**

List of acronyms and abbreviatons

BFGS Broyden, Fletcher, Goldfarb and Shanno. 50, 51, 55, 58, 59, 61, 108

CG Conjugate Gradient. 50, 51, 55, 58, 61, 102, 107

FCC Face-Centred Cubic. 3, 4

FIRE Fast Inertial Relaxation Engine. 48, 50, 51, 55, 57–59, 61, 65, 67, 94, 101, 105, 109–111

GGD Generalized Gaussian Distribution. 78

GL Gauss-Lobatto. 31, 33–35, 37–40, 69, 72

HCP Hexagonal Close Packed. 3, 4

HS Hard Sphere. 65–67, 94

HS+WCA Hard Sphere core plus Weeks-Chandler-Andersen. 66, 69, 70, 72, 83, 94, 99, 110

L-BFGS Limited memory BFGS. 40–42, 44, 45, 48, 50, 51, 55, 56, 58–62, 102, 108

LJ Lennard-Jones. 51, 53, 54, 59, 61, 65

LS Line Search. 105–108

MC Monte Carlo. 1, 19, 23–28, 30, 33, 35, 38–40, 44, 47, 67, 70, 72, 80

MD Molecular Dynamics. 10, 13, 17, 19, 23

MRJ Maximally Random Jammed. 14

PBC Periodic Boundary Conditions. 26, 36, 115

PEL Potential Energy Landscape. 1, 17, 19, 23, 30–32, 40, 41, 47

PT Parallel Tempering. 23, 26, 27, 39

RCP Random Close Packing. 4–8, 14, 15, 17, 22

SD Steepest Descent. 40–42, 44, 45, 48–50, 54, 55, 59–61, 106, 107, 110

WCA Weeks-Chandler-Andersen. 65, 66

Chapter 1

Introduction and motivation

In this work the configurational entropy of random jammed systems is calculated using Monte Carlo (MC) exploration of the Potential Energy Landscape (PEL). This study sheds light on the nature of this entropy and reveals some important characteristics of jamming phenomena in systems of soft repulsive particles. These model systems can be used to explore many of the physical properties of dispersed materials. Some examples of such material are emulsions, gels, foams, powders and colloids. Many of these systems can become “jammed” in an amorphous state at some point and lose the capability to respond to external forces due to changes in temperature, pressure and load. Roughly speaking, glasses are liquids that are cooled rapidly to form an amorphous solid; foams are jammed bubbles in a rigid structure; granular materials are non-thermal systems that can behave like solids, liquids or gases depending on the external conditions; colloids are mixtures where a substance is dispersed in another and may be solid, liquid or gaseous.

Granular materials are of particular interest as they are the second most manipulated material in industry after water and can be found in the form of powders in pharmaceutical, food, mechanosynthesis and semiconductor industries [76]. In fact, some industrial plants spend as much as 60% of their energy budget on handling and processing these materials[38]. Although the jamming transition may appear to be similar to a liquid-solid phase transition, it is in fact very different as the molecular structure of the material and the density do not change abruptly as they would in a first order phase transition. There is some speculation regarding the glass and jamming transitions and how they relate to each other and if they are merely the same physical phe-

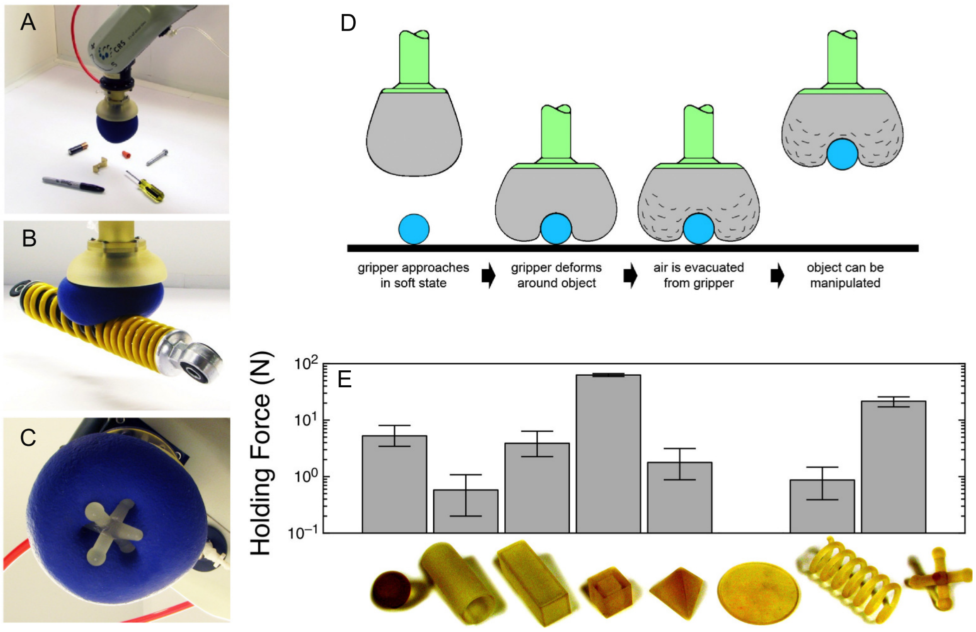


Figure 1.1: Universal gripper from [11]. (A) shows the gripper attached to a robotic arm. (B) shows the gripper lifting a shock-absorber coil. (C) shows the view from the underside. (D) shows a schematic of how the gripper works. (E) is a plot of the holding force for different objects. Figure from arXiv:1009.4444 [cond-mat.soft].

nomenon. Much work has been done and a general consensus has not been reached, although it has been shown, within a mean-field model, that they are different [57]. In the case of granular media, an inherently athermal system, the glass transition has no place.

In recent years researchers have used the jamming transition – the fact that grains become immobilised when their number density is increased – to their advantage and have created a *universal robotic gripper* [11]. The design of this gripper is not based on a fingered hand like most other models. Instead it uses granular media (in this case ground coffee) inside a flexible container (balloon) such that the air inside can be evacuated thus reducing the volume and making the system jammed. The way that it works is by starting in an unjammed state where the grains are free to move around inside the balloon. Then the gripper approaches the object that it wants to hold such that the object is surrounded by the balloon which deforms around the object, as in Figure 1.1. Then the air is evacuated, the grains become jammed and the object can be manipulated. The key feature of this gripper is that it is able to grip objects that may differ considerably in shape, size and weight without

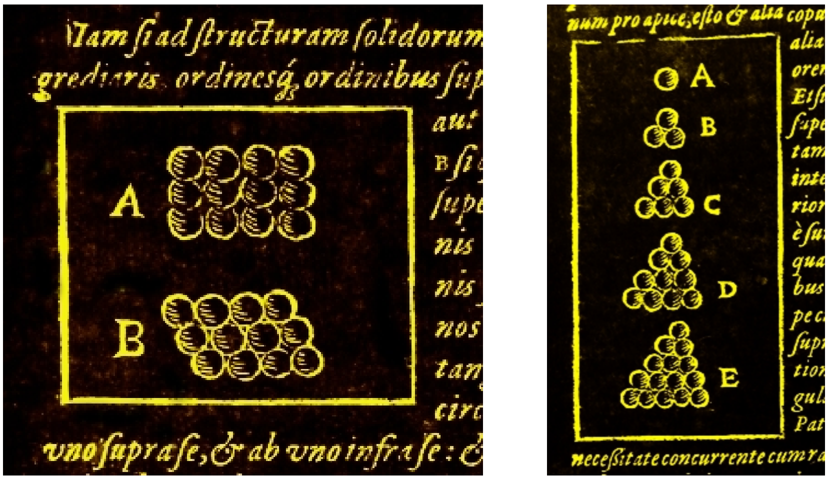


Figure 1.2: Regular close packing diagrams from Kepler's 1611 book *De nive sexangula* ("On the six-sided snowflake"). Images from [16].

prior knowledge of the kind of object that is to be manipulated without needing sensory feedback. As can be seen in Figure 1.1, panel E, the resulting force applied to different objects is different.

The packing density, or volume fraction, ϕ of a jammed system is defined as the fraction of space occupied by the particles. In the year 1611, after being asked how to stack cannon balls efficiently, German mathematician Johannes Kepler proposed that the densest 3D packing of equally sized spheres occupies $\phi = \pi/\sqrt{18} \approx 74\%$ of the total volume (see Figure 1.2). In this case the spheres form a Face-Centred Cubic (FCC) or Hexagonal Close Packed (HCP) lattice (see Figure 1.3). The hypothesis that a regular stacking of close-packed triangular layers forms the densest possible packing of spheres is known as the *Kepler conjecture* and proving it has been far from trivial. In 1831 Gauss was able to prove that the conjecture is true for spheres on a regular lattice and in 1953 Hungarian mathematician Tóth proposed that the problem could be solved by doing a finite but very large number of calculations [27, 88]. This proposition was pursued by Thomas Hales who started his work in 1992. By 1998 he had finished his proof [35] which is almost certainly correct but very complex as it took a dozen referees of *Annals of Mathematics* about four years to be "99% certain that the proof was sound" [88]. Despite the lack of a simple 100%-certain proof, Kepler's conjecture is widely accepted as being true.

What about the case of *random* packing? Is there an equivalent to Kepler's conjecture? The interest in Random Close Packing (RCP) started with Stephen Hales (1677-1761) who studied the random packing of green peas in his book *Vegetable Staticks*. Bernal and Mason's experiments on ball bearings in 1960 [8], designed to study liquid configurations, resulted in a quantitative description of RCP. They suggested that the maximum packing density for a RCP state is $\phi = 0.64$ and that this value must have a mathematical explanation:

> The figure for the occupied volume of random close packing–0.64–must be mathematically determinable, although so far as we know undetermined. [8]

Bernal and Mason define RCP as the densest possible disordered state. The idea is that to achieve higher densities, regions of ordered structures must be present in the system. Scott [78], on the other hand, finds that the same RCP density corresponds to the densest possible state obtainable via shaking ball bearings. These results have been extensively tested by others both experimentally [8, 45, 72] and using simulations [67, 2, 55, 70] where the consensus is that the RCP packing fraction is $\phi \approx 0.64$. Despite this, RCP remains mathematically undetermined.

Now some interesting questions arise: What happens if we pour spheres into a container? How do they pack? What kind of structure do they form? What is their packing density? How many distinct packings exist? These are the type of questions that must be answered to understand RCP.

If spheres are poured into a container and pressed down they will most certainly not be in a regular packed FCC or HCP state. This was already known quite a while ago and probably the first mention of the concept of RCP is in the Bible:

> Give, and it will be given to you. A good measure, **pressed down, shaken together and running over**, will be poured into your lap. For with the measure you use, it will be measured to you. (*Luke 6:38*)

This means that grains that have been pressed down and shaken until they cannot be compressed anymore have a unique packing fraction. Or at least this is what is suggested and needs to be investigated.

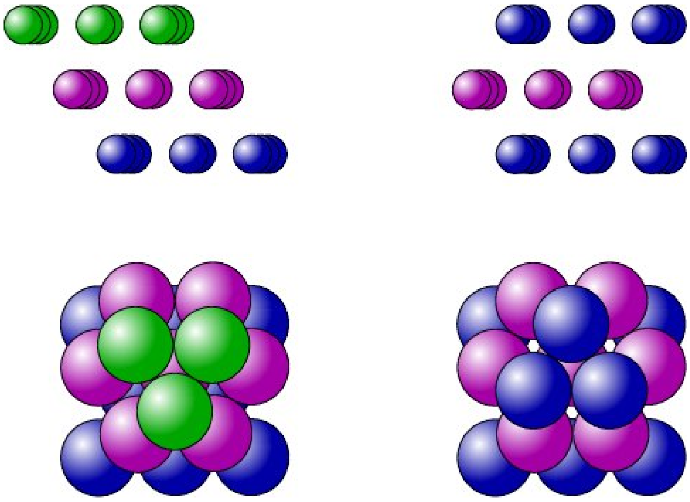


Figure 1.3: Regular close packing. FCC (left) and HCP (right) lattices. Image from [106].

The statistical mechanics of powders was introduced by Edwards and Oakeshott in 1989 [24] and states that the volume of a granular system plays a role equivalent to the energy of a thermal system and that the "compactivity" is equivalent to the temperature. Their hypothesis is that all states (or jammed configurations) are equally probable and that there is a well-defined entropy which "corresponds to the logarithm of number of ways the grains can be assembled to fill a volume V" [25]. This is sometimes known as the *Edwards conjecture*. This hypothesis has never been tested directly for a realistic model of granular materials.

An experimental system of friction-less, attractive emulsion droplets has recently been studied by Jorjadze and coworkers [41]. This system is of particular relevance because it is a real system that is surprisingly similar to the systems studied in this work using simulations. Joriadze et al. were able to measure the decrease in compactivity with global density. This is an example where Edwards' statistical theory can successfully be applied to experimental jammed systems.

It has been suggested that the concept of RCP cannot be properly (mathematically) defined. Torquato *et al.* [89] state that the packing fraction depends on the compression rate and that a denser state can always be found by waiting longer. This happens because the lower the compression rate the more time the system has to adjust itself, which produces regions with local order.

In their 2003 paper [67], O'Hern *et al.* studied jamming of two and three dimensional particles that interact via a repulsive, finite range potential (soft spheres) at zero temperature and applied stress. They found that as the number of particles increases, the jamming transition becomes sharper. This occurs at a well-defined density that happens to be the RCP packing fraction discovered by Bernal and Mason, $\phi \approx 0.64$ (see Figure 2.6). This means that RCP can be well defined. It is the point where, in the thermodynamic limit, the distribution of jamming densities has a discontinuous jump. That is to say that the RCP packing fraction is a property of three dimensional space.

The main results of this thesis are presented in Chapter 6 which describes a novel approach to measure the distribution of volumes of basins of attraction and ultimately the configurational entropy based on the techniques described in Chapter 4. The improvements to this method are discussed. The rest of this work is organised as follows: In Chapter 2 a review of the state of the art of RCP and jamming is presented including where this work fits into the current framework. Chapter 3 contains a description of the computational methods used to develop this work. Chapter 4 describes the method used to measure the volumes of the basins of attraction and some important caveats. In Chapter 5 the definition of basins of attraction with different minimisation algorithms is studied. In Chapter 7 the distribution of densities at which our system becomes jammed is studied. The general conclusions and final remarks of this work are presented in Chapter 8 along with a brief outlook. Finally in Appendix A an in-depth description of minimisation algorithms is given and in Appendix B the details regarding the centre of mass constraint in our simulations is discussed.

Chapter 2

The state of the art: literature review

This Chapter provides the context for subsequent chapters by summarising some of the relevant findings of experimental, numerical and theoretical studies of Random Close Packing (RCP) and the jamming transition. We will then discuss some of the open questions in the field and this allows us to explain how the work presented in the present thesis fits into the broader context of understanding the physical properties of granular materials.

First, a brief account of one of the earliest packing experiments will be given followed by a description of the pioneering work of Bernal and Mason. Computer simulation results will also be discussed, including the seminal work by Stillinger and Weber, who introduced the concept of *inherent structure* and the Lubachevsky-Stillinger algorithm to find jammed states. Edwards' statistical mechanics for granular media and the so-called *Edwards' hypothesis* are introduced. The latter hypothesis has been hotly debated within the community and its predictions have been tested for many different systems in experiments, simulations and theory. The results of these tests are inconclusive and it appears that the validity of the Edwards hypothesis is at the very least protocol dependent.

2.1 Experimental work

One of the earliest experiments on soft spheres was performed by the Reverend Stephen Hales and accounted for in his 1727 book *Vegetable Staticks:*

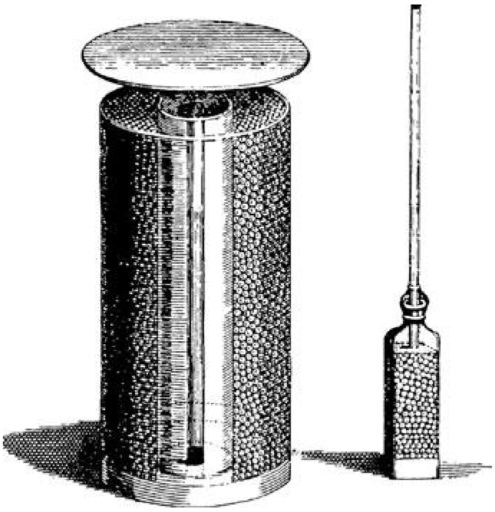


Figure 2.1: Experimental setup used by Stephen Hales, *Vegetable Staticks*, to study jammed green peas. Image from [103].

Or An Account of Some Statistical Experiments on the Sap Vegetables. Hales studied a system of green peas under pressure, Figure 2.1. He added water to dilate the peas and obtain a jammed state. The conclusion to this experiment was that the peas did not raise the weights holding them down, instead they filled the interstices and deformed into dodecahedrons [103]. Hales was wrong about the dodecahedrons as closely packed soft objects form irregular polyhedra [103].

It was not until 1960, and the experiments of Bernal and Mason [8] on the structure of liquids, that a quantitative study of RCP was performed. Bernal and Mason investigated the structures formed by spheres in a balloon (see Figure 2.2) using steel ball bearings. Ink was subsequently poured into the balloon and left to dry. This made it possible to perform an analysis of the contacts of each particle. The RCP volume fraction, 0.64, was defined as the densest possible random packing. Bernal was sure that it must be "mathematically determinable" [8] as seen in the quote in Chapter 1. This remark turned out to be somewhat of an understatement as it is still an unsettled issue today.

Another experimental investigation of RCP was published in the same issue of *Nature* as Bernal and Mason's work. Scott *et al.* [78] study an experimental system of steel ball bearings in different containers that were found to become jammed between the densities 0.63, dense random packing and 0.59, loose random packing. The dense random packing limit was obtained

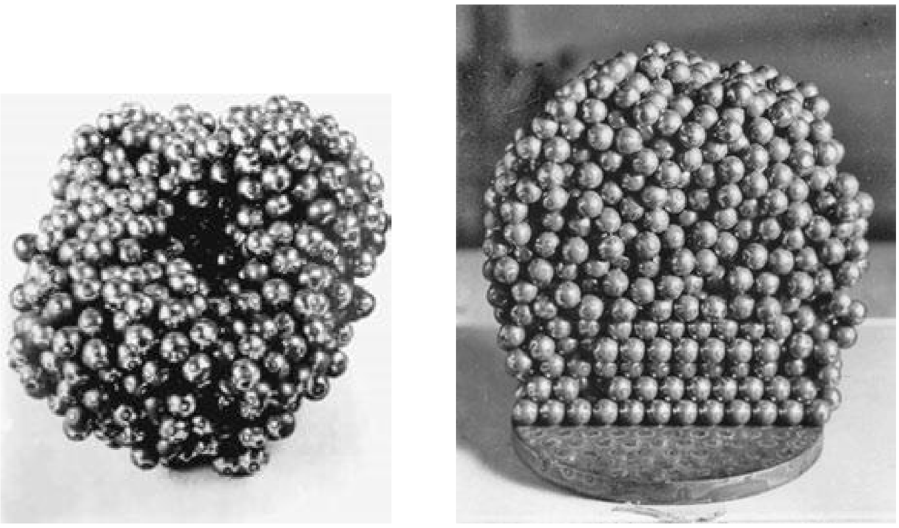


Figure 2.2: Bernal and Mason's experiments [8]. Image from [103]. The right hand image shows the partial crystallisation of the system near the borders in a cubic container.

by shaking the system. The loose packing limit is the result of pouring the particles into the container. At both limits the system was rigid, it could withstand pressure.

Knight *et al.* [45] studied a system of mono-disperse spherical glass particles (~ 2 [mm] diameter) jammed in a long thin cylinder (~ 1.88 [cm] diameter). The system started at a low packing density and was vibrated until the steady state density was found. The final packing fraction depends on the intensity of the vibrations, i.e. it is protocol dependent. Density relaxation was also studied and was found to be complex even for this simple system and depends on the size of the particles and the diameter of the cylinder.

The experiments of Philippe and Bideau [72] on vertically vibrated glass beads show that the evolution of the mean packing fraction and mean potential energy exhibit a stretched exponential law and that the compaction time follows an Arrhenius relation with the tapping intensity. The observed behaviour is in many ways reminiscent of the relaxation of glassy systems. The experiments reported in [72] studied glass spheres of 1 [mm] diameter in a cylinder of diameter 10 [cm] filled about 10 [cm] high. Stable packings were observed with volume fractions between 0.583 and 0.632 corresponding to the loose and dense packed limits found by Scott [78], mentioned above. The compaction or relaxation time of tapped granular matter has been discussed at length in [76] as there is some controversy whether the long-time relaxation is exponential or logarithmic.

2.2 Computer simulations

In 1972 one of the first simulations of jammed systems was performed by Adams and Matheson [2]. Closely packed configurations of up to 5402 hard spheres were produced and a packing density of $\phi \approx 0.628$ was found. Their method consisted of starting with a core of spheres already randomly packed and then adding spheres at specifically chosen sites. The result of this method is a spherical system of randomly packed hard spheres. The pair distribution function was computed using these results and compared with the ball bearing experiments. The computer simulated model was seen as a useful alternative to experiments because it is much easier to determine the positions of the particles and analyse the structure of the jammed state.

The concept of *inherent structure* was introduced by Stillinger and Weber in 1983 [84]. In their paper these authors define an inherent structure as a stable particle packing or potential energy minimum that can be reached by steepest descent. The dynamics of the transition between different minima was studied using Molecular Dynamics (MD). The fluid system evolves and at a given moment is quenched to the corresponding local energy minimum. Thereby the distribution of stable packings is determined with respect to their potential energy. A representation of an energy landscape showing some of its features including inherent structures and basins of attraction can be seen in Figure 2.3 from [85].

The simulations of Lubachevsky *et al.* [55] consider systems of up to 2000 disks and up to 8000 spheres. Their algorithm for obtaining jammed states starts with random initial positions and random velocities uniformly distributed between -1 and $+1$. As the system evolves via event-driven MD the particles grow monotonically until the system becomes jammed. The final state depends on the growth rate of the particles. This technique for generating jammed packings is often referred to as the *Lubachevsky-Stillinger* protocol. For jammed states found using this algorithm, it is possible to conclude that mono-disperse two dimensional systems, even for high growth rates, present short to intermediate-range order. This is not present in three dimensional systems. It is suggested that the number of jammed states has

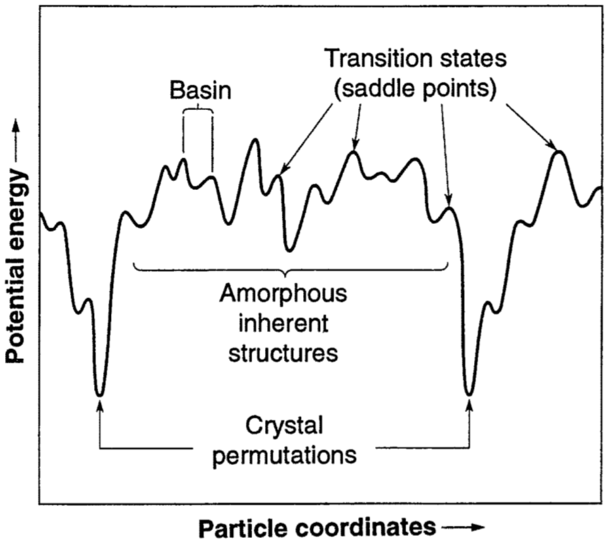


Figure 2.3: Representation of an energy landscape from [85] showing some of its features, basins of attraction of inherent structures are of particular importance for this work.

the asymptotically exponential form

$$S = \ln \Omega(N, d) = \alpha(d)N \qquad\qquad N \to \infty \tag{2.1}$$

were Ω is the number of distinct packings that depends on the number of particles N and dimension d. The fact that no order is present in three dimensional systems suggests that $\alpha(1) < \alpha(2) < \alpha(3) \cdots$. For three dimensional systems the packing fraction was found to be $0.63 \leq \phi \leq 0.65$ and for two dimensions $0.85 \leq \phi \leq 0.90$. The distribution of contacts between particles was studied and the mean contact number in three dimensions found to be 5.8295 for a mono-disperse system of 8000 particles, close to the isostatic condition discussed later in this Chapter. The exponential growth of the number of inherent structures with system size was further discussed by Stillinger [83] and is related to a key result of this work presented in Chapter 6.

2.3 Edwards' theory

In their ground-breaking paper *Theory of Powders* published in 1989, Edwards and Oakeshott [24] propose a "thermodynamics" for granular media. Specifically, the authors apply the concepts of statistical mechanics to granular materials. The volume of a powder plays the role of the energy in conventional statistical mechanics while the "compactivity", X replaces the temperature. The fundamental postulate of statistical mechanics says that an isolated system in equilibrium is equally likely to be in any of its accessible microstates. In Edwards' theory this is replaced by the assumption that all packings are equally likely when the system is at a fixed volume. For jammed structures prepared by energy minimisation, the validity of this assumption would imply that the distribution of volumes of the basins of attraction should be very sharply peaked meaning that most basins have similar volume.

In Edwards' statistical-mechanics approach, described above, all blocked configurations of grains in a given volume are considered to be equally likely, i.e. flat ensemble averages are considered. What has come to be known as *Edwards' hypothesis* is the fact that "all the metastable states in which a system can be trapped are equivalent for the dynamics" [15]. It is important to note that this approach is not justified from first principles and must be tested. Edwards' ideas can be applied to other systems with large numbers of metastable states [15] and it makes sense to consider glassy systems where the underlying energy landscape seems similar to that of granular systems. In the context of mean-field models and the study of ageing dynamics of glassy systems, it has been shown that the basins of attraction are explored with a flat measure [30] thus satisfying Edwards' hypothesis. The relevance of Edwards' measure has also been studied for zero-temperature spin-flip dynamics of ferromagnetic Ising models with blocked configurations [15]. In this thesis we will show that the energy-minimisation protocol does not lead to a uniform distribution over blocked configurations i.e. states are not sampled with Edwards' measure.

Barrat *et al.* [7] have shown that Edwards' hypothesis holds within the Kob-Andersen [46] model which is able to reproduce many aspects of glassy and granular systems and that it can be used to compute the thermodynamics. In this study all blocked configurations are treated as equivalent and Edwards' measure is generated explicitly using an "auxiliary model" in which particles have energy one if they would be allowed to move in the original model and zero otherwise. A method to count the blocked configurations is presented. The fact that Edwards' hypothesis only holds for certain protocols (or dynamics) is also discussed. The paper ends with a final remark that is worth including here:

> ...in the simplest cases the correspondence between Edwards' distribution and long-time dynamics is at best checked but does not follow from any principle. The situation is thus as if one would have checked that the microcanonical distribution gives good results for gases, without knowing Liouville's theorem that proves that such a distribution is indeed left invariant by equations of motion.

The work of Song *et al.* from 2005 [81] describes an experimental method to measure an effective temperature for jammed granular materials. The diffusion of tracer particles is measured in a slowly sheared granular system near the jamming point. The effective temperature is determined using the Einstein relation. This temperature is the variable that controls the way that the system explores different jammed configurations. The particles diffuse over distances "a few times" their diameter implying that they rearrange outside their cages [81]. Systems of different particles were found to equilibrate at the same temperature, given by their volume fraction.

In 2007, Brujić and co-workers measured the coordination number and entropy of a three dimensional jammed emulsion using confocal microscopy [13]. The contact network of the system was identified and the average contact number agrees with the isostatic condition described later in this Chapter (for frictionless, spherical particles on average six contacts in three dimensional space). Instead of calculating the entropy using the statistics of volumes, microscopic variables were used. The contact network and graph theory representations were used to calculate the Shannon entropy using the probability of observing a graph of a certain class i. This entropy is based on the probability to observe a jammed state with a certain topology and is an extensive property within the scope of their results.

Paillusson and Frenkel studied ergodicity in tapped granular systems using event-driven MD simulations of hard friction-less spheres under gravity [68]. Ergodicity is defined as the equivalence between "time" and ensemble averages. When comparing histograms generated by sampling long "time" trajectories to volume ensemble histograms non-ergodicity is observed. The overlapping histogram method is used to test Edwards' hypothesis within the scope of their system and found to hold over a limited range of packing densities.

Makse and Kurchan reported a simulation of sheared granular matter using a range of particle sizes [56] and were able to measure an effective temperature –via the Einstein relation– using the particles diffusivity and mobility showing that it coincides with Edwards' configurational temperature. In contrast, Ising model simulations show that the limiting distributions do not coincide with a Edwards' prediction of a uniform distribution over blocked configurations [15].

The experimental work of Lechenault and coworkers [48] describes the volume and free volume distributions of Voronoi cells for a system of vibrated, bi-disperse, jammed particles in two dimensions. Sub-extensive scaling of the free volume is found, suggesting the existence of correlations between the free volumes of the particles. Although this claim goes against Edwards' theory, an intensive compactivity can still be defined in a consistent way.

A recent experimental study on two-dimensional jammed configurations of photo-elastic disks was carried out by Pucket and Daniels [75]. The use of photo-elastic disks allows the forces to be measured at each contact and the *angoricity* to be calculated. The angoricity tensor –conjugate to the stress– is another analogue to temperature, defined by Blumenfeld and Edwards' [10]. Pucket and Daniels find that when putting two systems in contact with each other the compactivity does not equilibrate while the angoricity does.

2.4 Defining random close packing

Although the RCP problem has been studied for a long time, the lack of mathematical definition has stimulated the formulation of new theories and concepts. Torquato *et al.* [89] suggest that the concept of RCP is not well defined as it depends on the protocol used to obtain the packing. The idea is that the packing fraction can always be increased by introducing local order into the system and the new concept of Maximally Random Jammed (MRJ) states is proposed. Some results of these studies are shown in Figure 2.4 where we can see the reciprocal compression rate Γ^{-1} versus the packing fraction ϕ and a plot of the T-Q plane for hard spheres. T and Q are the translational and orientational order parameters respectively. It is clear that the packing fraction depends on the compression rate, i.e. the system shows

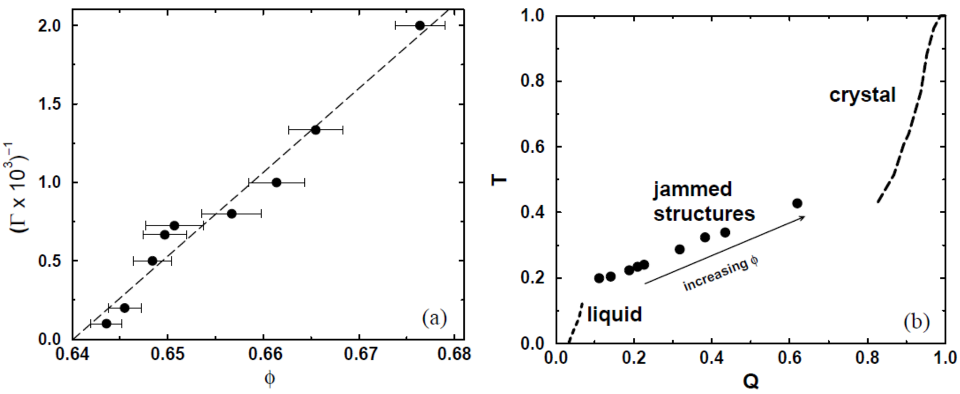


Figure 2.4: Results of Torquato *et al.* [89]. Figure (a) shows the reciprocal compression rate Γ^{-1} versus the packing fraction ϕ. (b) shows the T-Q plane for hard spheres. T and Q are the translational and orientational order parameters respectively.

protocol dependence. For a slower rate a denser packing is achieved.

Kamien and Liu [42] also propose that RCP can be defined in their study of the relation between the pressure and packing fraction of hard spheres. The rate at which states become forbidden is proportional to the pressure for increasing volume fraction. The pressure is found to diverge at a packing fraction $\phi_{\text{max}} = 0.640 \pm 0.006$ in good agreement with the RCP density.

The results of O'Hern and coworkers, presented in the next section, also suggest that the concept of RCP is well defined. At least within a certain protocol.

2.5 The jamming transition and phase diagram

Figure 2.5 is the speculative jamming phase diagram proposed by Liu and Nagel [52]. This schematic diagram suggests that different jamming phenomena are related and that a minimum density is necessary for a system to be jammed. It captures the observation that a system can become or cease to be jammed by changing one of the parameters, for example applying stress, raising the temperature or changing the density. Point J on the $1/$density axis at zero temperature and load is the volume fraction were athermal systems become jammed under these conditions. At jamming, the system is expected

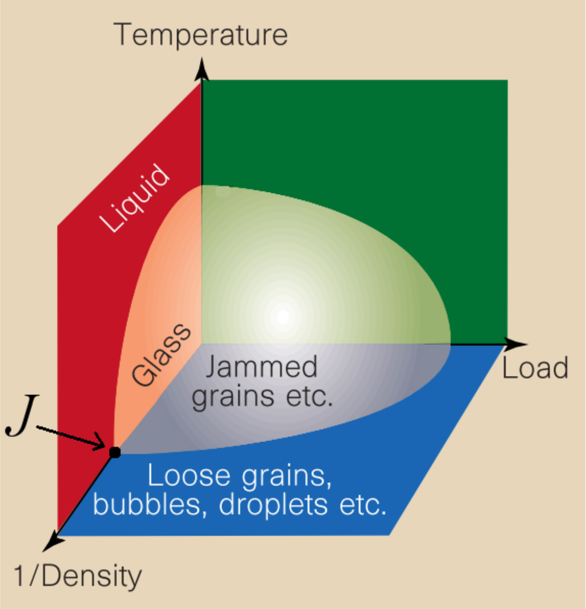


Figure 2.5: Jamming phase diagram proposed by Liu and Nagel from [52]. The J-point is shown and represents a system becoming jammed as its density is increased at zero temperature and load.

to be *isostatic*, as described in the next paragraph. Our research will probe systems near this point along the $1/$density axis.

In [3] Alexander describes the concept of *isostaticity* which occurs when a system has the same number of total contacts as force balance equations. This means that the system is at mechanical equilibrium. For a system of N frictionless soft spheres in dimension d, like the ones studied in this work, the number of degrees of freedom is Nd so the isostatic condition is $Z = 2d$ where Z is the average number of contacts per particle. The results of [67] state that the jamming transition coincides with the onset of isostaticity for all of the systems that they studied. The extension of these results to frictional spheres was performed by Henkes *et al.* [36]. Frictional spheres were found to not be isostatic at jamming unless a general form of isostaticity is introduced where fully mobilised contacts at the Coulomb friction threshold are treated as slipping contacts. In these systems with friction, the packing fraction and contact number at jamming are not unique and depend on the friction coefficient μ. The average contact number can vary between $d+1$ and $2d$ and it appears that the former is true only when $\mu \to \infty$ [92] and the packings are equilibrated very slowly. In most cases the number of contacts is larger than those needed for force balance and rigidity. This is called *hyperstaticity*.

O'Hern *et al.* [67] propose a method to simulate jammed states by quenching random initial states ($T = \infty$) to their corresponding potential energy minima ($T = 0$) using the conjugate gradient method. By starting with random initial states all phase space is sampled uniformly. A similar procedure had earlier been used by Stillinger and Weber in a different context [84] to find inherent structures. O'Hern *et al.* studied jamming of disks and spheres. The interaction between the particles is a repulsive, finite range potential (soft spheres, see Equation (4.19)). Systems near point J in the jamming phase diagram (Figure 2.5) that is to say at zero temperature and stress were investigated. At a given packing fraction (density), the energy minimisation can have two different outcomes: a minimum with no overlaps and zero energy or a minimum with one or more overlaps and positive energy. The latter is a forbidden configuration for a hard sphere system. O'Hern and coworkers calculated the distribution of the jamming density $P_j(\phi_c)$ for different system sizes in two and three dimensions (see Figure 2.6). As the system size increases, the jamming transition becomes sharper. The peak approaches a well defined density, namely $\phi_{\text{RCP}} = 0.64$ where the potential energy minima of most initial ideal gas states become forbidden (energy $\neq 0$). It is also interesting to note that these results do not depend strongly on the form of the continuous potential used to find the inherent structures. These results also suggest that RCP is a well-defined concept which is related to the discussion presented in Section 2.4.

Xu *et al.* [109] used a similar method to study small two dimensional systems. Instead of letting the system evolve via MD time steps, they successively compressed or expanded the soft particles and minimised the potential energy until a jammed state was found. The advantage of studying small systems is that it is possible to find and count all of the distinct packings. This method is often referred to as *direct enumeration*. The probability distribution of finding a RCP state at a certain packing fraction was decomposed into two parts: the density of jammed states and their frequency distribution. The latter depends on the protocol used to obtain the jammed state while the density does not. The number of distinct jammed states was found to grow exponentially with system size.

In 2011, Xu *et al.* proposed a method, used in this work, to directly determine sizes of basins of attraction on the Potential Energy Landscape (PEL)

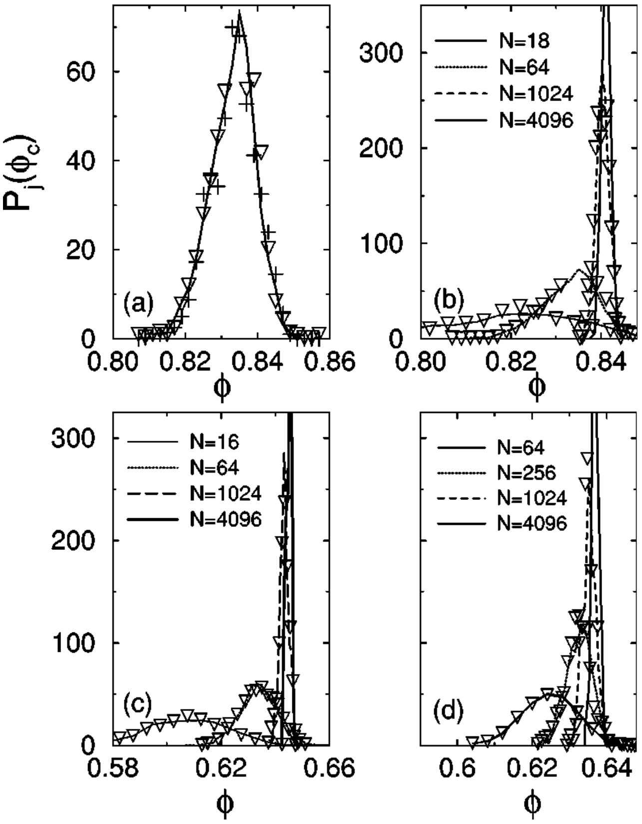


Figure 2.6: Figure illustrating the distribution of the jamming density $P_j(\phi_c)$ for different system sizes in two and three dimensions from [67]. From left to right: Top row: two dimensional bi-disperse system of 64 particles using different repulsive potentials ($\alpha = 2, 3/2, 5/2$ see Equation (4.19)), two dimensional bi-disperse particles ($\alpha = 2, 5/2$). Bottom row: three dimensional bi-disperse systems ($\alpha = 2, 5/2$) and three dimensional mono-disperse systems ($\alpha = 2, 5/2$).

via Monte Carlo (MC) sampling and thermodynamic integration [110]. The validation of the method comes from comparing the results with those found in previous studies using the direct enumeration method [109], described in the previous paragraph, where all basins of attraction (or inherent structures) must be found. This is only possible for small systems of up to ~ 10 particles. The entropy, defined as the logarithm of the number of distinct jammed states, can also be computed if the assumption of a Gaussian distribution for the logarithm of the basin volumes is made. This method is described in Chapter 4 as it is the foundation of this work.

A different study of basins of attraction has been presented by Ashwin and coworkers [5]. Here the focus of the study is on the *density landscape* instead of the PEL described above. The volume of the system is fixed so that a point in configuration space corresponds to a certain density, much like in the case of a PEL, were a point in configuration space has a certain energy. The basins of attraction on this landscape consist of all of the zero-density points in configuration space that end up at a certain mechanically stable packing (minima on the $1/\text{density}$ landscape) via a certain protocol. The basins have a highly complex structure consisting of a spherical core surrounded by long thin tentacle-like structures, as shown in Figure 2.7. Most of the volume of a basin is contained in these outer structures, not in the core, which means that the volume of the core is not strongly correlated to the probability of the packing.

Zhang *et al.* also studied the jamming transition of emulsions and granular materials using MD [112]. The model for emulsions consists of frictionless spheres interacting via potentials using experimental parameters. The grains are modelled as deformable spheres with friction. The results suggest that the jamming transition is characterised by power-law scaling of the pressure and number of excess contacts as the critical jamming density is approached from above. The critical jamming density is found to be protocol dependent, as noted in other studies.

In 2008, another jamming phase diagram was presented by Song, Wang and Makse [82]. Here the equation of state is calculated in terms of the average number of contacts Z assuming that all states are equally likely. In the limit where the compactivity goes to zero, $X \to 0$, the "ground state of jammed matter" the packing fraction is $\phi_{\text{RCP}} \approx 0.634$, random close packing. In the $X \to \infty$ limit the system approaches random loose packing $\phi_{\text{RLP}} \approx$

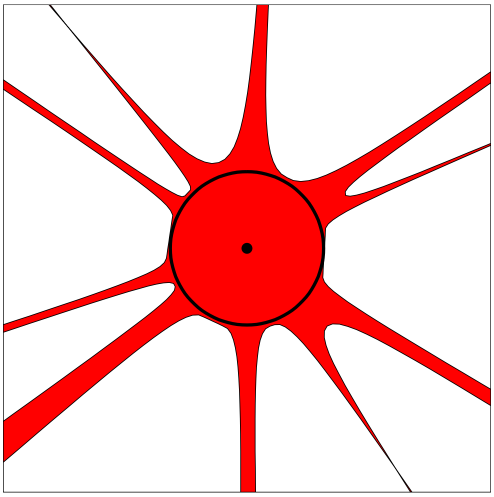


Figure 2.7: Schematic drawing of the shape of a basin of attraction on the density landscape [5]. The spherical core is the region contained by the black circle and the black point is the location of the corresponding mechanically stable packing. Figure modified by the author from: arXiv:1112.4234 [cond-mat.soft].

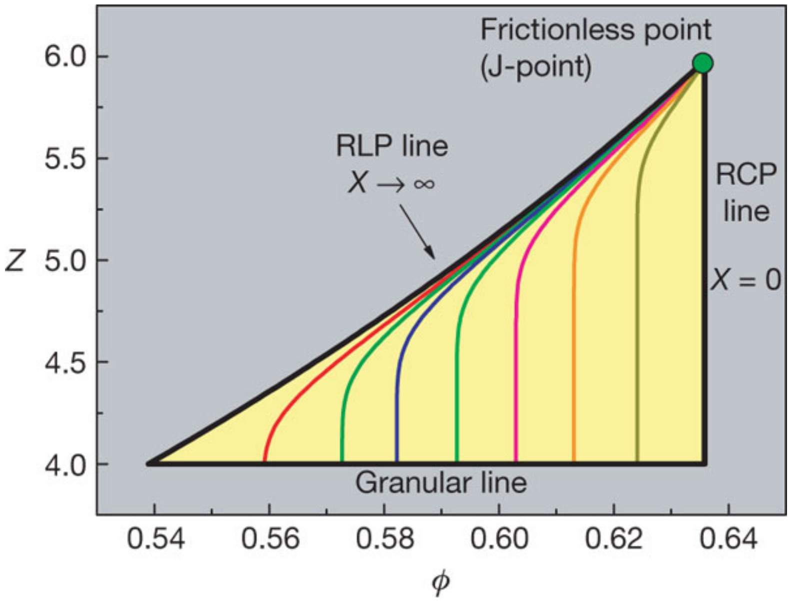


Figure 2.8: Jamming phase diagram from [82]. All disordered packing are within the triangle. The coloured lines represent iso-compactivity lines.

$Z/(Z+2\sqrt{3})$. The complete phase diagram is shown in Figure 2.8. Here the average contact number Z goes between the isostatic limit of no friction or J-point, described above, $Z=6$ and the infinite friction limit where $Z=4$.

The work of Chaudhuri, Berthier and Sastry, published in 2010 [17], shows that "jamming transitions in amorphous packings of frictionless spheres occur over a continuous range of volume fractions". The concept of "J-line" is introduced and represents the infinitely many J-points for different protocols (it has already been shown in this Chapter that for different compression rates different jamming densities are obtained). Although the findings show that the J-point is not unique, its critical properties remain so. In this sense it is not surprising that different experiments yield different results for the jamming density.

2.6 Mean-field theory and glasses

Many advances have been made in terms of mean-field theory of glasses in recent years and some of these results can be related to the jamming transition and the J-point introduced above. The glass transition is said to be related to "the appearance of many metastable glassy states in addition to the ideal glass one" [69]. These metastable states only appear above a certain density ϕ_d. For an overview of this topic refer to the review by Parisi and Zamponi [69]. Here the authors base their results on the assumption that "amorphous jammed packings of hard spheres can be identified with the infinite pressure limit of glassy states". Their results suggest that amorphous jammed packings exist in an interval of densities that depends on the protocol used to generate the packings, as previously observed by Krzakala and Kurchan where they position the so called J-point, described above, in the context of glass theory [47]. The random models (SAT and colouring) that were studied allowed them to ignore crystallisation and focus on the glassy characteristics. In the Lubachevsky-Stillinger type protocol, described above, the compression rate γ is fixed beforehand and the system equilibrates at a density ϕ_g where the relaxation time is of the order of γ. At this density the system responds to the compression by reducing the cages up to the jamming density $\phi_J(\gamma)$. ϕ_{th} is the value of the jamming density of the states that first appears at ϕ_d [69]. In mean-field theory it is possible to obtain jammed configurations in a whole range of densities $\phi_{th} \leq \phi \leq \phi_J$ [47, 111].

2.7 Final remarks

This Chapter has given a brief overview of recent studies (experiments, simulations and theories) of RCP and the jamming transition. In summary it seems that, in the context of an energy-landscape approach, RCP can be well defined: it is the point where, in the thermodynamic limit, there is a discontinuous jump in the number of allowed inherent structures. The location of this jamming transition does not seem to depend on the form of the continuous potential used to find these inherent structures as long as the protocol remains the same. The jamming transition is also related to the average contact number of the system.

The present thesis aims to shed light on the properties of jammed systems by extending the work of Xu *et al.* [110] described in Section 2.5. We show that it is possible to extend the approach of ref [110] to much larger systems of over 100 particles. Using this novel approach, it is possible to study the system-size dependence of the number of distinct packings of a system of poly-disperse soft disks. The simulations presented here provide strong evidence that it is possible to define a packing entropy that is extensive. This extensivity has often been assumed but, thus far, direct evidence was lacking.

Chapter 3

Computer simulations and methods

Molecular simulations can be separated into two broad classes: Molecular Dynamics (MD) and Monte Carlo (MC). MD uses a numerical method to integrate the equations of motion of a system. The result is a trajectory in time of the whole system. MC uses computer generated random numbers to sample *configurations* in phase space which are generally used to compute averages of observables.

In this section we will review those aspects of the simulation techniques that are of particular relevance for the research reported in this thesis. We use MC simulations to estimate the volume of the basins of attraction of the minima on the Potential Energy Landscape (PEL). These simulations require the determination of a free energy, a quantity that cannot be sampled directly in either MC or MD simulations due to the fact that the free energy of a system depends one the phase space volume accessible to it. Below we explain the MC technique that we use, including a scheme for improving sampling called Parallel Tempering (PT). We will also discuss the choice of the boundary conditions.

3.1 Basins of attraction

In this work, jammed states are studied by taking random initial conditions and minimising their interaction energy. Each distinct jammed state corresponds to a different minimum on the PEL (or “inherent structure” if we use

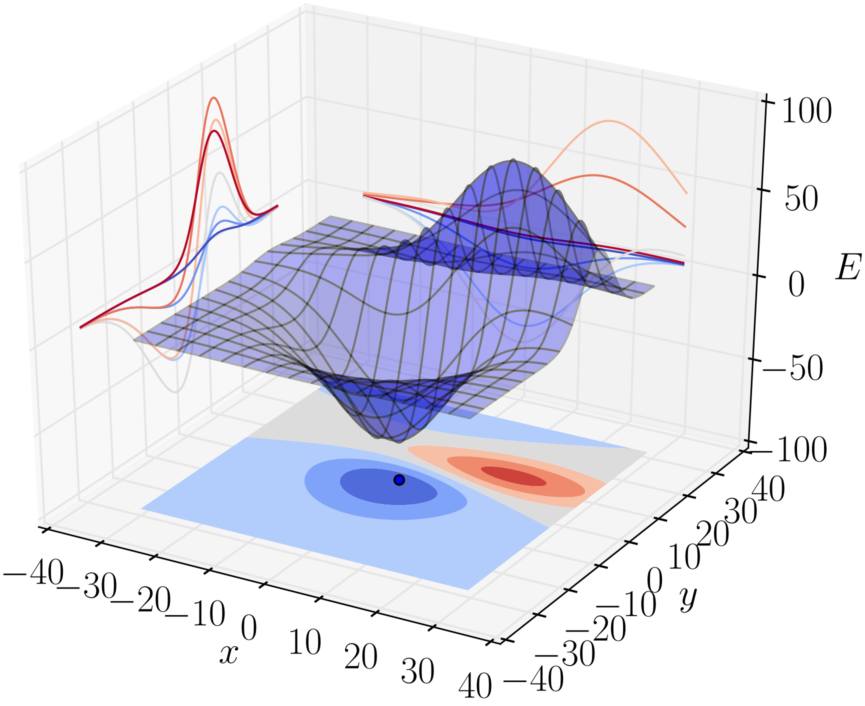


Figure 3.1: Example of a basin of attraction. This figure shows a three-dimensional plot of a function E with a minimum shown by the blue point on the xy plane. The contour plot on the xy plane is of the same function where the darker blue regions represent lower values of the function. The basin of attraction is the region delimited by the saddle points and maxima and ridges between them.

the language introduced by Stillinger and Weber [84]) and each minimum has its own basin of attraction. A basin of attraction is the "...volume of configuration space surrounding a local minimum from which steepest-descent paths converge to that minimum" [99]. This concept was introduced by James Clerk Maxwell around the year 1870 [99]. Examples of basins of attraction can be seen in Figures 3.1 and 2.3.

3.2 Monte Carlo

A MC simulation calculates averages of observables by sampling configuration space at random points. One of the simplest examples is a method for calculating π: Consider a circle of diameter d centred in a square of length

$l \geq d$. If the space is sampled randomly and the fraction of points that fall in the circle is measured, π can be calculated: $\pi = 4p$ (assuming $l = d = 1$ and p is the fraction of points that fell in the circle). Obviously the more points sampled the more accurate the result will be. If an infinite number of points could be sampled the result would be exact. This procedure is known as *random sampling* or *direct sampling*: probing configuration space by generating independent random configurations from a certain probability distribution. In the case of measuring π, discussed above, configurations are sampled from a uniform probability distribution. This however is not a good way to sample configuration space for most inter-molecular potentials because the Boltzmann factor varies sharply and in most cases is extremely small [32]. In fact for a system of 100 hard spheres at the freezing point the Boltzmann factor is nonzero for only $1/10^{260}$ configurations [32]. It is clear that it would be convenient to concentrate the sampling on states with a large Boltzmann factor, provided that we can correct for the bias in the sampling. A way to achieve this is described below and is known as *importance sampling*.

The MC importance sampling technique, introduced after the Second World War by Metropolis *et al.* [61] is known as Metropolis MC (in the rest of this text, we refer to this technique simply as MC). This method was developed at Los Alamos National Laboratory as computers became available for unclassified research.

As stated before, the main idea behind the Metropolis MC method is to sample the most probable states more than the least probable ones. If we can sample states with a probability that is proportional to their Boltzmann factor $\exp[-\beta U(\mathbf{r}^N)]$ where $\beta = 1/k_BT$ and k_B is the Boltzmann constant, T the temperature of the system and $U(\mathbf{r}^N)$ the potential energy as a function of particle positions, then an unweighted average over sampled points corresponds to a Boltzmann weighted average over configuration space. There are many ways to achieve this type of importance, the most common method was devised by Metropolis *et al.* and is described below [32].

1. Randomly select a particle and calculate the potential energy $U(\mathbf{r}^N)$ of the system.

2. Move the chosen particle by giving it a random displacement Δ: $r' = r + \Delta$.

3. Calculate the new potential energy of the system $U(\mathbf{r}'^N)$.
4. Accept the move with probability

$$\mathrm{acc}(\mathrm{o} \to \mathrm{n}) = \min\left(1, e^{-\beta[U(\mathbf{r}'^N)-U(\mathbf{r}^N)]}\right). \tag{3.1}$$

5. If the new move is accepted $\mathbf{r}^N = \mathbf{r}'^N$.
6. Go to step 1.

This procedure is repeated a large number of times (exactly how many depends on the system and the observables of interest) and at regular intervals all relevant quantities (for example positions, pressure, displacement, energy) are stored for future averaging.

3.3 Periodic boundary conditions

As the simulated systems are far from macroscopic, the choice of boundary conditions can significantly affect the measured properties. To illustrate this, imagine a three dimensional system of N particles. The fraction of all particles at the surface is proportional to $N^{-1/3}$ so for $N = 1000$, 49% are at the surface. Periodic Boundary Conditions (PBC) are used to repeat the simulation cell periodically. Interactions could in principle involve any periodic image of a particle. For short-ranged potentials is is convenient to use the so-called nearest-image convention: Each particle interacts with the nearest periodic image of the others in the simulation cell.

3.4 Parallel tempering

PT is a technique used to sample phase space more efficiently when performing MC simulations. It is useful for energy landscapes that have many local minima in which the system may be trapped. PT is implemented by doing MC runs of the same system at different values of a given parameter (most commonly temperature T). The system at the lowest temperature, T_1 will be the one most likely to get trapped in a deep basin, sampling only a small region of phase space [23]. The highest-temperature system, at T_M, will sample many regions of phase space. So the idea, introduced by Swendsen and

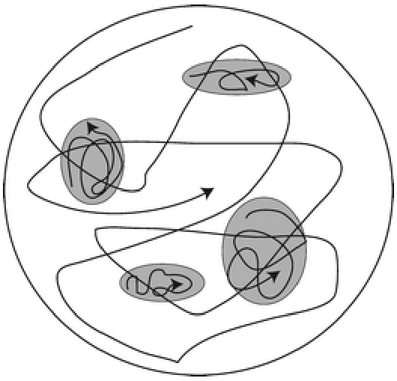


Figure 3.2: 2D representation of phase space from [23]. Low temperature systems are trapped near the energy minima (shaded areas) while high temperature systems sample many regions of phase space.

Wang [87], is to exchange configuration information between systems at T_1 and T_M so that the system at T_1 can sample a series of states near the minima (see Figure 3.2 for a schematic representation). A significant reduction in equilibration time is reported for low temperature systems meaning that less computer time is required.

The configurational swap move is performed with a probability similar to the probability of accepting a MC move (see Section 3.2, Equation (3.1)), ensuring that detailed balance is satisfied. Balance is satisfied if the swap is performed every fixed number of moves [23]. The probability of accepting a swap between systems i and j is:

$$\mathrm{acc}(i \leftrightarrow j) = \min\left(1, e^{(\beta_i-\beta_j)[U(\mathbf{r}_j^N)-U(\mathbf{r}_i^N)]}\right). \tag{3.2}$$

This means that if the temperature difference between systems i and j is large, the probability of accepting a trial swap of two configurations will be small. One solution for this problem is simulating intermediate temperatures such that in order of ascending temperature we have $T_1, T_2, \ldots, T_M$ where the system of interest is normally T_1. The choice of M and the values of the temperatures depend on the system to be studied. The distributions of energy must overlap so that the swaps have a good chance of being accepted. Normally only swaps between neighbouring systems are considered ($j = i+1$) but this is not necessary.

Some important characteristics of PT are:

- PT generates correct Boltzmann sampling of configuration space for each system i ($i = 1, \ldots, M$). [23, 32].

- It is easily parallelised: the energy of each system is known already from the MC calculations and communication only occurs if swaps are accepted. Even when swaps occur the communication is minimal as only the temperature is swapped rather than the entire configuration.

- Each system i at T_i is simultaneously undergoing a MC simulation and every certain number of MC moves a swap move is performed with the probability given by Equation (3.2). The systems are totally separate and do not interact energetically.

- Not only temperature can be used as the tempering parameter. In fact the calculations reported in this thesis did not use temperature, instead the spring force constant k was swapped. (this will be explained in Section 4.3).

3.5 Thermodynamic integration

One technique that can be used to calculate the free energy of a system is to find a reversible path that links a state of known free energy to the state in question. The change in F along this path is evaluated by thermodynamic integration, for example, by integrating the following equations:

$$\left(\frac{\partial F}{\partial V}\right)_{NT} = -P \tag{3.3}$$

$$\left(\frac{\partial F/T}{\partial 1/T}\right)_{VN} = E. \tag{3.4}$$

There are a limited number of convenient states of known free energy that can be used as a starting point. Two such states are the ideal gas and the low-temperature harmonic crystal. The system changes gradually from the initial (known free energy) to the final state by means of Kirkwood's coupling parameter λ [44]. In a system with potential energy U it is assumed that U depends linearly on λ such that $\lambda = 0$ corresponds to the reference system (I) and $\lambda = 1$ corresponds to the potential energy of the system being studied

(II). This can be written as:

$$\begin{aligned} U(\lambda) &= (1-\lambda)U_{\mathrm{I}} + \lambda U_{\mathrm{II}} && (3.5) \\ &= U_{\mathrm{I}} + \lambda(U_{\mathrm{II}} - U_{\mathrm{I}}) && (3.6) \end{aligned}$$

Here the potential energy of the reference system is recovered when $\lambda = 0$: $U(\lambda = 0) = U_{\mathrm{I}}$ and the potential of the system of interest is recovered when $\lambda = 1$: $U(\lambda = 1) = U_{\mathrm{II}}$. The partition function is given by:

$$Q(N, V, T, \lambda) = \frac{1}{\Lambda^{3N} N!} \int d\mathbf{r}^N \, \exp[-\beta U(\lambda)] \tag{3.7}$$

and the derivative of the Helmholtz free energy ($F = -k_B T \ln Q$) with respect to λ is:

$$\begin{aligned} \left(\frac{\partial F}{\partial \lambda}\right)_{N,V,T} &= -k_B T \frac{\partial}{\partial \lambda} \ln Q(N, V, T, \lambda) = -k_B T \frac{1}{Q} \frac{\partial Q}{\partial \lambda} && (3.8) \\ &= \frac{\int d\mathbf{r}^N \, (\partial U(\lambda)/\partial \lambda) \exp[-\beta U(\lambda)]}{\int d\mathbf{r}^N \, \exp[-\beta U(\lambda)]} && (3.9) \\ &= \left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_{\lambda} . && (3.10) \end{aligned}$$

The free energy difference can be obtained by integrating Equation (3.10) over λ from 0 to 1:

$$F(\lambda = 1) - F(\lambda = 0) = \int_{\lambda=0}^{\lambda=1} d\lambda \left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_{\lambda} . \tag{3.11}$$

So the free energy difference can now be calculated using an ensemble average that can be measured using simulations.

The Helmholtz free energy of a solid can be calculated via thermodynamic integration along a reversible path that transforms the solid into an Einstein crystal, where the particles are bound to their lattice sites by harmonic springs. When the spring constant is large, the solid resembles an Einstein crystal of known free energy [32].

The study of athermal jammed systems relies on calculating the free energy of solids that interact via discontinuous potentials, as described in Chapter 10 of [32], using a thermodynamic integration technique developed by Frenkel and Ladd [31] and later by Polson and Frenkel [73]. When studying hard-core interactions, defined by the potential U_0, it is not possible to construct a reversible path that transforms the system into a non-interacting Einstein crystal via a linear coupling scheme as was shown previously. A solution to this problem is to use a scheme where the the spring constants can be switched on while leaving the hard-core interactions unchanged [32]:

$$U(\lambda) = U_0 + \lambda U = U_0 + \lambda \sum_{i=1}^{N} (\mathbf{r}_i - \mathbf{r}_0, i)^2 \tag{3.12}$$

for a system of N particles where $\mathbf{r}_0, i$ is the position of particle i at its lattice or reference site. The free energy difference is then

$$F_{\text{HS}} = F(\lambda_{\text{max}}) - \int_0^{\lambda_{\text{max}}} d\lambda \, \langle U(\mathbf{r}^N, \lambda) \rangle_\lambda. \tag{3.13}$$

For high enough values of λ the system behaves like a non-interacting Einstein crystal. This is ensured by the choice of λ_{max}. When the particles are tethered to their lattice sites their is no translation of the system as a whole. On the other hand, when $\lambda \to 0$, the whole system is free to drift and the integrand in Equation (3.13) becomes sharply peaked at $\lambda = 0$. This presents a problem for numerical integration schemes as many values of λ close to zero would be needed to calculate the integral precisely. This problem can be avoided if the centre of mass is kept fixed. In MC simulations, where a particle is moved randomly, this can be done by moving the rest of the particles in the opposite direction so that the centre of mass is kept fixed. Another technical detail that is important to keep in mind is that when keeping the centre of mass fixed, with periodic boundary conditions, the particles should never be put back into the "original" box because this would cause a discontinuous jump in the centre of mass position [32]. For a brief description of the centre of mass constraint, see Appendix B

In Section 4.2, a modified version of these techniques is employed to describe the method used in this work: MC sampling of a random configuration tethered to its corresponding potential energy minimum ("lattice sites") via springs ("Einstein crystal"). This method ultimately uses this free energy calculation to determine the volumes of basins of attraction on a PEL.

Chapter 4

The basin volume method

The direct enumeration method [109], described in Chapter 2, determines the probability of each jammed state by minimising random initial conditions until all states are sampled many times. It is clear that this technique only works for systems where the total number of potential energy minima, or mechanically stable jammed states, can be sampled within a reasonable amount of time. It turns out that this method is currently feasible for systems of up to $N \sim 12$ particles due to the huge number of minima. The approach presented in this Chapter, first developed by Xu *et al.* [110], does not rely on direct enumeration, which is why it is possible to estimate the number of potential energy minima while only sampling a small subset of them.

4.1 Introduction

A method to calculate the configurational packing entropy by studying the volumes of basins of attraction on the Potential Energy Landscape (PEL) is used. Each minimum (or inherent structure to use the language of Stillinger and Weber [84]) of the interaction potential corresponds to a jammed state and can be related to a physical system such as gels, emulsions, foams and granular materials as has been previously mentioned in this work. Once the distribution of volumes is calculated the packing entropy can be determined.

In this Chapter the techniques and methods used to calculate volumes of basins of attraction are accounted for. First the details of the simulations are explained. In the next section the parameters, inter-particle potential used and the Gauss-Lobatto (GL) integration algorithm are presented. The following section is an outline of the complete procedure used to calculate the volumes of the basins of attraction, followed by the results and conclusions of this Chapter.

4.2 The volume calculation

In this section, the method for calculating the configurational entropy by measuring the volumes of basins of attraction proposed in [110] is explained. Measuring the volume of a basin of attraction of a minimum on the PEL is equivalent to determining the free energy associated with that high-dimensional volume. To determine this free energy, we will use thermodynamic integration described in Section 3.5. Once a sufficient number of volumes have thus been measured i.e. the distribution of free energies has been sampled, an estimate of the packing entropy of the system can be made.

The volume of a basin of attraction for a given minimum can be defined as:

$$V = \int_{\text{basin}} d\mathbf{r}^N \tag{4.1}$$

where $\mathbf{r}^N = (\vec{r}_1, \ldots, \vec{r}_N)$ is the $d \times N$ dimensional vector of the coordinates of all particles and $\mathbf{r}_0^N$ is the coordinate of the minimum in phase space. The integral is over the coordinates in the basin.

To calculate this integral we perform a thermodynamic integration that links the free energy associated with this volume to that of a harmonic oscillator centred at the minimum. The harmonic potential can be defined as follows:

$$U(\mathbf{r}^N) = \frac{1}{2}k(\mathbf{r}^N - \mathbf{r}_0^N)^2. \tag{4.2}$$

A spring of force constant k is attached to the minimum (or "lattice sites" to use the language introduced in Section 3.5), the partition function can be defined as:

$$Z(k) = \int_{\text{basin}} d\mathbf{r}^N \exp\left[-\frac{1}{2}k(\mathbf{r}^N - \mathbf{r}_0^N)^2\right] \tag{4.3}$$

Note that the volume of the basin is $V = Z(k = 0)$ and the free energy is $F(k) = -\ln Z(k)$ so determining the volume of a basin of attraction is now a free energy calculation and can be calculated using thermodynamic integration. If $u = |\mathbf{r}^N - \mathbf{r}_0^N|$ is the displacement of the system in configuration space then:

$$\begin{aligned}\frac{dF(k)}{dk} &= -\frac{1}{Z(k)}\frac{dZ(k)}{dk} &(4.4)\\ &= \frac{1}{Z(k)}\int_{\text{basin}} d\mathbf{r}^N\,\frac{1}{2}u^2\exp\left[-\frac{1}{2}ku^2\right] &(4.5)\\ &= \frac{1}{2}\langle u^2\rangle_k &(4.6)\end{aligned}$$

Where k corresponds to the coupling constant λ introduced in Section 3.5. This ensemble average is calculated by performing Monte Carlo (MC) simulations (see Section 3.2).

The volume of the basin is $V = \exp[-F(0)]$ so calculating the free energy at $k = 0$ is equivalent to calculating the volume. This can be done by integrating Equation (4.6) from 0 to k_{max}:

$$F(0) = F(k_{\text{max}}) - \int_0^{k_{\text{max}}} dk\,\frac{1}{2}\langle u^2\rangle_k. \tag{4.7}$$

which is equivalent to Equation (3.13), used to calculate the free energy of a system of hard spheres. Each part of the free energy at $k = 0$, $F(0)$, the right hand side of Equation (4.7), is calculated separately. As stated above, $\langle u^2\rangle_k$ is obtained via MC sampling, the integral is calculated using GL quadrature (see Section 4.3.2) and $F(k_{\text{max}})$ is determined using direct sampling considering the probability density

$$f(\mathbf{r}^N - \mathbf{r}_0^N) = \exp\left[-\frac{k_{\text{max}}}{2}(\mathbf{r}^N - \mathbf{r}_0^N)^2\right] \tag{4.8}$$

for the randomly generated displacements in dN-dimensional space. The fraction of points that fall in the basin of attraction is

$$p = \frac{\int_{\text{basin}} \exp\left[-\frac{k_{\text{max}}}{2}(\mathbf{r}^N - \mathbf{r}_0^N)^2\right]\,d\mathbf{r}^N}{\int_{\text{all space}} \exp\left[-\frac{k_{\text{max}}}{2}(\mathbf{r}^N - \mathbf{r}_0^N)^2\right]\,d\mathbf{r}^N} = \frac{N_{\text{in}}}{N_{\text{total}}}. \tag{4.9}$$

The denominator can be calculated analytically

$$\int_{\text{all space}} \exp\left[-\frac{k_{\text{max}}}{2}(\mathbf{r}^N - \mathbf{r}_0^N)^2\right]\,d\mathbf{r}^N = \left(\frac{2\pi}{k_{\text{max}}}\right)^{dN/2} \tag{4.10}$$

so the partition function at k_{max} is

$$\begin{aligned} Z(k_{\text{max}}) &= \int_{\text{basin}} \exp\left[-\frac{k_{\text{max}}}{2}(\mathbf{r}^N - \mathbf{r}_0^N)^2\right] d\mathbf{r}^N && (4.11)\\ &= p\int_{\text{all space}} \exp\left[-\frac{k_{\text{max}}}{2}(\mathbf{r}^N - \mathbf{r}_0^N)^2\right] d\mathbf{r}^N && (4.12)\\ &= p\left(\frac{2\pi}{k_{\text{max}}}\right)^{dN/2} && (4.13)\end{aligned}$$

and the free energy is

$$F(k_{\text{max}}) = -\ln Z(k_{\text{max}}) = -\ln p - \frac{dN}{2}\ln(2\pi/k_{\text{max}}) \tag{4.14}$$

which are the first two terms of the right hand side of Equation (4.18).

To calculate the integral in Equation (4.7) with the GL method (explained in Section 4.3.2) a variable substitution must be performed such that the integral goes between -1 and 1 and the integrand must be a slowly varying function in the desired interval (see Figure 6.7).

We achieve this by using a coordinate transformation in Equation (4.7), such that that the product of $\langle u^2\rangle_k$ and the Jacobian of the coordinate transformation is roughly constant. For large enough k, we have $\langle u^2\rangle_k = dN/k$, because the integrand is only appreciable inside the basin. In contrast, for small k, $\langle u^2\rangle_k$ is independent of k but determined by the basin boundaries. We can use the value of $\langle u^2\rangle_{k=0}$ to define an "effective spring constant" κ: [110] such that

$$\langle u^2\rangle_{k=0} \equiv \frac{dN}{\kappa} \Rightarrow \kappa = \frac{dN}{\langle u^2\rangle_{k=0}} \tag{4.15}$$

It is clear that the product

$$\langle u^2\rangle_k \frac{k+\kappa}{dN} \approx \text{constant.} \tag{4.16}$$

(see Figure 6.7). The variable substitution that yields the appropriate Jacobian is

$$t = \frac{2\ln(1+k/\kappa)}{\ln(1+k_{\text{max}}/\kappa)} - 1 \tag{4.17}$$

And Equation (4.7) can be written as

$$\begin{aligned} F(0) &= F_{\mathrm{HO}} - \ln p - \int_0^{\ln(1+k_{\mathrm{max}}/\kappa)} \frac{1}{2}\langle u^2\rangle(k+\kappa)\, d[\ln(1+k/\kappa)] \\ &= F_{\mathrm{HO}} - \ln p - \int_{-1}^{1} \frac{1}{2}\ln(1+k_{\mathrm{max}}/\kappa)\frac{1}{2}\langle u^2\rangle(k+\kappa)\, dt \qquad (4.18) \end{aligned}$$

where $F_{\mathrm{HO}} = -(dN/2)\ln(2\pi/k_{\mathrm{max}})$ is the free energy of a dN-dimensional harmonic oscillator and p is the fraction of brute force MC moves that fall into the basin of attraction for k_{max}, see Equation (4.9).

A 6-point GL quadrature has been used. An n-point GL quadrature is exact for polynomials up to degree $2n-1$ which is why it is important that the integrand is a smooth, slowly varying function (for an example see Figure 6.7).

It is important to note that direct sampling (non-Markovian) MC, described in Section 3.2, cannot be used to efficiently sample the volumes of the basins of attraction. This technique only works for large enough k where a significant fraction of Gaussian-generated random points fall inside the basin. For systems with low k, most Gaussian-generated points would fall outside the basin and would therefore be rejected, resulting in poor statistics. At $k = 0$ this effect is most notorious because the random numbers are generated uniformly and the fraction of points that fall inside the basin is extremely small. Our method overcomes this by sampling points along a random walk within the basin and using the Metropolis acceptance criteria that gives a certain probability to accept a move according to the energy difference involved. We stress once again that the energy that enters the Metropolis acceptance rule is not the inter-molecular interaction between the soft spheres but the harmonic potential energy of a spring attached to the minimum and the current point in configuration space. In addition, a trial move is also rejected if it would move the system outside the original basin of attraction. During every trial move, we have to test whether this move would keep the system within the original basin. This is tested by minimising the potential energy for every trial move.

Parameter	**Symbol**	**Values**
Dimension of space	d	2
Number of particles	N	$8 \to 128$
Volume/packing fraction	ϕ	$0.86 \to 0.9$
Size of square box (volume$^{1/d}$)	L	1
Diameter ratio	D_{ratio}	1.4
Total number of steps (MC and PT)	n_{step}	$10^5 \to 10^6$
Number of equilibration steps (MC and PT)	n_{eq}	$10^4 \to 10^5$
Number of steps (Direct sampling)	n_{DS}	$10^4 \to 10^5$
Gradient tolerance (minimisation)	`EPS`	10^{-7}
Position tolerance (comparing minima)	r_{tol}	$10^{-5}/\sqrt{N}$

Table 4.1: Parameters used in the simulations of soft harmonic spheres described in this Chapter.

4.3 Details of the model

In this section the interaction between the particles, their size distribution and the initial conditions are described. Most of these parameters were taken from [110] in order to compare, at least at a preliminary stage, the current calculations to their results.

4.3.1 Soft spheres

To model a jammed system configurations are created consisting of N bi-disperse circular particles in two dimensions with Periodic Boundary Conditions (PBC). The bi-disperse systems have a diameter ratio of 1.4. These systems have been chosen because a two-dimensional mixture of disks with this size ratio is hard to crystallise[1] [67, 55]. Uniformly random initial positions were used i.e. ideal gas configurations. Starting with completely random positions ensures that phase space is sampled uniformly [67]. Table 4.1 shows the parameters used for the simulations carried out for this work. The values of particle diameters are determined by the size of the box L, the volume fraction ϕ and the diameter ratio D_{ratio}. The volume fraction is set higher than the RCP limit which means that there will always be overlaps. This ensures that the soft-sphere system is jammed.

---

[1] A mixture of an equal number of large and small disks were used. The sizes are determined by the diameter ratio.

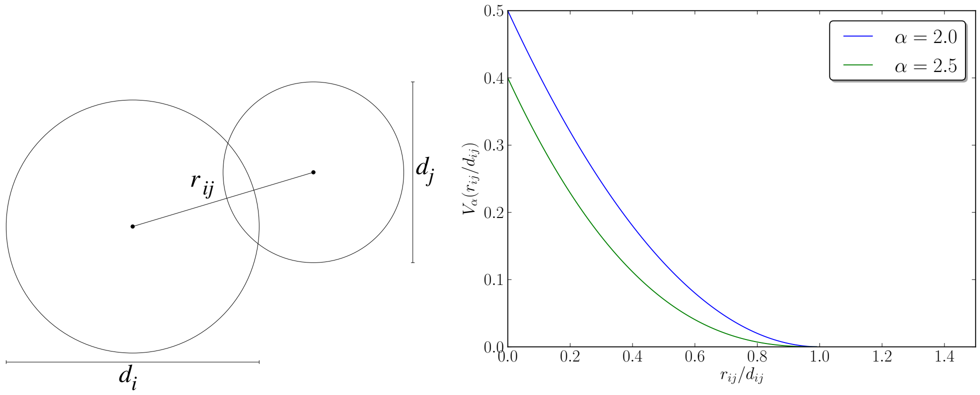


Figure 4.1: Interaction between particles for soft potential. Left: Diagram showing the distances that determine the value of the potential energy between two overlapping particles. Right: Plot of the finite-range repulsive interaction between particles (Equation (4.19)) for $\alpha = 2$ (harmonic) and $\alpha = 5/2$ (Hertzian).

The interaction between particles i and j is given by Equation (4.19) where d_i and d_j are the diameters of particles i and j and r_{ij} is the distance between their centres (see Figure 4.1).

$$V_\alpha(r_{ij}) = \begin{cases} \dfrac{\varepsilon}{\alpha}\left(1 - \dfrac{r_{ij}}{d_{ij}}\right)^\alpha & r_{ij} < d_{ij} \\ 0 & r_{ij} \geq d_{ij} \end{cases} \qquad d_{ij} = \frac{d_i + d_j}{2} \tag{4.19}$$

where ε determines the energy scale and α determines the strength of the repulsion between particles. $\alpha = 2$ is Harmonic and $\alpha = 5/2$ Hertzian, see Figure 4.1 for examples. In this work $\alpha = 2$ has been used as it is more efficient to compute. In Chapter 2 we learnt that the shape of the potential does not affect the jamming transition, [67].

Jammed states are found by taking uniformly random particle positions and minimising the potential energy. Each distinct jammed state corresponds to a different minimum and each minimum has its own basin of attraction.

4.3.2 Gauss-Lobatto integration

The method chosen to perform the integration of $\langle u^2 \rangle_k$ (Equation (4.7)) is GL quadrature. Gaussian quadratures give us the ability to choose not only the weighting coefficients, as in a Newton-Cotes quadrature, but also the abscissas where the function is evaluated [74]. Examples of Newton-Cotes quadrature methods are the trapezoid rule and Simpson's rule. The GL method can be expressed as

$$\int_{-1}^{1} f(x)\,dx = \frac{2}{n(n-1)}[f(1)+f(-1)] + \sum_{i=1}^{n-1} w(x_i)f(x_i) + R_n \qquad (4.20)$$

where $w(x_i)$ is the weighting function

$$w(x_i) = \frac{2}{n(n-1)[P_{n-1}(x_i)]^2} \qquad (4.21)$$

and $P_n(x)$ are the Legendre polynomials. The remainder or error of the GL integration is

$$R_n = \frac{-n(n-1)^3 2^{2n-1}[(n-2)!]^4}{(2n-1)[(2n-2)!]^3} f^{(2n-2)}(\xi) \qquad -1 < \xi < 1. \qquad (4.22)$$

For a table of values of x_i and $w(x_i)$ see [1].

4.3.3 Volume calculation: the procedure

In this section the procedure used to calculate the volumes of the basins of attraction is presented as a list of steps. This is done so that the reader is aware of the details of the simulations and the order in which the different parts are carried out.

1. Initialise particle positions in a box.

2. Minimise to find the corresponding potential energy minimum (jammed configuration). This minimum will be used throughout the procedure and the volume of its basin of attraction will be measured.

3. Perform direct (Gaussian, brute force) MC sampling to determine k_{max}, $\langle u^2\rangle_{k_{\text{max}}}$ and p:

    (a) Choose an arbitrary k_{max}.

    (b) Start at the minimum and move the system by choosing a random Gaussian displacement.

    (c) Minimise.

   (d) If the minimisation returns to the same minimum, accept. If not, reject. If accepted measure the distance u to the minimum.
   (e) If not the last step, n_{DS}, return to (b).
   (f) If p is too large, reduce k, if p is to small, increase k. Now return to (b) if k was changed. If not, p is in the correct range (which is predefined and somewhat arbitrary) and the MC sampling for k_{max} is finished, move on to step 4..

4. Run Markov chain MC routine to determine $\langle u^2 \rangle_{k=0}$.
   (a) Start at the minimum.
   (b) Choose a random particle and give it a random displacement (with a maximum step size).
   (c) Minimise.
   (d) Check if the minimum reached is the same as the initial one (This is done by comparing particle positions. If it is, the move is accepted, if not it is rejected. Measure the distance u to the minimum.
   (e) Go back to (b) and repeat until step n_{step}.

5. Determine the ks for GL integration (see Section 4.3.2).

6. Run MC for remaining values of k (determined in the previous step by the GL algorithm) using Parallel Tempering (PT) moves.
   (a) Separately for each k, start at the minimum.
   (b) For each k, choose a random particle and give it a random displacement (with a maximum step size).
   (c) For each k, accept the move with the probability given in Equation (3.1). If the move is rejected go to (b).
   (d) Minimise.
   (e) Check if the minimum reached is the same as the initial one (this is done by comparing particle positions). If it is, the move is accepted, if not, it is rejected.
   (f) Every 100 steps try to swap adjacent (in k) configurations with the probability given in Equation (3.2).

   (g) Go back to (b) and repeat until step n_{step}.

7. Perform GL integration (Equation (4.18)) to determine the free energy.

This procedure is repeated to calculate the volumes of basins of attraction of different minima. Once a certain number of basins have been sampled and their volumes measured, a histogram of their free energies can be calculated and analysed. This process is described in Chapter 6.

4.3.4 Other technical details

When using energy minimisation on a PEL to find minima that correspond to jammed states, it is possible that a certain number of particles, known as “rattlers” or “floaters” are left unjammed surrounded by cages of stable, jammed particles. This is possible as long as the structure remains mechanically stable to external forces. These particles are free to move without any energy cost and are normally removed from simulations [67, 20]. In the study presented in this Chapter, floaters have been identified by particles with two contacts or less. This criterion seems fair as the system is composed of similarly sized particles and almost certainly a particle with three contacts will be stable whereas two contacts is almost always a floater and one and zero contacts is automatically a floater. These particles are not considered when comparing two minima to test whether a certain point in configuration space belongs to a basin of attraction.

Energy minimisation is another important aspect of the method described in this Chapter. It must be performed at every MC step of the simulation (see section 4.3.3) and needs to be as efficient and accurate as possible. This means that the algorithm must be able to find the correct minimum within a very narrow tolerance using the least amount of time. The chosen methods are Steepest Descent (SD) and Limited memory BFGS (L-BFGS) [53] which is based on the approach by Broyden, Fletcher, Goldfarb and Shanno optimised for limited memory so suited for high dimensional problems.

4.4 Results and discussion

In this Section the results of this Chapter are presented. A histogram of basin free energies, calculated using 1000 minima for 64 particles is shown (Figure

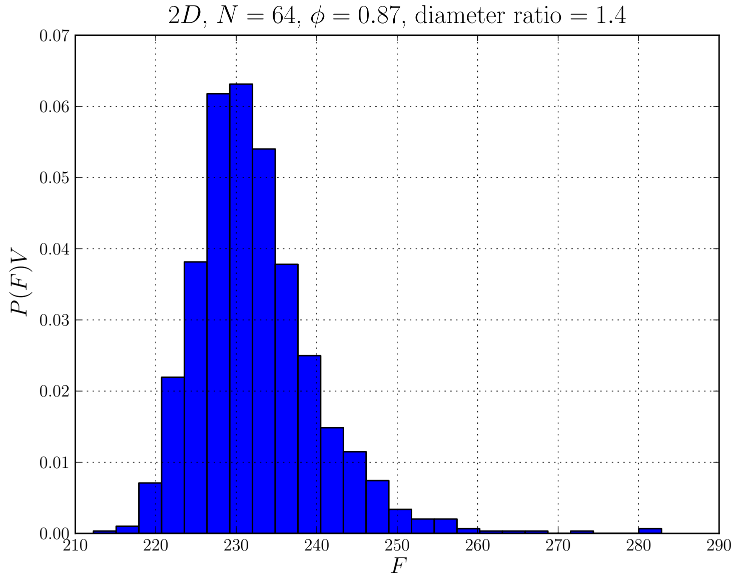


Figure 4.2: Histogram showing extremely small volumes possibly created by L-BFGS.

4.2) to highlight the fact that basins with very high free energy, corresponding to extremely small volumes, are very likely. This is because there is an inherent bias in the probability to fall into a basin given that bigger basins, having a bigger catchment area, are more likely. Thus the pure fact that these basins appear at all suggest that there must be many of them. The minimisation algorithm L-BFGS was used for these simulations given its efficiency versus more traditional methods like SD. As these apparent tiny basins present on this PEL are really many orders of magnitude smaller than average, a test was devised to figure out if this was merely an artefact of the minimisation scheme.

The difference between a projection of a basin of attraction defined by two different minimisation schemes, L-BFGS and SD is studied in the following way: A $2N$-dimensional basin of attraction of a minimum of the potential energy of a system of jammed disks interacting via a repulsive, harmonic potential (described in Section 4.3.1) has been projected onto the Cartesian coordinates of different particles. In Figure 4.3 a basin of attraction has been projected onto the two dimensional Cartesian coordinates of two small particles in a $N = 8$ particle bi-disperse ($50 : 50$ small and large particles) system with diameter ratio $D_{\text{ratio}} = 1.4$ in two dimensions. The left hand column

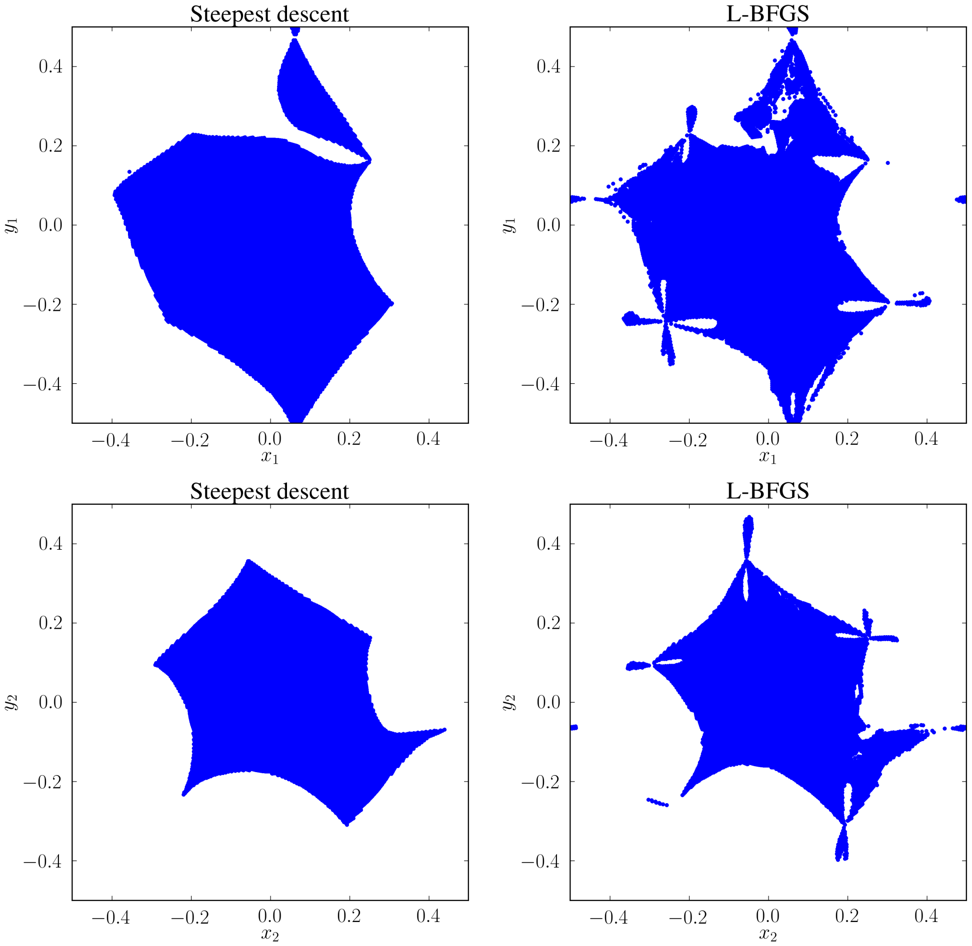


Figure 4.3: Difference between basins of attraction of the same minimum probed with different minimisation algorithms (projected onto two coordinates).

shows the results using SD and the right hand column shows the results obtained with L-BFGS minimisation. Figure 4.4 show the same minimum but now the basin has been projected onto the coordinates of one of the large particles.

The differences between basins of attraction found with L-BFGS and SD can be observed in Figures 4.3 and 4.4. The importance of sampling the minima “properly” when measuring the volumes of their basins of attraction becomes apparent. If a random walk is performed and one of the small, disconnected regions seen in Figure 4.4 happens to be sampled, this region will be defined as the entire basin and therefore the conclusion is that the volume of the basin is merely the volume of this small region. This leads to spurious

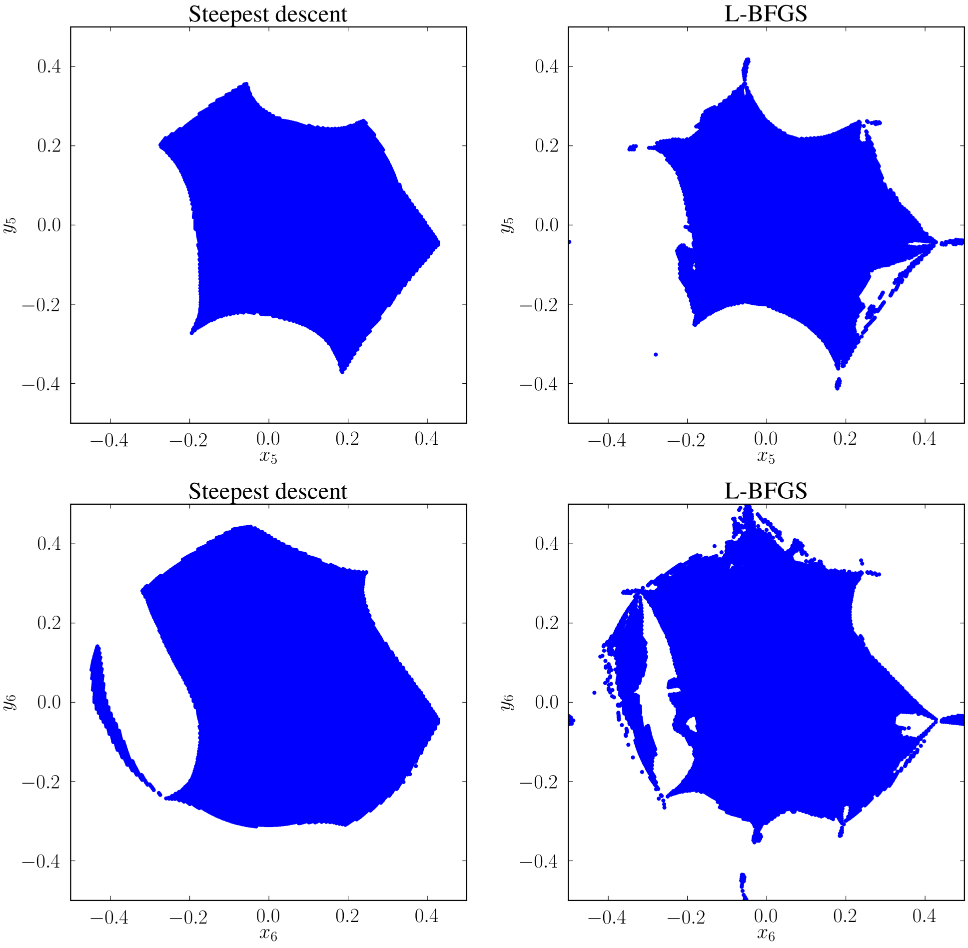


Figure 4.4: Difference between basins of attraction of the same minimum probed with different minimisation algorithms (projected onto two coordinates).

results that do not correctly represent the physics of the system. This finding prompted a rigorous study of how different minimisation algorithms define basins of attraction, the topic of Chapter 5. Another important difference between the two energy minimisation methods is the number of energy and force calculations needed to reach a minimum, which determines the time it takes. This issue is also addressed in Chapter 5.

Another result that has been uncovered during this study is that the convergence of the mean square displacement, u^2 is very slow for $k = 0$, when the system is free to move away from the minimum ("lattice sites"), with no penalty. This can be seen in Figure 4.5. Here a steady value is not reached, even after 10^5 steps. This is made even worse by the fact that starting two

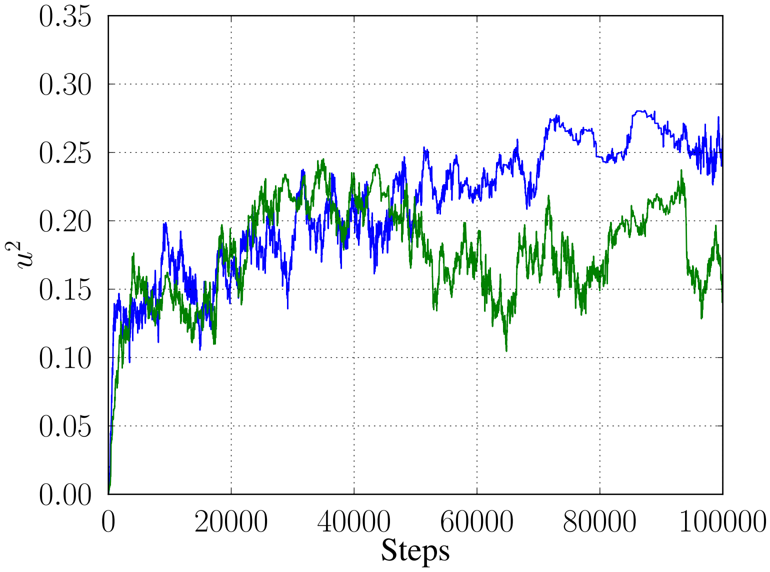


Figure 4.5: Mean square displacement as a function of MC steps for $k = 0$. Note how different MC runs starting from the same point do not converge to the same value even after 10^5 steps.

different MC runs from the same point gives totally different values for $\langle u^2 \rangle$ at $k = 0$ for the same basin, even after 10^5 steps. This is illustrated in Figure 4.5 by the blue and green curves. This suggests that long equilibration times are needed for the sampling at $k = 0$, adding to the total simulation time to calculate the basin volumes.

4.5 Conclusions

Basins of attraction defined using L-BFGS are clearly very different from those defined with SD. This is evident in Figure 4.3 and 4.4 where the same basin is mapped with both minimisation methods. L-BFGS works near the centre of the map but produces strange effects at the borders.

The difference between the basins produced with different minimisation methods led us to use SD as it will always find the “correct” (by definition, as seen in Chapter 1) basin, at least for infinite precision (infinitely small step size). The main problem with SD is that a typical MC run takes 25 times longer than a run using L-BFGS. This is due to the fact that L-BFGS is able to find a more efficient direction (see Appendix A for a detailed description) on the potential energy surface and therefore reaches the minimum in fewer steps, as discussed in Chapter 5.

In conclusion, L-BFGS affects the morphology of the basins of attraction and therefore their volumes. The equilibration of the MC sampling is also affected by L-BFGS because the system becomes trapped in regions of small volume "created" by L-BFGS (not present when using SD). The volume is not sampled properly as the system cannot explore the basins uniformly. This is a possible explanation for the presence of basins with extremely small volumes (~ 50 orders of magnitude smaller than the average volume) as shown in the free energy histogram presented in Figure 4.2.

With respect to the problem of an extremely erratic mean square displacement at $k = 0$, shown in Figure 4.5, a novel method will be described in Chapter 6 that reduces (to some extent) the equilibration time needed to converge to a steady value of $\langle u^2 \rangle_{k=0}$.

The failure of the L-BFGS method to generate compact basins of attraction makes the method less attractive for the present project. The difference found between the basins of attraction defined by the two minimisation algorithms tested in this Chapter motivated us to perform a more systematic, quantitative study of the differences between basins defined by various minimisers. This is the topic of the next Chapter where we test some state-of-the-art algorithms against SD to see how they perform both in terms of efficiency and in terms of the definition of the basins.

Chapter 5

Defining basins with different minimisers

Energy minimisation is an important aspect of this work as it must be performed at every Monte Carlo (MC) step of the simulation to check if a certain point in configuration space belongs to a particular basin of attraction (the details of the algorithm are described in Section 4.3.3). Therefore it needs to be as efficient and accurate as possible. This means that the algorithm must be able to find the correct minimum within a very narrow tolerance using the least amount of time.

The motivation for looking at the way different minimisation algorithms define basins of attraction of minima comes from the first preliminary results of this project. As has been seen in Section 4.4, the method used to determine the configurational entropy relies on measuring volumes of basins of attraction on the Potential Energy Landscape (PEL). When these volumes were first measured there were some extremely small volumes, orders of magnitude smaller than the next smallest ones observed. This finding seemed paradoxical: volumes that are extremely small (often by more than 20 orders of magnitude than the most common volumes) are extremely unlikely to be sampled, unless there are huge numbers of them (and there is certainly no *a priori* reason to suspect that). The alternative explanation is that the observation of small basin volumes is an artefact of the minimisation algorithm used. This question led to the visual examination of the boundaries of a basin of attraction when a particle is displaced from its position at the potential energy minimum.

A study of the basins of attraction for potential energy minima defined by different minimisation algorithms for an atomic system is reported in this chapter. The findings indicate that whereas some minimisation algorithms produce connected basins, others produce basins with complex boundaries or basins consisting of disconnected parts. Such basins deviate from the "correct" basin of attraction defined by Steepest Descent (SD) pathways, and the differences can be controlled to some extent by adjustment of the maximum step size. The choice of the most convenient minimisation algorithm depends on the problem at hand. It has been shown that while Limited memory BFGS (L-BFGS) is the fastest minimiser, the Fast Inertial Relaxation Engine (FIRE) algorithm is also quite fast, and can lead to less fragmented basins of attraction.

5.1 Introduction

Optimisation problems are ubiquitous in the physical sciences and beyond. In the simplest case optimisation refers to the search for the minimum or maximum values of an objective function. Global optimisation involves searching for the highest maximum or lowest minimum in a certain domain. In contrast, local optimisation procedures identify the first minimum or maximum that is found by a given algorithm when starting from an arbitrary point in parameter space.

In the study of energy landscapes, the properties of stationary points and the connections between them are of central importance [99]. These stationary points represent key features of the landscape. In chemical reactions, saddle points are geometric transition states [65] along the reaction coordinate. Glassy systems are trapped in meta-stable local minima [43, 19, 18] and, similarly, jammed states can also be viewed as local potential energy minima. The study of these minima and the pathways connecting them can be carried out using geometry optimisation techniques [99, 100], and local minimisation is the focus of the current Chapter.

When faced with the task of numerically optimising a smooth function there are many algorithms from which to choose. Which algorithm is best suited for the purpose depends on factors such as speed, memory usage and ease of implementation. All algorithms follow a general procedure starting with the user supplying an initial point, which can be an informed guess or an arbitrary point in parameter space. The algorithm generates a sequence of iterates that terminates when a solution is found within a predefined accuracy, such as when the gradient is near zero, or when the value of the function stops changing. Recent work has shown how convergence criteria can be chosen according to a certification procedure [60]. Different algorithms have different ways of proceeding from one iteration to the next. The formulations considered here involve the value of the function that is being optimised, its derivatives, and the results of previous iterations.

In general, two different algorithms can converge to different minima starting from the same initial conditions. In this work, the interest lies in identifying the configuration space that leads to a particular minimum as a function of the optimisation algorithm. This connection is important for applications such as calculation of thermodynamic properties using the superposition approach, where the global partition function is written as a sum over contributions of local minima [97, 22, 86] and the method used in Chapter 4 of this work. In this context, the SD algorithm occupies a unique position, since it defines basins of attraction [62] for local minima that must be compact, in the sense that any point in the basin (e.g. the minimum) must be connected to any other point in that basin (e.g. the starting position) through a path that is completely inside the basin. This result follows because SD paths are defined using a linear first-order differential equation, for which the uniqueness theorem applies [71]. Importantly, when the SD is started from any point on a pre-existing trajectory that was generated by SD, then the new SD path will follow the pre-existing path. This means that any point on a trajectory between an arbitrary origin inside the basin and the minimum of that basin necessarily belongs to the basin. In other words, SD generates compact basins. However, SD minimisation is very inefficient compared to more sophisticated algorithms, which are normally preferable. The latter methods generally employ non-linear equations to determine the steps, and the corresponding basins of attraction can exhibit complex boundaries [96, 98] and, more to the point, need not be compact. In contrast, the basins for other algorithms can exhibit re-entrant, interpenetrating boundaries [96, 98]. For applications such as basin-hopping global optimisation [101, 51] this structure is probably unimportant. However, the simplicity of the boundaries associated with SD paths is relevant if we are interested in partitioning the configuration space, for example, when measuring the size of basins of attraction using the technique described in Chapter 4. In the present work, the basins of attraction defined by SD are regarded as a reference to which other methods will be compared. The purpose of this Chapter is to make these comparisons rigorously and to provide criteria for choosing the most appropriate and convenient minimisation algorithm for a given problem.

5.2 Methods

The minimisation algorithms considered here are SD, L-BFGS, FIRE, Conjugate Gradient (CG), and Broyden, Fletcher, Goldfarb and Shanno (BFGS) (for a detailed description of different minimisation techniques see [66]). The L-BFGS algorithm is tested using two different methods to determine the length of the steps. In the first approach, a line search routine is used to choose the step size. In the second approach, the step size guess of the L-BFGS algorithm is accepted subject to a maximum step size and the condition that the energy does not rise. This is the default procedure in the global optimisation program `GMIN` [102] and the `OPTIM` program for locating transition states and analysing pathways [94]. We have only compared gradient-based minimisers in the present work, because they represent the most efficient class of algorithms.

SD, sometimes referred to as gradient descent, uses the gradient as the search direction (this is the steepest direction). The step size can be chosen using a line search routine. In this Chapter, a fixed step size ($\Delta = 0.005$ in reduced units) is used for all of the SD calculations. It is worth noting that the definition of basins of attraction in Section 5.1 applies to SD minimisation in the limit of infinitesimal step size.

BFGS, named after its creators Broyden [12], Fletcher [28], Goldfarb [33], and Shanno [79] is a quasi-Newtonian optimisation method, which uses an approximate Hessian to determine the search direction. The approximate Hessian is built up iteratively from the history of steps and gradient evaluations. The implementation used in this Chapter is from `SciPy` [40] and uses a line search to determine a step size. The line search used is the `Minpack2` method `DCSRCH` [64], which attempts to find a step size that satisfies the Wolfe conditions (discussed briefly in Appendix A), ensuring a certain decrease in the objective function. The maximum step size is fixed to be 50 times the initial guess returned by the BFGS algorithm.

L-BFGS is a limited memory version of the BFGS algorithm described above and was designed for large-scale problems, where storing the Hessian would be impractical. Rather than saving the full approximate Hessian in memory it only stores a history of M previous values of the function and its gradient with which it computes an approximation to the inverse diagonal components of the Hessian [53]. For a system with N variables, $\mathcal{O}(N^2)$ memory and operations are needed when using BFGS, while L-BFGS scales as $\mathcal{O}(MN)$, which is significantly smaller if $M \ll N$, and is linear in N. Two versions of L-BFGS were used in this Chapter. The first is from the `SciPy` [40] optimisation library ”L-BFGS-B” [113, 14, 63]. This routine uses the same DCSRCH line search as the BFGS implementation, but with slightly different input parameters. For example, the maximum step size is adaptively updated. The second L-BFGS implementation is included in the `GMIN` [102] and `OPTIM` [94] software packages and adapted from Liu and Nocedal [53]. In this version there is no line search. The step size returned by the L-BFGS algorithm is accepted subject to a maximum step size constraint and the condition that the energy does not rise. In both of these versions, the diagonal components of the inverse Hessian are initially set to unity.

The fast inertial relaxation engine, known as FIRE, is a minimisation algorithm based on ideas from molecular dynamics, with an extra velocity term and adaptive time step [9]. Stated simply, the system slides down the potential energy surface gathering “momentum” until the direction of the gradient changes, at which point it stops, resets, and resumes sliding.

The CG method uses information about previous values of the gradient to determine a conjugate search direction [37]. It only stores the previous search direction. The implementation considered here is from `SciPy` [40], and the step size is determined using the same line search as the BFGS routine.

A detailed description of the optimisation algorithms mentioned above is given in Appendix A along with a modification to the FIRE algorithm that is able to minimise an inter-molecular potential with hard cores.

In order to test the accuracy of the minimisers, a three-particle system in which the inter-particle interactions are given by a Lennard-Jones (LJ) potential plus a three-body Axilrod–Teller term [49, 6] is used:

$$\begin{aligned} V &= 4\varepsilon \sum_{i<j} \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right] + \\ & \qquad Z \sum_{i<j<k} \left[\frac{1 + 3\cos\theta_1 \cos\theta_2 \cos\theta_3}{(r_{ij} r_{ik} r_{jk})^3} \right] \end{aligned} \tag{5.1}$$

Here θ_1, θ_2 and θ_3 are the internal angles of the triangle formed by particles i, j, k; r_{ij} is the distance between particles i and j; and Z is the strength of the three-body term. This three-particle system is chosen because, for $Z > 0$, it has four local minima. It is practical to choose a small system with only a few degrees of freedom in order to visualise the basins of attraction in two dimensions. In one of the minima, the atoms are arranged in an equilateral triangle and the other three linear minima are related by permutations of the atoms. Reduced units with one parameter are used, as in [95, 21, 96, 98]:

$$Z^* = \frac{Z\sigma^9}{\varepsilon}. \tag{5.2}$$

Without loss of generality, we can define axes such that the three particles are in the xy plane with one particle at the origin, another along the x axis, and the third in the upper half plane. Now only the three internal coordinates r_{ij} are needed to describe the system. A projection onto the page was chosen to visualise the basins of attraction in such a way that the basins of the four minima are present in the plane [96, 98]. In internal coordinates, the projection plane is chosen to be perpendicular to the vector $\vec{n} = (1, 1, 1)$ at a distance $\sqrt{3}\alpha$ from the origin. Points in the plane have the property $r_{12} + r_{23} + r_{13} = 3\alpha$. An arbitrary vector can be defined such that $\vec{v} = (0, 0, 1)$ and the plane is spanned by the unit vectors

$$\hat{x}_2 = \frac{\vec{n} \times \vec{v}}{|\vec{n} \times \vec{v}|}, \qquad \hat{x}_1 = \frac{\vec{n} \times \vec{x}_2}{|\vec{n} \times \vec{x}_2|}. \tag{5.3}$$

The projection of an arbitrary vector $\vec{a} = (r_{12}, r_{23}, r_{13})$ onto the plane is $\vec{a}_p = (\vec{a} \cdot \hat{x}_1, \vec{a} \cdot \hat{x}_2)$. The equilateral triangle minimum is at the origin in terms of the projected coordinates x_1 and x_2, as shown in Figure 5.1. For more details regarding the projection see [96, 98].

The following results were produced using the projection described above

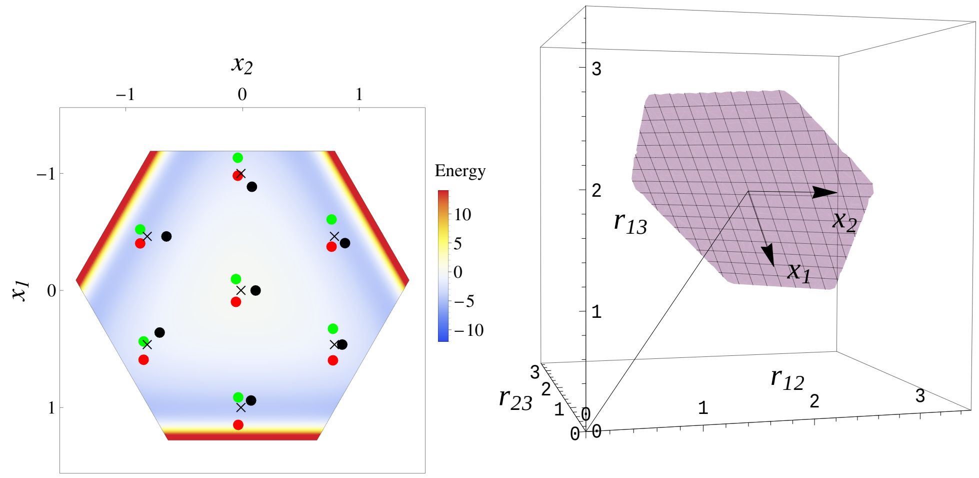


Figure 5.1: Projected plane in its own coordinate system. Points on this plane were used as the initial configurations for the minimisations. Examples of starting configurations are shown for several points marked by "×". The plot is coloured by the energy of the starting configuration.

with $\alpha = \sqrt{3}R_e$, where $R_e = 2^{1/6}\sigma$ is the LJ equilibrium separation, and $Z^* = 2$. For this choice there is a linear minimum with energy $-2.219\,\varepsilon$ in addition to the equilateral triangle with energy $-2.185\,\varepsilon$. There are three distinct permutations of the linear minimum, since any of the three atoms can reside in the central position. A 700×700 grid of initial points was taken with x_1 and x_2 between $-\alpha$ and α. All of the minimisations were terminated when the root mean square (RMS) gradient was smaller than 10^{-3} reduced units. Under some conditions, such geometry optimisations could appear to converge to a saddle point [91], so the geometries were also checked, as well as the RMS gradient. As in [96, 98], each pixel in the resulting plots corresponds to a different initial configuration and is coloured according to the minimum that is found after optimisation.

The efficiency of each algorithm was also tested. Here there is no need to be constrained to small systems by the need for visualisation, so a more interesting system size was chosen, namely 38 LJ atoms. The average number of function calls needed to get to the nearest local minimum from $1,000$ random starting configurations was measured. The number of function calls is a fairer test than wall clock time because for most real word calculations computing the energy and gradient will be the time bottleneck and it avoids measuring differences in implementation efficiency. The results are reported

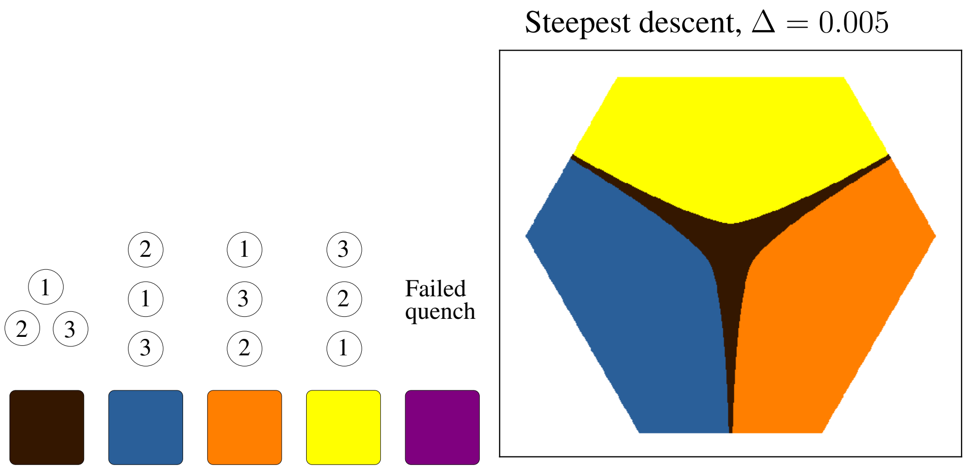


Figure 5.2: The panel on the right shows the same (x_1, x_2) plane as Figure 5.1, which defines the starting configuration of the three particles that interact via the LJ potential plus a three-body Axilrod-Teller term, as described in Section 5.2. The plane is coloured according to the final configuration after a SD minimisation. The colour coding is displayed in the panel on the left. Black corresponds to the atoms in the triangle configuration, while blue, orange, and yellow correspond to the three possible linear configurations. A failed quench means that the final coordinates are not close enough (according to a certain tolerance) to the equilateral triangle or linear configurations. Δ is the maximum step size. This figure, corresponding to the SD results, will serve as our reference for comparing the other minimisers.

in Table 5.2.

5.3 Results and Discussion

Figure 5.2 shows the colour scheme used to identify the results of local minimisation in the subsequent figures. The white region of the Figure corresponds to forbidden geometry and refers to the points in the plane that correspond to initial geometries that do not satisfy the triangle inequality or have excessively high energy. Failed quench means that the quenched coordinates are not close enough (according to a certain tolerance) to the equilateral triangle or linear configurations, that is, the algorithm failed to reach a minimum [91].

The figures that follow show the basins of attraction of the four minima described above determined using the different minimisation techniques and parameters, as described in Section 5.2. As expected, SD is the slowest (see

Table 5.1), most robust minimiser, and it produces well-defined basin boundaries (Figure 5.2). This result holds as long as the step size is kept relatively small. Smaller step sizes are always more robust when using SD. The usual definition for the basin of attraction in the context of energy landscapes is the set of points in configuration space that converge to a certain minimum for a SD quench [62, 99]. Hence this approach produces a useful reference against which to compare the other algorithms.

L-BFGS is the fastest algorithm tested here (see Table 5.2), although the basin boundaries are not always well defined (see Figure 5.3 and Figure 5.6). In the case of L-BFGS without line search it is clear that reducing the step size does not necessarily improve the definition of the basin boundaries. In this case, the resolution of the basin boundaries improves with increasing maximum step size until it reaches an optimum length, beyond which the resolution decreases. This effect is clearly visible in Table 5.1 and Figure 5.3. Removing the line search does not improve the resolution of the boundaries, but it does reduce the number of failed quenches. The effect of changing the parameter M, the number of previous values of the function and gradient used to build the approximate Hessian was tested. Increasing M between 1 and 10 makes the resolution of the basins worse (see Figure 5.3) but produces faster convergence (see Table 5.1). This result arises due to the fact that increasing M increases the degree of non-linearity of the algorithm.

Several values of the maximum step size for FIRE were tested. For a small value of the step size the boundaries of the basins of attraction are well defined and similar to the results for SD. The method only ends up in the wrong basin when starting from points that lie very close to the boundaries between two basins (Figure 5.4, top left). Using a larger value of the step size leads to many artefacts and failed quenches, which are evident in the bottom half of Figure 5.4.

Some other popular algorithms were also tested, namely, CG (Figure 5.5) and BFGS (Figure 5.6). The `SciPy` implementation of these methods uses the same line search routine to determine the step size at each iteration, and both of them produce similar ill-defined basin boundaries. In both cases, the boundary artefacts are caused by the line search returning a step size that is large enough to move into a different basin. Furthermore, the failed quenches at the edges are due to step sizes sufficiently large that particles end up so far apart that the gradient is small enough to satisfy the termination condition.

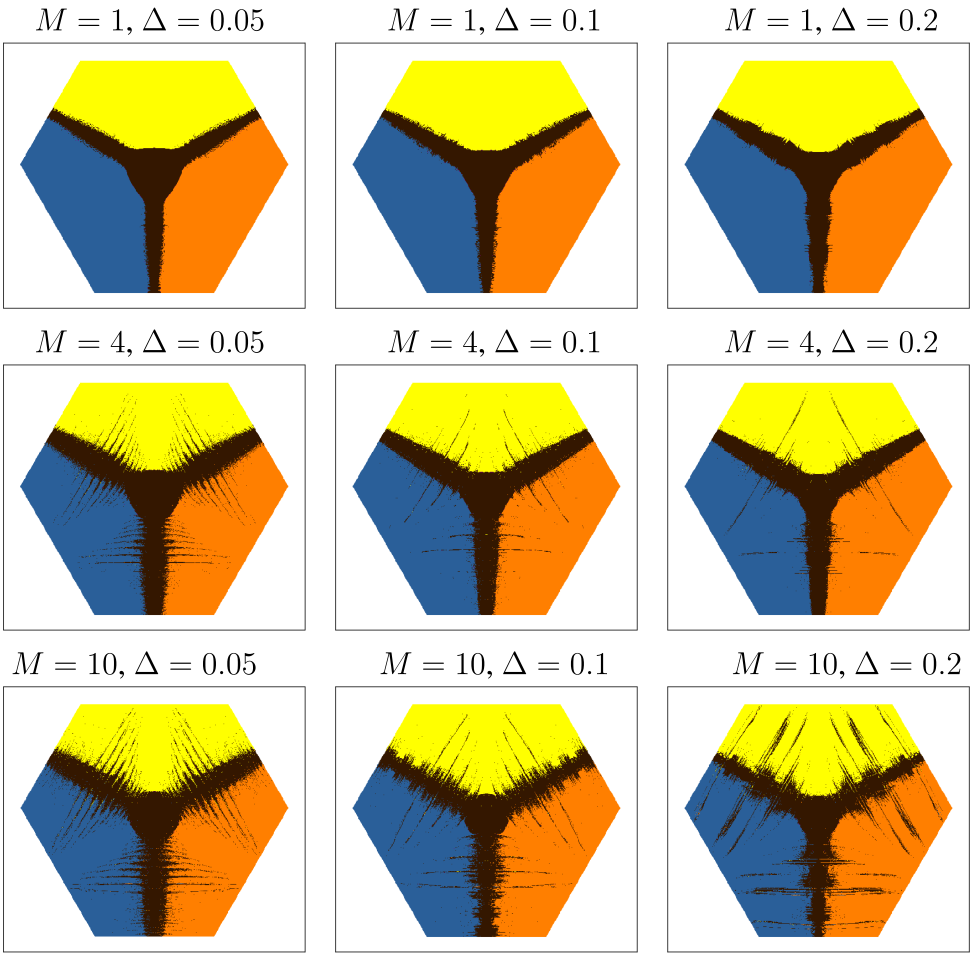
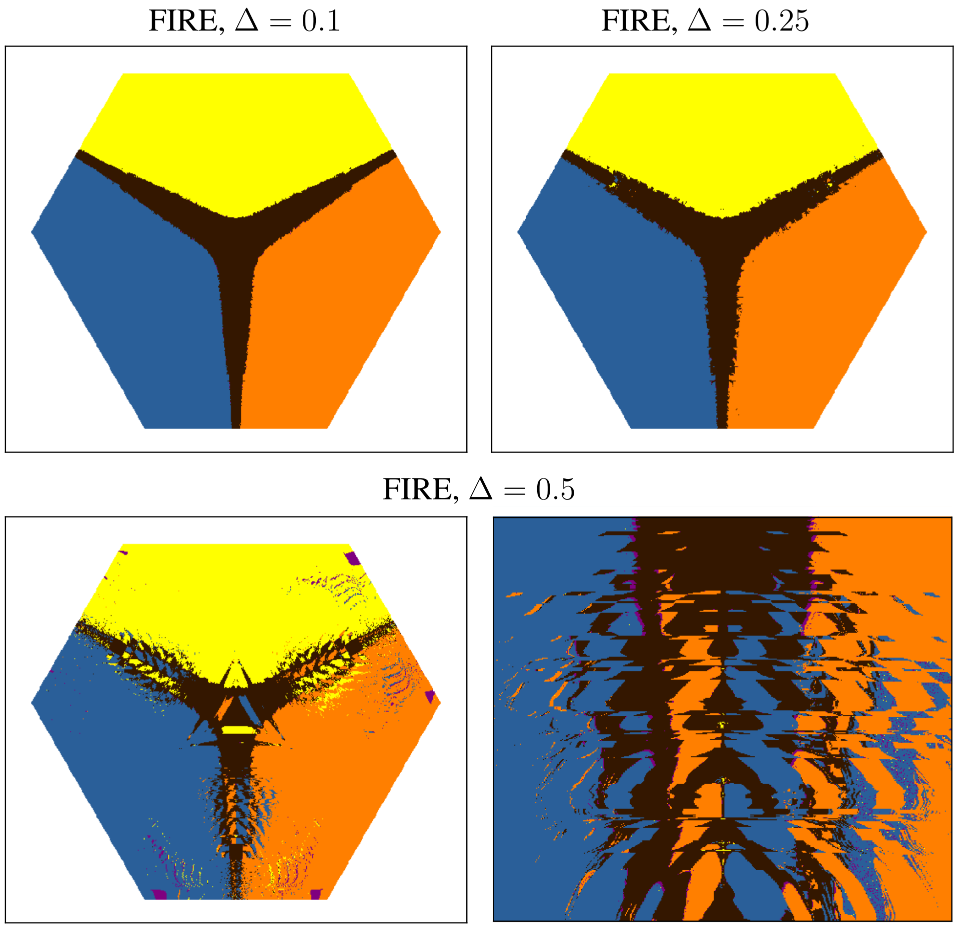


Figure 5.3: Results for the L-BFGS algorithm for different values of M (the number of previous steps used to construct the next step) and the maximum step size Δ. See Figure 5.2 for a detailed explanation of the figure.



FIRE, $\Delta = 0.1$

FIRE, $\Delta = 0.25$

FIRE, $\Delta = 0.5$

Figure 5.4: Results for the FIRE algorithm for several values of the maximum step size Δ. The bottom right panel is a magnification of a portion of the bottom left image. See Figure 5.2 for a detailed explanation of the figure.

Conjugate gradient, with line search (`SciPy`)

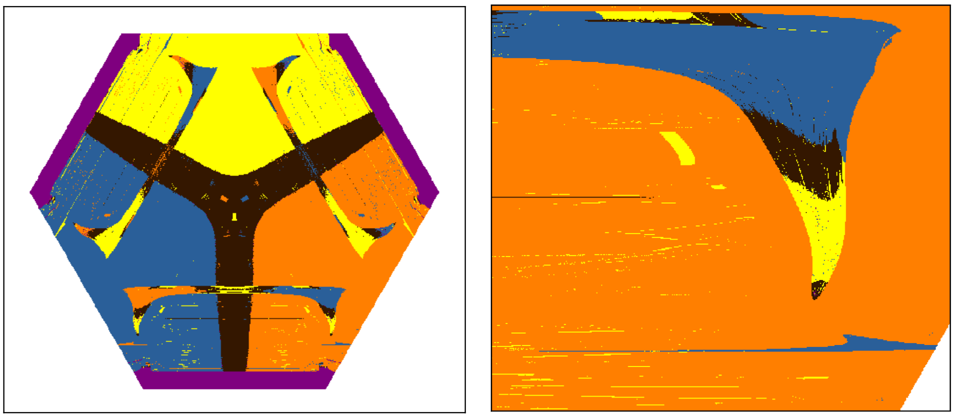


Figure 5.5: Results for the CG algorithm with line search (`SciPy` implementation) are shown in the left panel. The right panel shows a magnification of the left panel. See Figure 5.2 for a detailed explanation of the figure.

The line search algorithm is not entirely responsible for the imprecise basin boundaries. The initial guess for the step size passed to the line search by the CG and BFGS algorithms is often large enough to step to the next basin by itself. To check this effect, this line search routine was also tested with the L-BFGS algorithm. The results (not shown for brevity) produce quite reasonable basin boundaries with most of the above artefacts absent. An interesting question is why the L-BFGS algorithm produces an accurate guess for the step size while BFGS tends to overestimate the step size. The answer, most likely, is that the initial Hessian in L-BFGS is scaled [53], while in BFGS it is fixed to unity.

The effect of maximum step size on the conjugate gradient and BFGS algorithms was tested; however the `SciPy` routines do not accept parameters for adjusting the maximum step size. This adjustment was introduced by modifying the source code of the line search routine used in each case. With these modifications and a maximum step size of $\Delta = 0.1$, the CG routine produced reasonably accurate basin boundaries. The penalty for this improved precision was roughly 50% more function evaluations. No significant improvements for the BFGS routine were obtained.

The basins of attraction determined by some of the methods mentioned above (in particular FIRE with a large step size, CG and BFGS) display complex structures, as shown in Figure 5.4 (bottom right) and Figure 5.5 (right). Here it is clear that the basin boundaries still have structure, even as the length

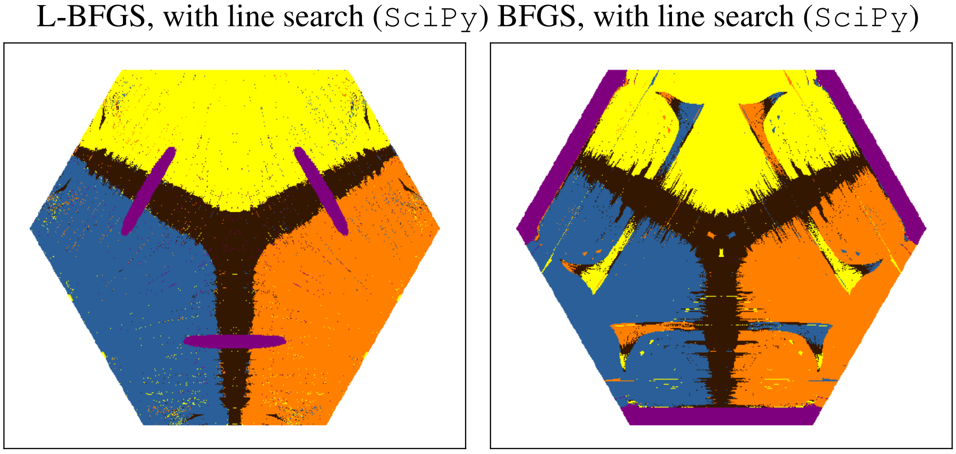


Figure 5.6: Results for L-BFGS and BFGS with line search (`SciPy` implementations). See Figure 5.2 for a detailed explanation of the figure.

scale is reduced. As noted in previous work, the structure may be fractal [34, 96, 98], although this possibility was not investigated in detail.

The difference between the outcomes of the basin mapping for different minimisers has been quantified by counting the number of starting structures for which a basin different from the one obtained using SD is found, as shown in Table 5.1. This difference corresponds to the number of different structures (from a total of $170,607$ valid starting structures) when comparing the minimum produced by the corresponding algorithm with the minimum produced by SD (Figure 5.2).

Table 5.2 reports the performance of the algorithms tested here in terms of the average number of times that the energy and the force were evaluated $\langle$FCs$\rangle$. The average is taken over a sample of $1,000$ random initial states for a 38-particle LJ cluster. The number of evaluations ultimately determines the time it takes to find a minimum, as this is generally the most time consuming part of any minimisation algorithm. The results show that that L-BFGS is the fastest and FIRE is about three to four times slower, while SD is orders of magnitude slower, as expected.

5.4 Conclusions

In this Chapter, the basins of attraction of a simple system have been mapped onto a plane to compare a number of minimisation algorithms. We are able to compare the different approaches both visually and quantitatively, building

Algorithm	**Parameters**	**Err (%)**
L-BFGS	$M = 1, \ \ \Delta = 0.05$	6.41
(without LS)	$M = 1, \ \ \Delta = 0.1$	6.46
	$M = 1, \ \ \Delta = 0.2$	7.16
	$M = 4, \ \ \Delta = 0.05$	14.75
	$M = 4, \ \ \Delta = 0.1$	10.84
	$M = 4, \ \ \Delta = 0.2$	10.13
	$M = 10, \Delta = 0.05$	16.78
	$M = 10, \Delta = 0.1$	12.72
	$M = 10, \Delta = 0.2$	14.81
L-BFGS		
(`SciPy`, with LS)	$M = 10$	12.79
FIRE	$\Delta = 0.1$	3.68
	$\Delta = 0.25$	4.53
	$\Delta = 0.5$	11.23
BFGS		
(`SciPy`, with LS)		23.45
CG		
(`SciPy`, with LS)		22.93
SD	$\Delta = 0.001$	0.00

Table 5.1: Table showing the quantitative analysis of the basins of attraction defined by different minimisers for the three-body system. The third column, Err, is the percentage of starting configurations that minimised to a different minimum compared to the SD method. Δ is the maximum step size and M is the length of the history used by L-BFGS, as defined in Section 5.2.

Algorithm	**Parameters**	**⟨FCs⟩**
L-BFGS	$M = 1, \ \ \Delta = 0.1$	273
(without LS)	$M = 4, \ \ \Delta = 0.05$	369
	$M = 4, \ \ \Delta = 0.1$	241
	$M = 4, \ \ \Delta = 0.2$	225
	$M = 10, \Delta = 0.1$	228
	$M = 20, \Delta = 0.1$	216
L-BFGS		
(`SciPy`, with LS)	$M = 10$	215
FIRE	$\Delta = 0.1$	822
	$\Delta = 0.5$	3,185
BFGS		
(`SciPy`, with LS)		1,233
CG		
(`SciPy`, with LS)		837
SD	$\Delta = 0.001$	31,672

Table 5.2: Benchmarks for the algorithms tested in this Chapter in terms of the average number of function calls ⟨FCs⟩ needed to minimise a 38-particle LJ system from a random configuration. The stopping condition is that the maximum force on any atom is less than 0.01 in reduced units.

upon previous work, where the focus was mainly on transition state searches [96, 98]. Some of the more complex algorithms (CG, BFGS, L-BFGS and FIRE) depending on the choice of parameters produce basins that consist of disconnected parts. Such basins deviate from the "correct" basin of attraction defined by SD pathways, especially at the basin boundaries, where complex interpenetrating patterns can appear. These patterns generally do not disappear as the length scale is reduced, as can be seen in Figure 5.4 and Figure 5.5, making the basins ill-defined. In particular, overestimates for the step size are primarily responsible for the complex basin boundaries. Imposing a maximum step size can mitigate this problem for some algorithms, at the cost of slightly higher computational effort and an additional, system dependent, parameter. An appropriate value for the maximum step size can be chosen based on length scales in the system: for atomic systems, a good choice is about one tenth of the inter-atomic pair equilibrium distance.

In conclusion, if assignment of a starting configuration to the basin of attraction defined by SD is important, then FIRE may be the most convenient algorithm, due to its speed and precision, provided that an acceptable maximum step size is chosen. If finding a minimum quickly is more important, then L-BFGS is clearly the best choice.

Chapter 6

Calculating the entropy

In this Chapter, a numerical approach is used to compute the number of ways in which N particles can pack into a given volume V. This technique extends the method of Xu *et al.* [110] in such a way that it can be applied to much larger systems than before (over 100 particles). Many of the caveats of this method, discovered in Section 4.5, are addressed. Using this novel approach, the system size dependence of the number of distinct packings of a system of poly-disperse soft disks is studied in a regime where direct enumeration would fail completely. We show that, even though granular particles are distinguishable, we have to deal with the Gibbs paradox. The simulations presented here provide strong evidence that the packing entropy, when properly defined, is extensive. This extensivity has often been assumed but, thus far, direct evidence was lacking. We argue that, as different packings are created with unequal probabilities, the packing entropy should not be expressed as the logarithm of the number of packings but as $S = -\sum_i p_i \ln p_i$. We show that we can compute this quantity reliably and that, it too, is extensive.

We focus on the question of is it possible to define and compute a meaningful 'granular entropy' that is extensive? We note that the concept of a granular entropy was introduced over 20 years ago by Edwards [24] as the basic concept in the 'thermodynamic' theory of powders. Edwards suggested that all jammed states at a given packing fraction are equally likely and that the logarithm of the number of packings provides a definition of the granular entropy. We note that Edwards hypothesis includes several ingredients: 1) the assumption that all packings occur with equal frequency and 2) the assumption that the granular entropy is extensive. Over the years, the validity of Edwards' hypothesis has been hotly debated (this has been discussed in Chapter 2), but one thing has, to our knowledge, never been done: computing it explicitly for an off-lattice model, and for a number of particles that is large enough to study the system-size dependence of the granular entropy.

As previously mentioned, for relatively simple systems, the number of jammed configurations has been estimated. This was done for the Kob-Andersen lattice model [7]. For off-lattice models, the calculation is more challenging: estimates of the number of distinct minima in the energy landscape of very small systems of soft particles have been obtained by direct enumeration [109, 110] as seen in previous Chapters. However, such a direct enumeration becomes intractable for larger systems and, in addition, it suffers from an inherent sampling bias that privileges minima with large basins of attraction. If the number of packings can be counted, a second question can be addressed: is the granular entropy (i.e. the logarithm of the number of packings) extensive (i.e. if $\ln \Omega(N, \phi) \sim N$)?

In this Chapter, an estimate of the entropy of quasi-rigid disks over a range of system sizes is presented. To do so, the method introduced in Chapter 4 has been extended and modified in order to get an estimate of both $\Sigma(\phi)$ and the probability to be in the basin of any given minimum for a wide range of system sizes. In the thermodynamic limit, this method allows us to get an estimate of both $\ln \Omega(N, \phi)$ and the probability to be in the basin of any given minimum. Our main conclusions are: 1) the probability to observe different packings is not a constant 2) we can nevertheless define a granular entropy that is extensive, but it is not simply $-\sum_i p_i \ln p_i$.

6.1 Details of the model

In order to compute the number of distinct minima $\Omega(N, \phi)$ on the energy landscape of a system of N particles at packing fraction ϕ, the techniques described in Chapter 4 are used with modifications put in place to improve the issues mentioned in Section 4.5.

The mean volume $\langle v \rangle (N, \phi)$ of basins of attraction of minima on this energy landscape can be defined via:

$$\langle v \rangle (N, \phi) \equiv \frac{1}{\Omega(N, \phi)} \sum_{i=1}^{\Omega(N,\phi)} v_i = \frac{V_{\text{acc}}(N, \phi)}{\Omega(N, \phi)}, \tag{6.1}$$

where v_i is the volume of basin $i \in [1, \Omega(N, \phi)]$. The basins tile the accessible configuration space of the hard-core parent system. This accessible phase-space volume $V_{\text{acc}}(N, \phi)$ is, except for constant factors, equal to the partition function of the hard-core fluid. This volume can be computed if the free energy of the hard-core fluid can be determined. By inverting Equation (6.1), we obtain $\Omega(N, \phi) = V_{\text{acc}}(N, \phi)/\langle v \rangle(N, \phi)$. Hence $\Omega(N, \phi)$ can be estimated if we know $\langle v \rangle(N, \phi)$. The important point is that the latter average can be obtained by sampling a small subset of all possible jammed configurations. In practice, we first generate $\mathcal{N}_f$ well-equilibrated configurations of the hard-core fluid and perform a Stillinger quench on each of them. As explained in [110], the volume $v_i(N, \phi)$ of a basin i can be calculated using thermodynamic integration from a harmonic reference state

The modifications made to the algorithm described in Chapter 4 include:

- The use of the Fast Inertial Relaxation Engine (FIRE) minimisation algorithm, based on the results of Chapter 5.
- A new criterion to correctly find floater particles in the system, presented in Section 6.1.2.
- Polydispersity to further reduce the symmetries of minima.
- New interaction potential with Hard Sphere (HS) cores. This improves the sampling of the mean square displacement u^2 for $k = 0$, described in Section 6.1.1.
- Initial conditions now given by an equilibrated HS fluid. More “physical” configurations. Closer to granular media.

The parameters used in the simulations performed in this Chapter are shown in Table 6.1.

6.1.1 Hard cores with strongly repulsive shell

In this section we describe the interaction potential devised to improve the equilibration times for sampling the mean square displacement within the basins of attraction.

The Weeks-Chandler-Andersen (WCA) potential [104] is a short-range repulsive potential that results from truncating the Lennard-Jones (LJ) potential at the minimum ($2^{1/6}\sigma$) and shifting by the depth of the well, ε.

Parameter	**Symbol**	**Values**
Dimension of space	d	2
Number of particles	N	$8 \to 128$
Volume/packing fraction	ϕ	$0.86 \to 0.9$
Size of square box (volume$^{1/d}$)	L	1
Polydispersity (disk area)	δ	0.2
Total number of steps (MC and PT)	n_{step}	$10^5 \to 10^5$
Number of equilibration steps (MC and PT)	n_{eq}	$10^4 \to 10^5$
Number of steps (Direct sampling)	n_{DS}	$10^4 \to 10^5$
Gradient tolerance (minimisation)	`EPS`	10^{-7}
Position tolerance (comparing minima)	r_{tol}	$10^{-5}/\sqrt{N}$
Shell thickness (HS+WCA potential)	$d_{ij}^{S}/d_{ij}^{\text{HD}}$	1.12, 1.4

Table 6.1: Parameters used in simulations.

A modified version of the WCA potential is used here. A HS core is included and therefore the WCA potential is shifted so that it diverges at the distance between the HS cores:

$$V_{\text{HS+WCA}}(r_{ij}) = \begin{cases} \infty & r_{ij} \le d_{ij}^{HS} \\ 4\varepsilon \left[\left(\frac{\sigma_{ij}}{r_{ij} - d_{ij}^{\text{HD}}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij} - d_{ij}^{\text{HD}}} \right)^{6} \right] + \varepsilon & d_{ij}^{\text{HD}} < r_{ij} < d_{ij}^{S} \\ 0 & r_{ij} \ge d_{ij}^{S} \end{cases} \tag{6.2}$$

where $d_{ij}^{S} = d_{ij}^{\text{HD}} + 2^{1/6}\sigma_{ij}$ and $\sigma_{ij} = (2^{5/6}/5)d_{ij}^{\text{HD}}$. The distances are shown in Figure 6.1 (left). For poly-disperse systems the width of the potential is different for each pair of particles. The shape of this potential is shown in Figure 6.1 (right).

We call this potential Hard Sphere core plus Weeks-Chandler-Andersen (HS+WCA). The initial conditions correspond to a configuration taken from an equilibrated (to ensure uniform sampling of configuration space) hard sphere fluid at the density of the hard sphere cores of the particles.

The volumes (areas) of the particles in the poly-disperse systems were taken from a Gaussian distribution with standard deviation 0.2 and mean 1.0 and then re-scaled in order to obtain the correct volume fraction ϕ. The use of poly-disperse particles not only makes crystallisation less likely it also removes some of the symmetry present in the system meaning that no two basins have exactly the same structure and energy as any permutation of two

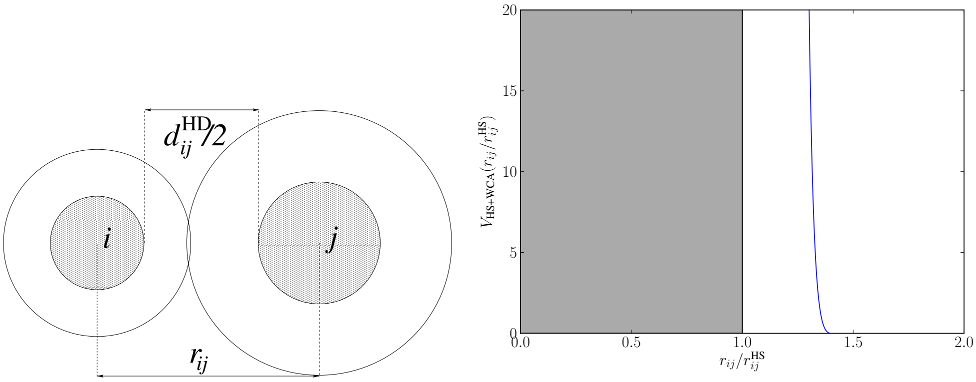


Figure 6.1: Interaction between particles for HS+WCA potential. Left: Diagram showing the distances that determine the value of the potential energy between two overlapping particles. Right: Plot of the potential energy $V_{\text{HS+WCA}}(r_{ij})$ as a function of the distance between particles i and j, r_{ij} in Equation (6.2).

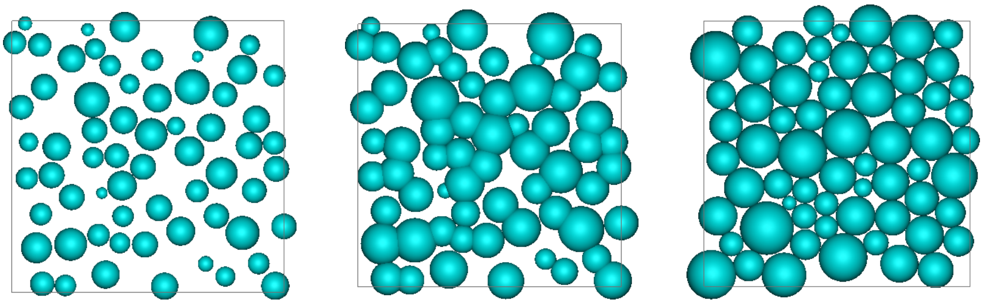


Figure 6.2: 2D Poly-disperse 64 particle system with $\phi_{\text{HD}} = 0.45$ and $\phi_{\text{HS+WCA}} = 0.88$. Left: Initial HS fluid. Middle: Initial HS fluid plus WCA shell. Right: Potential energy minimum.

or more particles will lead to a different structure.

It is important to note that the system is only allowed to sample configurations that belong to the HS fluid defined by the hard cores. This means that there can never be a move that would make the hard cores overlap. This is true during all of the Monte Carlo (MC) sampling and the minimisation algorithm. The FIRE algorithm has been modified to accommodate this situation because the original method takes steps that cause an increase in potential energy and ultimately an overlap of the HS cores (for implementation details see Section A.5.2). The initial conditions and a snapshot corresponding to a potential-energy minimum can be seen in Figure 6.2.

6.1.2 Floaters or rattlers

Once a jammed configuration is found, it is possible that some of the particles do not form part of the rigid structure and are free to move within a cage defined by neighbouring particles. These free particles are called *floaters* or *rattlers*: they can move (over short distances) without energy cost. In a perfectly minimised system, this definition is equivalent to having no overlapping contact particles. As the minimisations performed are not perfect in the sense that there are always some overlaps left even though they are very small, a more relevant definition for floaters has been devised that is an improvement over the method that relies only on counting the number of contacts, described in Section 4.3.4.

Counting the number of contacts or overlaps a particle has is not good enough if the system is not mono-disperse because all of the contacts could be on a single hemisphere of the particle. this means that the particle is free to move in the direction opposing that hemisphere.

An algorithm has been developed that is based on the sum of the angles between successive contacts and is able to correctly identify floaters in 2-dimensional systems.

For each particle in the system the first step is to identify its contacts and create a list of the particle numbers of these contacts. Starting from the first contact in this list, which is called contact A, the nearest contact is found and named B. This chooses a direction that is defined by the sign of $\overrightarrow{OA} \times \overrightarrow{OB}$, see Figure 6.3. All of the angles between successive contacts are measured and added together:

$$\sum \theta = \arccos\left(\frac{\overrightarrow{OA} \cdot \overrightarrow{OB}}{|\overrightarrow{OA}||\overrightarrow{OB}|}\right) + \cdots + \arccos\left(\frac{\overrightarrow{OX} \cdot \overrightarrow{OA}}{|\overrightarrow{OX}||\overrightarrow{OA}|}\right) \tag{6.3}$$

where X is the last contact. If all of the contacts are on the same hemisphere then one of the angles will be larger than π but the dot product will always measure the interior angle between two vectors such that $\sum \theta < 2\pi$ when the particle is a floater and $\sum \theta = 2\pi$ when the particle is not a floater. These situations can be seen in Figure 6.3, on the left a floater with all contacts on the same hemisphere. Clearly the particle is free to move in the direction of the arrow and its overlaps are determined by the stopping criteria of the energy minimisation algorithm. On the right side of Figure 6.3 we can

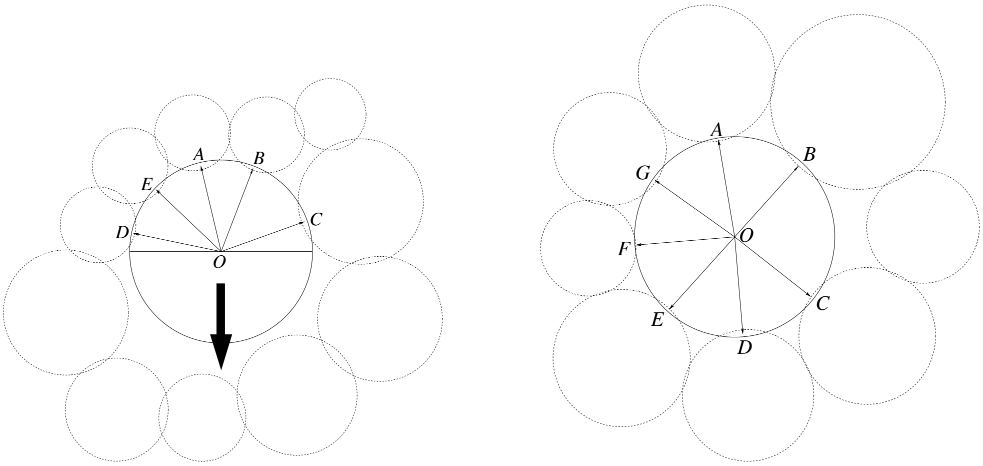


Figure 6.3: How to detect floaters. Left: Floater with all contacts on the same hemisphere, free to move on the direction of the arrow. Right: Stable, trapped particle with contacts on both hemispheres.

see a non-floater particle that is trapped by its contacts which are on both hemispheres.

6.2 Results

In this section the results of this Chapter are presented. The behaviour of the square displacement and its mean value are used to implement the thermodynamic integration method using the Gauss-Lobatto (GL) quadrature. Later the basin volume distributions are discussed and the resulting entropy is revealed.

6.2.1 Behaviour of the mean square displacement

The instantaneous square displacement as a function of steps is presented in Figure 6.4. The curves, from top to bottom correspond to $k = 0$, 1154.08, 6070.63, 22842.44, 62103.68 (for $N = 32$, $\phi = 0.9$, $\delta = 0.2$, $d = 2$, HS+WCA potential with $d_{ij}^{S}/d_{ij}^{\mathrm{HD}} = 1.12$). This is a log-log plot. As we can see the higher the value of k, the lower the displacement.

The problem of slow convergence of the mean square displacement for $k = 0$, seen in Figure 4.5 and discussed in Section 4.4, has improved drastically by using the new HS+WCA potential described in Section 6.1.1. This can be seen by looking at the square displacement u^2 as a function of steps for

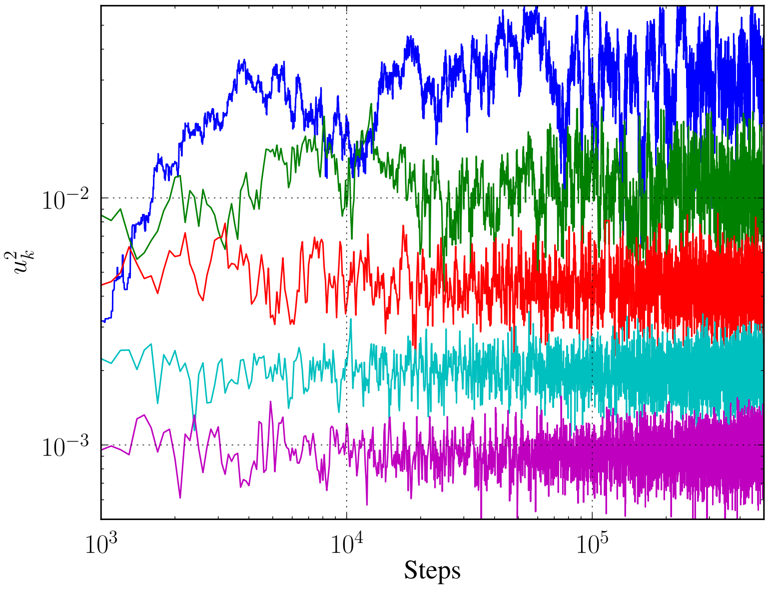


Figure 6.4: Log-log plot of u^2 vs steps for different values of k for the same minimum. From top to bottom, $k = 0$, 1154.08, 6070.63, 22842.44, 62103.68 ($N = 32$, $\phi = 0.9$, $\delta = 0.2$, $d = 2$, HS+WCA potential with $d_{ij}^S/d_{ij}^{\text{HD}} = 1.12$).

different MC runs starting from the same initial point, in this case the minimum. If the value of u^2 equilibrates then each different run should converge to the same average value. We have studied the equilibration of u^2 at $k = 0$ for 10 independent MC runs all sampling the basin of attraction of the same minimum. This procedure was carried out for both of the potentials in question, namely the soft harmonic interactions described in Section 4.3.1 and the HS+WCA potential described in Section 6.1.1. Table 6.2 shows the average $\langle u^2 \rangle_{k=0}$ for both systems and the respective deviation from the mean taken over all 10 runs. Figure 6.5 shows the evolution of $u_{k=0}^2$ as a function of MC steps. Although on average the relative deviation from the mean is similar for both potentials, we can see that, in the case of the soft potential, some of the runs get trapped in particular branches of the basin and never leave (at least not during the simulation time, for example the two green curves and the yellow curve in the top plot in Figure 6.5). This means that for different runs, even long ones, we could find drastically different values of the mean square displacement. In the case of the HS+WCA potential the square displacement fluctuates around a universal average value and even if these fluctuations are large, the average tends towards the same value.

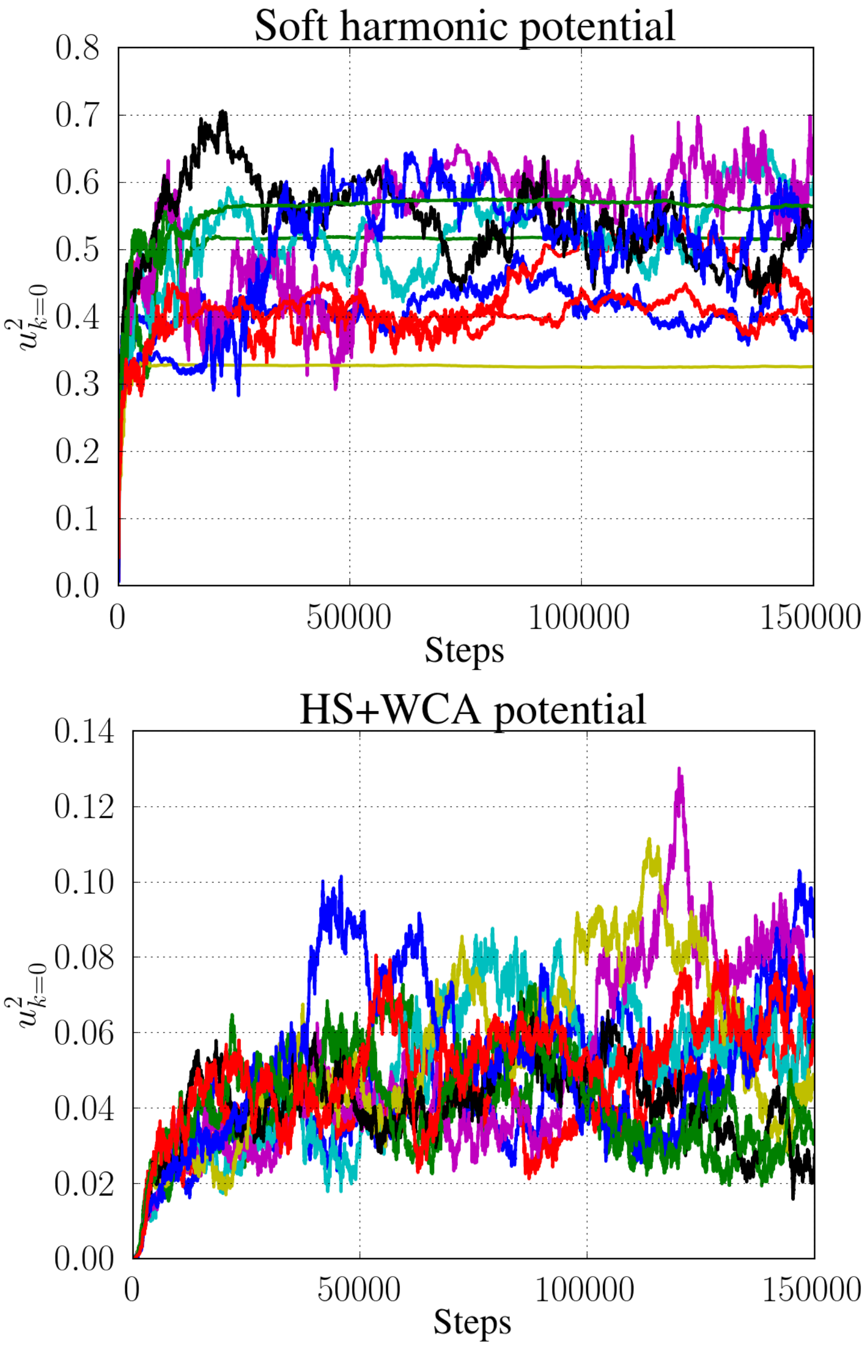


Figure 6.5: Square displacement u^2 as a function of MC steps for $k = 0$ for 10 different runs starting from the same minimum. The average values are shown in Table 6.2. The upper panel shows the results for the soft harmonic potential and the lower panel shows results for the HS+WCA potential.

HS+WCA	Deviation	Soft	Deviation
0.0535	0.00123	0.419	0.07144
0.0406	0.01168	0.518	0.02748
0.0455	0.00671	0.458	0.03244
0.0598	0.00754	0.539	0.04918
0.0609	0.00864	0.593	0.10320
0.0648	0.01252	0.327	0.16326
0.0419	0.01029	0.524	0.03423
0.0556	0.00332	0.546	0.05596
0.0419	0.01032	0.570	0.08025
0.0580	0.00575	0.407	0.08315

Table 6.2: Comparison of the mean square displacement for $k = 0$, $\langle u^2 \rangle_{k=0}$, for 10 different runs starting from the same minimum for the HS+WCA system and the soft harmonic system for $N = 64$, $\phi = 0.88$, $\delta = 0.2$, and $d_{ij}^S / d_{ij}^{\text{HD}} = 1.12$ for the HS+WCA potential. The deviation from the mean value is also shown.

6.2.2 Behaviour of $\langle u^2 \rangle$ vs k

In this section the results for the mean square displacement are presented, $\langle u^2 \rangle$, calculated from the MC simulations. In Figure 6.6 we can see the points calculated from simulations (red points), the harmonic approximation (blue curve) and the approximation using κ, the effective spring constant at $k = 0$ (Equation (4.16)) (green curve).

In Figure 6.7 we can see the GL integrand, Equation (4.18), as a function of the integration variable t from Equation (4.17). This should be a smooth and slowly varying function so that the GL integration is accurate, as the systematic quadrature error is proportional to the tenth derivative of the function being integrated. A rough estimate of the statistical error in the integral is $\sqrt{\sum w_i^2 \sigma_i^2}$ where w_i are the weights given by the GL method and σ_i^2 correspond to the variance of the MC sampling of the integrand shown in Figure 6.7. For the example case shown in Figure 6.6 and Figure 6.7, the value of the integral is 252.057 with an associated error of ± 15.072.

6.2.3 Free energy distributions

In this section the results of the study of the free energy distributions of the HS+WCA system described in Section 6.1.1 for two shell thicknesses $d_{ij}^S / d_{ij}^{\text{HD}}$ are presented. In what follows, *thick shell* refers to $d_{ij}^S / d_{ij}^{\text{HD}} = 1.4$

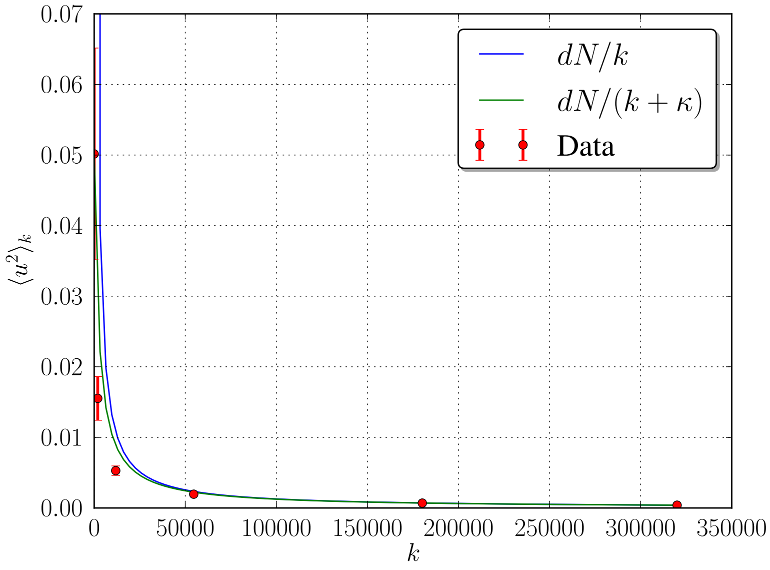


Figure 6.6: Mean square displacement $\langle u^2 \rangle_k$ as a function of k, $N = 64$, $\phi = 0.88$, $\delta = 0.2$, $d = 2$, HS+WCA potential with $d_{ij}^S/d_{ij}^{\mathrm{HD}} = 1.12$. The red points are the data from the MC calculations. The blue curve is the harmonic approximation, $\langle u^2 \rangle_k = dN/k$. The green curve is the approximation $\langle u^2 \rangle_k = dN/(k+\kappa)$ from Equation (4.16).

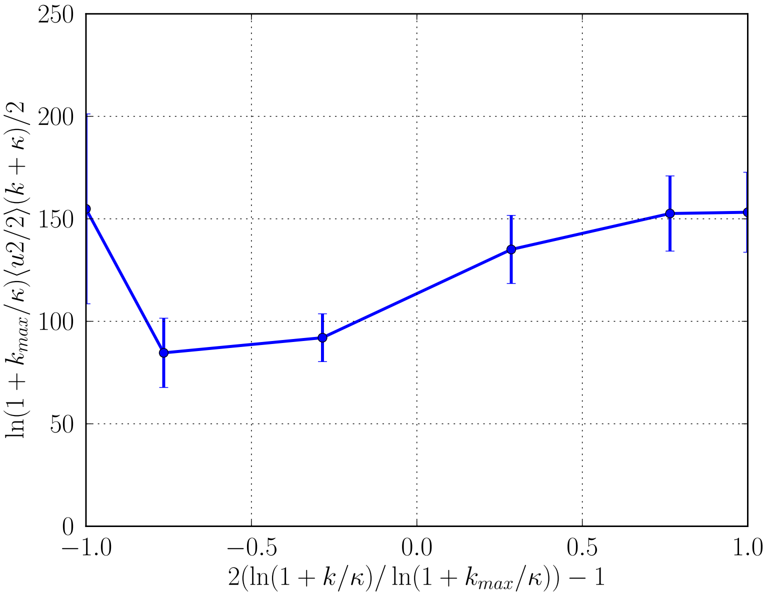


Figure 6.7: GL integrand (Equation (4.18)) as a function of t (Equation (4.17)) showing that it is a smooth function ($N = 64$, $\phi = 0.88$, $\delta = 0.2$, $d = 2$, HS+WCA potential with $d_{ij}^S/d_{ij}^{\mathrm{HD}} = 1.12$).

and *thin shell* refers to $d_{ij}^S/d_{ij}^{\mathrm{HD}} = 1.12$.

The results for $d_{ij}^S/d_{ij}^{\mathrm{HD}} = 1.4$ and $d_{ij}^S/d_{ij}^{\mathrm{HD}} = 1.12$ are presented in Figure 6.8 and Figure 6.9 respectively. The free energy histograms have been plotted using the volume calculation results of ~ 1000 minima.

6.2.4 The bias problem

Our first objective is to compute the number of distinct packings. After that, we turn to the granular entropy. We have to discuss granular entropy separately because the very fact that different basins have different volumes already implies that, as the probability with which we sample basins is proportional to their volume, all basins will not be equally populated. If this is the case, it is not meaningful to use Edwards' definition of the granular entropy: $S = \ln \Omega$. Rather, a more appropriate definition of entropy would be analogous to the information entropy: $S^* = -\sum_i p_i \ln p_i$. In fact, as we shall see, even this expression will have to be modified to obtain an extensive entropy (but this is no different in the canonical ensemble of statistical mechanics).

We now first focus on the computation of the number of distinct packings of a granular medium of N particles at a prescribed density. As previously mentioned in this Chapter, and in Section 4.4, we can compute this number if we know the *unbiased* distribution of basin volumes. Hence, we need to get rid of the bias from the 'as-sampled' volume distribution.

Such unbiasing was not relevant in earlier studies [109] where the small system sizes allowed direct enumeration of *all* minima. However, the bias becomes crucial for system sizes larger than $N = \mathcal{O}(10)$ where we can only sample a tiny fraction of all basins. In what follows, we will denote the biased and unbiased distributions of free energies by $\mathcal{B}(F|N,\phi)$ and $\mathcal{U}(F|N,\phi)$ respectively. Examples of such distributions are shown in Figure 6.10. The biased distribution of the $\mathcal{N}_f$ sampled volumes allows us to compute $\mathcal{B}(F|N,\phi)$ that, *in principle*, can be related to $\mathcal{U}(F|N,\phi)$ via:

$$\mathcal{U}(F|N,\phi) = \mathcal{C}(N,\phi)\mathcal{B}(F|N,\phi)e^F \ , \tag{6.4}$$

where $\mathcal{C}$ is a normalising constant. It is easy to see that:

$$\langle v\rangle(N,\phi) = \mathcal{C}(N,\phi) = \left[\int_{F_{\min}}^{+\infty} dF\ \mathcal{B}(F|N,\phi)e^F\right]^{-1} \ , \tag{6.5}$$

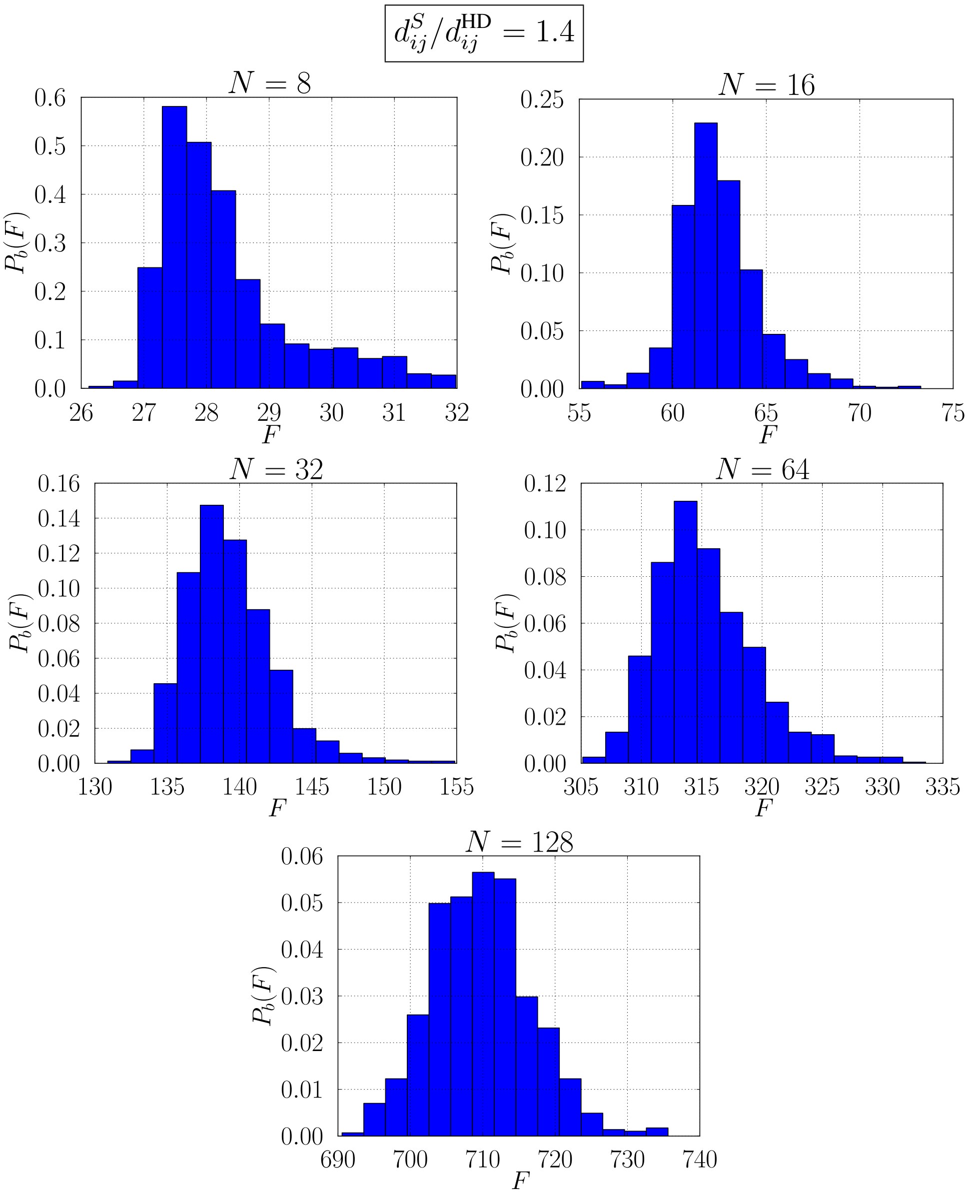


Figure 6.8: Histogram of free energies for ~ 1000 basins in two dimensions for different system sizes N. $\phi = 0.88$, $\delta = 0.2$, $d = 2$, HS+WCA potential with $d_{ij}^S/d_{ij}^{\mathrm{HD}} = 1.4$.

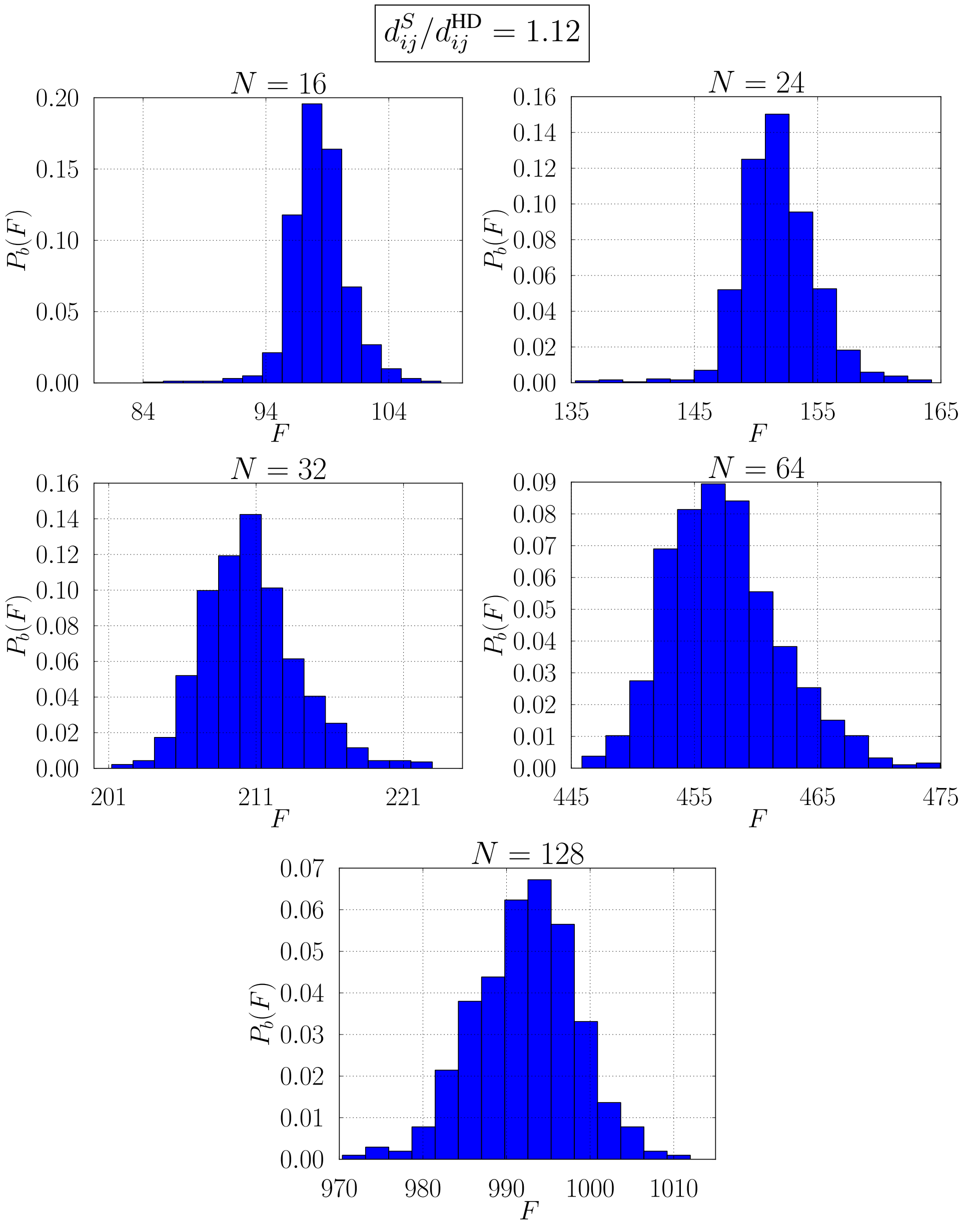


Figure 6.9: Histogram of free energies for ~ 1000 basins in two dimensions for different system sizes N. $\phi = 0.88$, $\delta = 0.2$, $d = 2$, HS+WCA potential with $d_{ij}^S/d_{ij}^{\mathrm{HD}} = 1.12$.

which is the expression used below.

Following the approach described in Chapter 4, with the modifications presented in Section 6.1, the basin-volume weighted distribution $\mathcal{B}(F|N,\phi)$ of free energies as a function of N can be obtained for the thick and thin shell systems (see Figure 6.8 and Figure 6.9).

Because small basin volumes are inadequately sampled, Equation (6.4) cannot be applied directly to obtain an unbiased distribution. In fact, from Figure 6.8 and Figure 6.9, it is clear that there is a typical spread in free energies ΔF of at least 20 in the sampled basins. Since only $\mathcal{N}_f \sim 1000$ minima are sampled in the histogram, it is implied that the most probable –biased– basins are about $\mathcal{O}(10^3)$ more probable than the smallest ones. Upon unbiasing, this ratio is multiplied by a factor e^{-20} and thus the smallest basins (with high free energies) observed in Figure 6.8 and Figure 6.9 are in fact overwhelmingly more numerous than the large ones. As a consequence, a large fraction of the unbiased distribution is obscured by noise and cannot therefore be normalised.

The seriousness of this problem can be seen in Figure 6.10. The left hand side shows the histogram calculated directly from the simulation data, namely $\mathcal{B}(F|N,\phi)$. The right hand side (blue points and curve) shows the unbiased distribution calculated only from the simulation data via Equation (6.4). The green curve shows what an unbiased distribution would look like if the full distribution of basins would have been sampled in simulations. This curve was obtained by fitting the simulation data using the procedure described below in Section 6.2.5.

To arrive at an estimate of the unbiased distribution $\mathcal{U}(F|N,\phi)$, we make use if the fact that the number of basins, even if very large, is finite. Hence, both the biased and unbiased distributions must be normalisable. As we show below, we find that our simulation data is compatible with the assumption that, for large enough F, $\mathcal{B}(F|N,\phi)$ decays as $e^{-cF^{\nu}}$ where $\nu > 1$ and $a, c > 0$. Furthermore, we assume that the unbiased distribution $\mathcal{U}$ is unimodal. This assumption is supported by the direct enumeration results for small systems [110]. The observed distributions $\mathcal{B}$ are therefore fitted with a 3-parameter Generalised Normal Distributions, described in the next section.

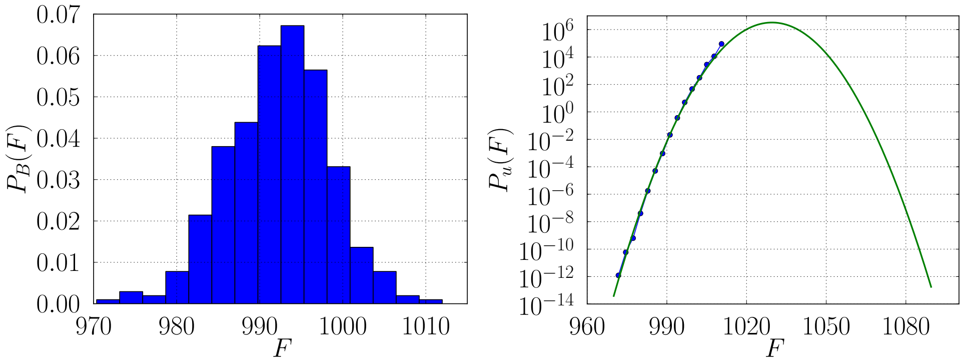


Figure 6.10: Histogram of free energies for 1000 basins in two dimensions for $N = 128$, $\phi = 0.88$, $\delta = 0.2$, $d = 2$, HS+WCA potential with $d_{ij}^S/d_{ij}^{\text{HD}} = 1.12$. Left: biased distribution as sampled in simulations. Right: unbiased distribution. The green curve shows what an unbiased distribution would look like if the full distribution of basins were sampled in simulations. This curve was obtained by fitting the simulation data using the procedure described in Section 6.2.5.

6.2.5 Fitting the free energy distribution

We have fitted the distributions to a functional form and calculated the error bars using the bootstrapping technique outlined in Section 6.2.6. Figure 6.10 shows that, using a Generalised Normal Distribution, we can fit the unbiased distribution and subsequently unbias it by dividing the probability density by the volume $V = \exp(-F)$. Without a functional fit, unbiasing is not possible. In 'thermal' Monte Carlo simulations, unbiasing is often possible using a technique such as umbrella sampling (see e.g. [32]). The reason for this difference is that in normal Metropolis sampling, points that are close in configuration space tend to have similar order parameters that can be used in the biasing. In contrast, adjacent basin volumes are not spatially correlated in such a way that the sampling in configuration space can easily be biased towards, say, smaller volumes. It is important to note that, because we do not sample the complete distribution in our simulations, we *must* fit the distribution in order to normalise it. To illustrate this point, we show in Figure 6.10 a free energy distribution that has been unbiased by dividing the probability density by the volume $V = \exp(-F)$.

The free energy distributions obtained using the method described above are fitted to a Generalized Gaussian Distribution (GGD). The probability den-

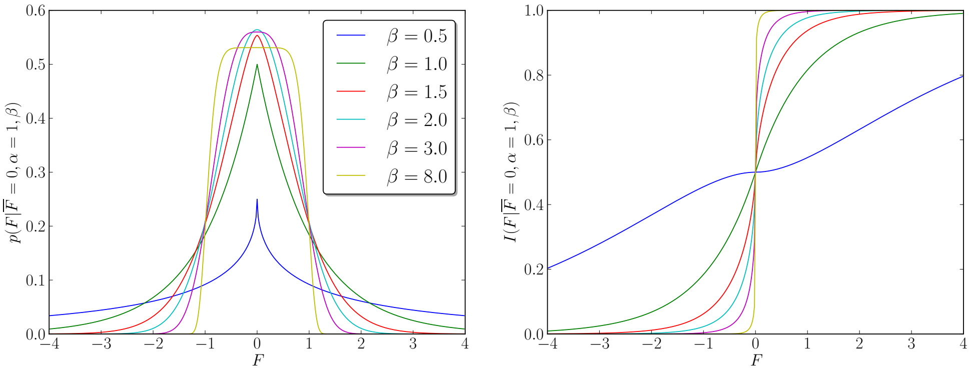


Figure 6.11: Generalised Gaussian Distribution for different values of β with $\alpha = 1$ and $\overline{F} = 0$. Left: probability density function, Equation (6.6). Right: cumulative density function, Equation (6.7).

sity function is given by:

$$p(F|\overline{F}, \alpha, \beta) \equiv \frac{\beta}{2\alpha\Gamma(1/\beta)} e^{-\left(\frac{|F-\overline{F}|}{\alpha}\right)^{\beta}} \tag{6.6}$$

and the cumulative density function is:

$$I(F|\overline{F}, \alpha, \beta) \equiv \frac{1}{2} + \text{sgn}(F - \overline{F}) \frac{\gamma\left[1/\beta, \left(\frac{|F-\overline{F}|}{\alpha}\right)^{\beta}\right]}{2\Gamma(1/\beta)} \tag{6.7}$$

where Γ is the Euler Gamma Function defined by

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t}\, dt,$$

γ is the Incomplete Gamma Function, defined by

$$\gamma(z, x) = \int_0^{x} t^{z-1} e^{-t}\, dt,$$

$\overline{F}$ is the mean, α^2 is a measure for the variance and β is a parameter that determines the shape of the distribution, as can be seen in Figure 6.11. When $\beta = 2$ we get a Gaussian (Normal) distribution with variance $\alpha^2/2$ and if $\beta = 1$ we recover the Laplace distribution. For more information see [105] and references therein.

It is important to stress the fact that the fitted distribution must decay faster than the bias imposed by sampling each basin with a probability proportional to its volume. This means that the fitted β must be larger than one.

The `scipy.optmize.curve_fit` routine from [40] is used to fit the cumulative distribution (Equation (6.7)). This routine fits using the Levenberg-Marquardt algorithm [50, 58].

The results for $d_{ij}^{S}/d_{ij}^{\text{HD}} = 1.4$ and $d_{ij}^{S}/d_{ij}^{\text{HD}} = 1.12$ are presented in Figure 6.12 and Figure 6.13 respectively. The free-energy histograms have been plotted using the volume calculation results of ~ 1000 minima.

In Figure 6.14 the system-size dependence of the best fit parameters α^2 and β is shown. We note that α^2 appears to become extensive with N while β tends to be 2 in the large-N limit. We also note that it is far from obvious that the variance of the free energy distribution should be extensive.

In the thermodynamic limit, $\mathcal{B}(F|N, \phi)$ apparently tends to a Gaussian, an observation that is compatible with the findings of [110] for a much smaller system. Using the fitted functional form for the volume distribution, we can now proceed to estimate the unbiased distribution of volumes and use Equation (6.5) to obtain an estimate for the average basin volume and, from that, of the number of distinct packings.

6.2.6 Bootstrapping

The *bootstrapping* technique is used to calculate the errors of the parameters $\overline{F}$, α, β from Equation (6.6) and Equation (6.7) fitted to the free energy distributions of jammed states (see Section 4.4) obtained from the MC simulations.

Bootstrapping is a re-sampling technique that consists of constructing a new data set by randomly sampling with replacement from the original data [26]. If this process is repeated many times, we will have a distribution for each parameter. The mean value of these distributions is assigned to the parameter and the standard deviation is the error.

6.2.7 Configurational entropy

To obtain an estimate of the number of packings, we combine our information about $\langle v \rangle_{\phi,N}$ with our knowledge of the accessible volume of the parent system (i.e. the poly-disperse hard-disk fluid):

$$-\ln V_{\text{acc}}(N, \phi_{\text{HD}}) = -\ln V^N + N f_{\text{ex}}(\phi_{\text{HD}}) \tag{6.8}$$

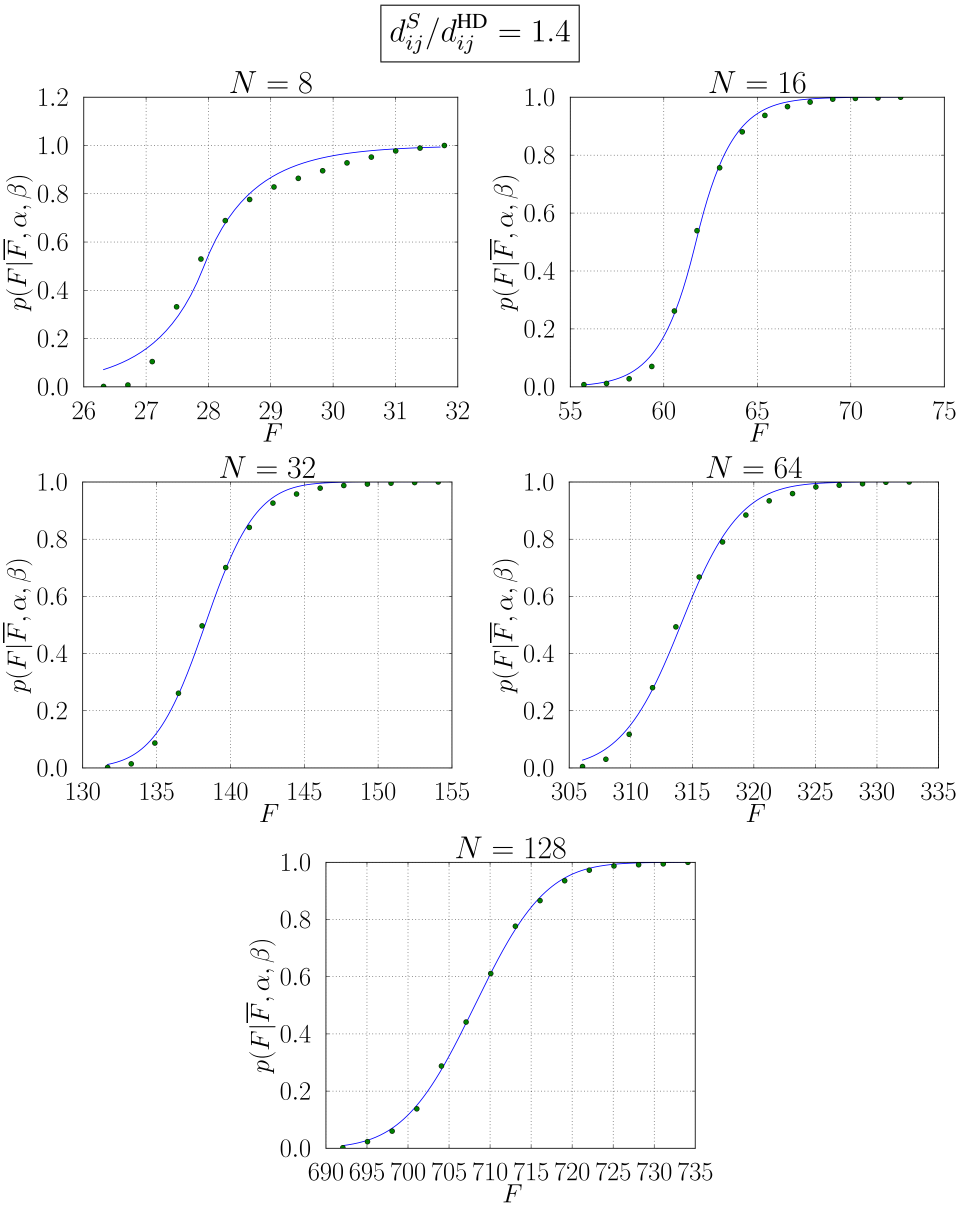


Figure 6.12: Fitted cumulative histogram of free energies for ~ 1000 basins for different system sizes N. $\phi = 0.88$, $\delta = 0.2$, $d = 2$, HS+WCA potential with $d^S_{ij}/d^{\mathrm{HD}}_{ij} = 1.4$.

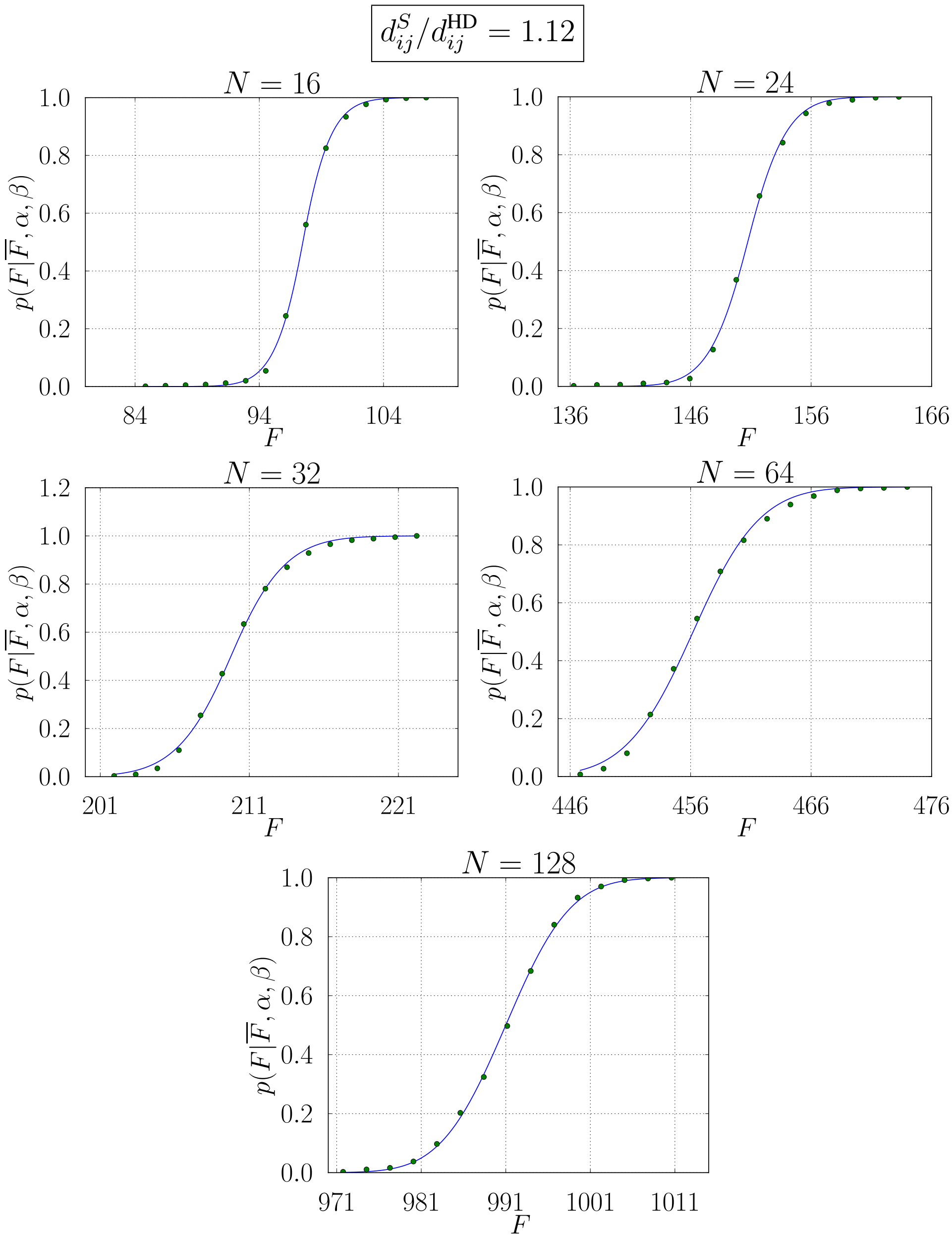


Figure 6.13: Fitted cumulative histogram of free energies for ~ 1000 basins for different system sizes N. $\phi = 0.88$, $\delta = 0.2$, $d = 2$, HS+WCA potential with $d_{ij}^S/d_{ij}^{\mathrm{HD}} = 1.12$.

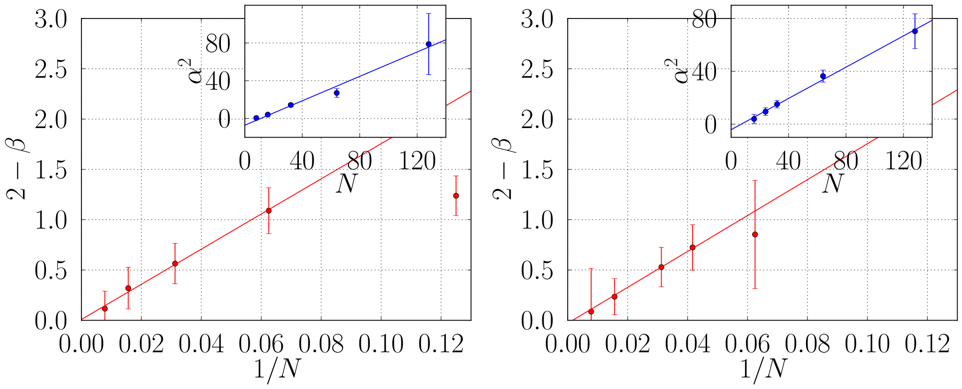


Figure 6.14: Fitted generalised Gaussian distribution parameters for a HS+WCA system with $d^S_{ij}/d^{\text{HD}}_{ij} = 1.4$ (Left) and $d^S_{ij}/d^{\text{HD}}_{ij} = 1.12$ (Right) for $\phi = 0.88$, $\delta = 0.2$, $d = 2$. $2-\beta$ plotted as a function of the inverse system size $1/N$ (points) and compared to a linear function. The inset shows the best fitting parameter α^2 as a function of N.

where the first term on the right hand side is the contribution from an ideal gas of distinguishable particles where V is the volume of the box (which, in our units is equal to one). $f_{\text{ex}}(\phi_{\text{HD}})$ is the excess free energy term of the hard-disk fluid at volume fraction ϕ_{HD}.

Our estimate for $\Omega(N,\phi)$ is then obtained by combining Equation (6.1), Equation (6.5) and Equation (6.8). Coming back to Equation (6.8), since we work with a unit box, only f_{ex} needs to be computed. We can get it directly via thermodynamic integration:

$$f_{\text{ex}}(\phi_{\text{HD}}) = \int_0^{\phi_{\text{HD}}} d\phi \, \frac{(Z(\phi)-1)}{\phi} \tag{6.9}$$

where $Z(\phi) \equiv P/\rho k_B T$ is the compressibility factor of the poly-disperse hard-disk fluid. For poly-disperse hard disks, Z is well approximated by

$$Z(\phi) = pZ_m(\phi) + \phi(1-p)/(1-\phi) \tag{6.10}$$

from [77], where $p \equiv \sqrt{\langle d^2_{\text{HD}}\rangle}/\langle d_{\text{HD}}\rangle$ and [54]:

$$Z_m(\phi) \approx \left[1 - 2\phi - \frac{\phi^2(1-2\phi_{\text{max}})}{\phi^2_{\text{max}}}\right]^{-1} \tag{6.11}$$

where $\phi_{\text{max}} = \pi/\sqrt{12}$.

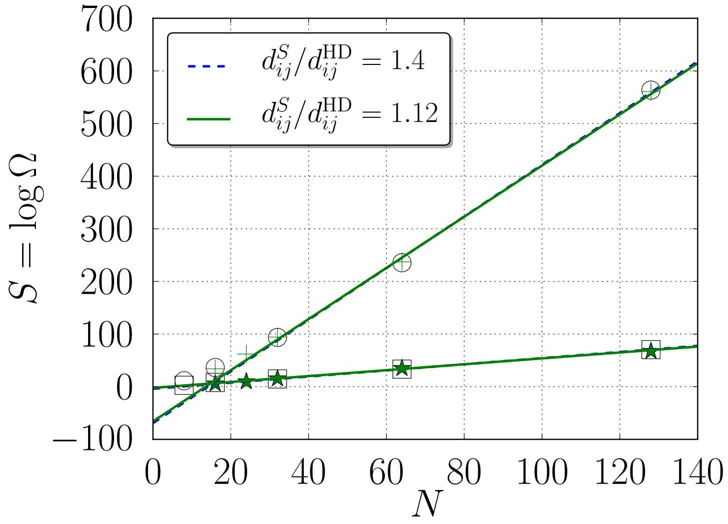


Figure 6.15: Configurational “Edwards” entropy for $\phi = 0.88$ and two different $d^S_{ij}/d^{\mathrm{HD}}_{ij}$. $S(N) = \ln \Omega(N)$ (open circles and crosses for $d^S_{ij}/d^{\mathrm{HD}}_{ij} = 1.4$ and 1.12 respectively) calculated using the fitted free energy distributions. As the straight-line fit shows, $S(N)$ is not extensive while $S(N) - \ln N!$ (open squares and stars for $d^S_{ij}/d^{\mathrm{HD}}_{ij} = 1.4$ and 1.12 respectively) appears to be extensive. We also note that the results for systems prepared with $d^S_{ij}/d^{\mathrm{HD}}_{ij} = 1.4$ and 1.12 are virtually indistinguishable.

Figure 6.15 shows the direct numerical calculations of $S(N)$ from Equation (6.1), Equation (6.5) and Equation (6.9). It is perhaps redundant to stress that the number of granular packings that we find for larger systems is very large. For instance, for 128 particles, we find 10^{250} distinct packings. It should be clear that direct enumeration would be utterly useless to compute such a large number.

The dashed lines show the limiting behaviour that we should expect if S is asymptotically linear in N. As can be seen from the upper curves of Figure 6.15, we observe large deviations from such a linear relation for systems containing fewer than $\mathcal{O}(30)$ particles. We will find a similar problem if we define $S = -\sum_i p_i \ln p_i$. Below, we shall argue that the analogy with the statistical mechanics of classical particles suggests that extensivity can only be achieved if we divide Ω by $N!$.

Before we discuss the extensivity of the granular entropy, we first focus on the second issue raised at the beginning of Section 6.2.4, namely that it is inconsistent to use $\ln \Omega$ as a measure of the entropy if the different packings of the system are not equally populated. In that case it is advisable to replace the

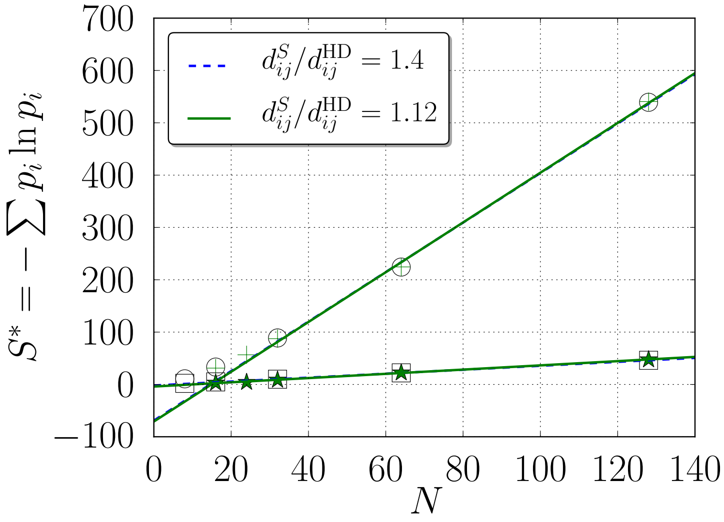


Figure 6.16: Redefined configurational entropy for $\phi = 0.88$ and two different $d^S_{ij}/d^{\mathrm{HD}}_{ij}$. $S^*(N) = -\sum_i p_i \ln p_i$ (open circles and crosses for $d^S_{ij}/d^{\mathrm{HD}}_{ij} = 1.4$ and 1.12 respectively) calculated using the fitted free energy distributions. As the straight-line fit shows, $S^*(N)$ is not extensive while $S^*(N) - \ln N!$ (open squares and stars for $d^S_{ij}/d^{\mathrm{HD}}_{ij} = 1.4$ and 1.12 respectively) appears to be extensive. We also note that the results for systems prepared with $d^S_{ij}/d^{\mathrm{HD}}_{ij} = 1.4$ and 1.12 are virtually indistinguishable.

micro-canonical expression for the entropy ($S = \ln \Omega$) with the 'canonical' form

$$S^* \equiv -\sum_{i=1}^{\Omega} p_i \ln p_i \tag{6.12}$$

where p_i denotes the probability to generate packing number i. Of course, if all packings are equally likely, Equation (6.12) reduces to the 'micro-canonical' form $S = \ln \Omega(N, \phi)$. However, in general, it is different. For the present procedure to generate packings, it is known that $p_i = v_i/V_{\mathrm{acc}}$. Hence,

$$S^* = -\langle \ln v/V_{\mathrm{acc}} \rangle = \langle F \rangle_B - N f_{\mathrm{ex}}(\phi_{\mathrm{HD}}) \tag{6.13}$$

where the subscript B denotes an average over the volume-weighted, biased distribution. The important point to note is that this expression depends only on the basin volumes *as sampled* and hence requires no additional assumptions to achieve unbiasing. Hence, we can obtain reliable estimates for S^*.

When we plot S^* as a function of N (Figure 6.16), we note (again) that its N-dependence is not very linear, in other words, this form of S^* is also not extensive. This should come as no surprise because also in equilibrium statistical mechanics, the partition function of a system of N distinct objects is not extensive. We must follow Gibbs' original argument about the equilibrium between macroscopically identical phases of classical particles and this implies that we should subtract $\ln N!$ from S^*. We note that there is much confusion in the literature on this topic, although papers by van Kampen [93] and by Jaynes [39] clarify the issue. Indeed, if we plot $S^* - \ln N!$ versus N we obtain an almost embarrassingly straight line that, moreover, goes through the origin. Previous studies on the entropy of jammed systems, such as the simulations of Lubachevsky *et al.* [55] presented in Chapter 2, ignored the $N!$ term. We note that the assumption that all packings are equally likely was an unnecessarily strong condition to construct a granular entropy and, from that, granular thermodynamics. Edwards' hypothesis of the existence of a meaningful granular entropy therefore survives even when the condition that all granular packings are equally likely is dropped.

Our finding that Edwards' entropy, once properly defined, is indeed extensive is highly significant. Extensivity of the granular entropy is absolutely crucial if this quantity is to be used to define quantities such as the compactivity and angoricity [59] that play a central role in Edwards' theory of granular transport. Now that we can compute the entropy, we can start to test these theories. Of course, it would be interesting to test if other protocols to generate jammed structures also find an extensive granular entropy.

One interesting observation is the following: we find that the plots for the entropy of jammed packings at $\phi = 0.88$ that were generated from two very different parent systems (one with a short-ranged the other with a long-ranged WCA potential, thick and thin shells defined above) are almost on top of one another (see Figure 6.15 and Figure 6.16). Again, this finding is not obvious *a priori*. It seems to apply that effectively all minima that are generated in the thick-shell system with a low initial density are also permissible for the thin-shell, high-density parent system. We do not expect that such protocol independence of jammed structures at a given density will hold in general. However, it suggests that the results that are reported in this section may also apply to hard, rather than soft disks.

6.3 Conclusions

In this Chapter we reported a numerical study of the packing entropy of soft repulsive, poly-disperse disks. Using a novel technique to sample the volumes of the basins of attraction of potential-energy minima, it was possible to study the basin-volume distribution for much larger systems sizes than was hitherto possible (up to $N = 128$). We find that the distribution of basin volumes becomes log-normal in the large N limit.

Using our approach we can estimate the number of distinct packings of a system of poly-disperse hard disks up to system sizes of 128 particles. We find that the number increases approximately (but not exactly) exponentially with system size. For $N = 128$, we estimate that the number of distinct states is of order 10^{250}, a number well outside the reach of any direct enumeration scheme.

If, in our definition of the granular entropy, we take into account that different packings are not *a priori* equally likely, we should use the 'information' form for the entropy ($S^* \equiv \sum_i p_i \ln p_i$). This entropy can be computed accurately and without fitting to a particular form. We find that the behaviour of S^* is very similar to that of $\ln \Omega$ and even the numerical value differs but little.

We note that N-dependences of $\ln \Omega$ and S^* are not very linear. However, the same would be true of the logarithm of the partition function of a system of N distinct particles in classical statistical mechanics. In classical statistical mechanics, we must divide Ω by $N!$ to ensure that two systems in identical macroscopic states can be in equilibrium under the exchange of particles, even though these particles are not indistinguishable in the quantum-mechanical sense.

We use the same $1/N!$ correction to define a granular entropy that *is* extensive. We note that our modifications of Edwards' recipe for the granular entropy have resulted in a quantity that has exactly the extensivity properties assumed by Edwards'. We have thus established that the key quantity in Edwards' theory of granular media is physically meaningful. The observed robustness of the extensivity of the Edwards' entropy for our system of soft, repulsive particles may explain why experiments on soft jammed granular matter [41] find good agreement with Edwards' theory, even though in that case there is also no reason to assume that all packings are equally likely.

Chapter 7

Jamming density distribution and compactivity

In this chapter a study of the distribution of jamming densities is presented. The distribution of densities at which a system becomes jammed depends on the protocol used to generate the jammed states and on the phase space of the initial conditions. It is possible to define a compactivity that relates the distribution of packing densities to the phase-space volumes of the states of the parent system that jam at a given density.

7.1 Introduction

The density and structure of a jammed packing of N particles depends on its preparation protocol Π and on the phase-space coordinates Γ of the original phase space. Therefore a unique, universal jamming density does not exist. This protocol can be considered as a procedure that projects a part of the original phase space of the N-particle system onto a particular packing. We can separate protocols into two broad classes: *deterministic* and *stochastic*. Deterministic protocols, like our energy minimisation quenches, always map a point in phase space to a particular packing. All of the points that end up at a packing i form the volume v_i. Stochastic protocols are different because there is a certain probability $\rho_i(\Gamma)$ that a coordinate in phase space will map on to packing i. The volume v_i is nothing but

$$v_i = \int d\Gamma \, \rho_i(\Gamma). \tag{7.1}$$

For deterministic protocols, $\rho_i(\Gamma) = 1$ inside the basin v_i and $\rho_i(\Gamma) = 0$ outside. If the original distribution in phase space is uniform v_i determines the probability of mapping onto packing i

$$P_i = \frac{v_i}{\mathcal{V}}. \tag{7.2}$$

where $\mathcal{V}$ is the total phase space volume of accessible initial coordinates. If the original distribution is not uniform then the probability of mapping onto i is given by

$$P_i = \int d\Gamma \, w(\Gamma)\rho_i(\Gamma) \tag{7.3}$$

where $w(\Gamma)$ is a normalised weight function. We shall focus on deterministic packings as the protocols described in this thesis fall into this category. In the case of our protocols to generate packings, by means of energy minimisation quenches, the original phase space distribution is uniform and the packings are generated with a probability proportional to v_i. In general different jammed configurations will have different packing fractions. The sum of all volumes $v_i(\phi)$ that map onto a packing with volume fraction ϕ defines the total phase space volume of the original system that ends up in a packing with volume fraction ϕ:

$$\mathcal{V}(\phi) = \sum_i v_i(\phi). \tag{7.4}$$

ϕ is a continuous variable so $\mathcal{V}$ is really a phase space density

$$\frac{\mathcal{V}(\phi)\, d\phi}{\mathcal{V}} = \mathcal{P}(\phi)\, d\phi. \tag{7.5}$$

The entropy is defined as

$$S_J(\phi) \equiv \ln \mathcal{V}(\phi) \tag{7.6}$$

where J denotes the jammed phase. If the original phase space is not uniform a partition function can be defined as

$$Q_J = \int_\phi d\Gamma \, w(\Gamma) \tag{7.7}$$

where the integration is over the region of phase space that maps onto packings with volume fraction ϕ and $w(\Gamma)$ is a normalised weighting function.

Here we will focus on the relation between the "parent system" (the set of initial conditions), the jamming protocol Π and the density of the final jammed structure, $\phi_J = N/V_J(N,\Pi)$. It is well known that the volume V_J is self averaging [67, 17] so for finite systems, ϕ_J fluctuates but in the thermodynamic limit the relative magnitude of these fluctuations vanishes at a rate $\sim 1/\sqrt{N}$. Hence there is a one-to-one relation between ϕ_J and the initial configuration. To our knowledge this relation has not yet been predicted.

The question is whether there is a simple relation that expresses how a protocol affects the resulting jamming density? We can assume that the compactivity X plays the same role as the ratio T/P in the NPT ensemble, large pressures result in higher densities and for jammed systems low compactivity protocols produce denser jammed packings. In the NPT ensemble we have

$$\mathcal{P}(V) = \frac{Q(N,V,T)e^{-\beta PV}}{\int dV\, Q(N,V,T)e^{-\beta PV}} = \frac{Q(N,V,T)e^{-\beta PV}}{Z(N,P,T)}, \tag{7.8}$$

which defines the integral $Z(N,P,T)$. The imposed "NPT" pressure P can be related to the "canonical" virial pressure P':

$$P'(N,V,T) \equiv k_B T \frac{\partial \ln Q(N,V,T)}{\partial V} \tag{7.9}$$

and in equilibrium $P = \langle P' \rangle$:

$$\begin{aligned}
\langle P' \rangle &= \frac{\int dV\, Q(N,V,T)e^{-\beta PV} k_B T \frac{\partial \ln Q(N,V,T)}{\partial V}}{Z(N,P,T)} && (7.10)\\
&= \frac{k_B T \int dV\, e^{-\beta PV} \frac{\partial Q(N,V,T)}{\partial V}}{Z(N,P,T)} && (7.11)\\
&= \frac{-k_B T \int dV\, Q(N,V,T) \frac{\partial e^{-\beta PV}}{\partial V}}{Z(N,P,T)} && (7.12)\\
&= P \frac{\int dV\, Q(N,V,T)e^{-\beta PV}}{Z(N,P,T)} && (7.13)\\
&= P. && (7.14)
\end{aligned}$$

The equivalent identity in the discussion of the compactivity is the only way to determine the compactivity of a system. Following from Equation (7.8) the free energy per particle $f(\rho,T)$ is defined by

$$F(N,V,T) = N f(\rho,T) = -k_B T \ln Q(N,V,T). \tag{7.15}$$

the volume per particle is defined as $v \equiv V/N$ and the distribution is

$$\mathcal{P}(v) = \frac{Ne^{-N\beta(f+Pv)}}{Z(N,P,T)}. \tag{7.16}$$

We have obtained the well-known result that the most likely value of v does not depend on N and that the variance in v vanishes as $1/\sqrt{N}$. For this result to hold, we need to assume that the volume of the composite system can be written as the sum of its sub-volumes. In the case of granular media the conserved quantity is the volume $V_J(N;\Pi)$ of the jammed system which is equal to the sum of the volumes of any sub-systems. This assumption alone is not enough to get Equation (7.16) as the extensivity of the granular entropy S_J is also needed and will be taken as a hypothesis as it has been shown to hold in Chapter 6.

$$S_J(N,V_J;\Pi) = Ns_J(V_J/N;\Pi). \tag{7.17}$$

If a jammed system is in contact with a reservoir the distribution of volume per particle v_j is:

$$\mathcal{P}(v_J;X) \sim \exp[N(s_J(V_J/N;\Pi) - v_J/X)]. \tag{7.18}$$

which defines the compactivity X for a particular protocol and reservoir. The imposed compactivity is equal to the average compactivity of the system and in general

$$\left(\frac{\partial S_J(N,V_J,\Pi)}{\partial V_J}\right) \equiv \frac{1}{X(\Pi;v_J)}, \tag{7.19}$$

is not equal to the imposed value $1/X$. If a protocol can impose a compactivity, mimicking the action of a reservoir, then the distribution of jamming volumes should obey an “overlapping distributions” relation:

$$\ln\mathcal{P}(v_J;X) = \ln\mathcal{P}(v_J;X') - Nv_J\left(\frac{1}{X'} - \frac{1}{X}\right) + \text{constant}. \tag{7.20}$$

The difference with the NPT ensemble is that it is not obvious that a protocol can impose a compactivity and some studies suggest that this property does not hold [68]. The existence of an overlapping-distribution relation is a necessary but not sufficient condition for Edwards’ ergodic hypothesis to be valid. It means that $\mathcal{P}(v_J; X)$ can be expressed as the product of a distribution $Z(V_J)$ that does not depend on X and a factor $e^{-V_J/X}$. The definition of compactivity is not unique as we can always write:

$$Z(V_J)e^{-V_J/X} = Z(V_J)e^{-V_J/X_0}e^{-V_J(1/X-1/X_0)} = Z'(V_J)e^{-V_J/X'}. \quad (7.21)$$

where $1/X' = 1/X - 1/X_0$, for any X_0. Edwards noticed that there is a lower bound of the jamming density for a given protocol at which $Z(V_J)$ vanishes much like the entropy vanishes at $T = 0$. This knowledge is not enough to fix X_0 because for any finite value of X_0, $1/X = \infty$ is equivalent to $1/X' = \infty$. If Edwards' hypothesis is valid, $Z(V) = \Omega(V)$ is the number of packings with volume V, which is independent of the protocol used to find the packings and

$$\mathcal{P}(V_J; X) \sim \Omega(V_J)\exp(-V_J/X). \quad (7.22)$$

From Equation (7.20) it would follow that the logarithm of the ratio of the distribution functions that result from two different protocols should be of the form

$$\ln\left[\frac{\mathcal{P}(v_J; X)}{\mathcal{P}(v_J; X')}\right] = aV_J + b. \quad (7.23)$$

In the following section the protocols used in this work to find the jamming density are described. Once the jamming density distributions are calculated for different protocols, the relation in Equation (7.23) can be tested.

7.2 Finding the jamming density

The goal of the simulations described in this Chapter is to find the volume (or density) at which a particular initial configuration becomes jammed (or unjammed), V_J. This corresponds to finding the so called J-point, introduced in Chapter 2. If this is done for a number of protocols it may be possible to determine the relative compactivity, described above in Section 7.1.

The first protocol (Π_I) used to find the jamming density, consists in starting from a known jammed configuration (found by a potential energy minimisation) at a packing fraction $\phi_\text{max} = 0.88$ which is higher than the jamming threshold and deflating the particles by a fixed amount to reach a density ϕ_min.

If the system becomes unjammed – this is tested by minimising the potential energy and checking for particle overlaps – then the particles are inflated to the midpoint between these densities. If at this density, ϕ_i the system is jammed then $\phi_{\text{max}} = \phi_i$ and the midpoint between this density and ϕ_{min} is tested. If instead at ϕ_i the system is not jammed then $\phi_{\text{min}} = \phi_i$ and the midpoint between ϕ_i and ϕ_{max} is tested. This procedure is repeated until the interval $[\phi_{\text{min}}, \phi_{\text{max}}]$ is small enough and $\phi_J \in [\phi_{\text{min}}, \phi_{\text{max}}]$ is determined with sufficient accuracy.

The second protocol (Π_{II}) used to find the jamming density, similar to the one described in [67], consists in instantaneously inflating a Hard Sphere (HS) fluid as in Figure 6.2 and finding the minimum of the Hard Sphere core plus Weeks-Chandler-Andersen (HS+WCA) potential energy given by Equation (6.2) (shown in Figure 6.1) with $d_{ij}^{S}/d_{ij}^{\text{HD}} = 1.12$. The packing fraction of the final inflated state is $\phi = 0.88$ which is higher than the jamming threshold. The next step is to decrease the density by decreasing the shell that surrounds the HS core by multiplying $d_{ij}^{S}/d_{ij}^{\text{HD}}$ by a factor ΔD. Then the potential is minimised again and if the resulting configuration has positive energy, which means that $V < V_J$, this step is repeated until the particles are just unjammed, the energy is zero and $V \approx V_J$. The factor ΔD is reduced as the jamming density is approached. This is to ensure that no energy barriers are crossed on the way to the jamming threshold [67] and so that V_J can be found with sufficient accuracy. The Fast Inertial Relaxation Engine (FIRE) algorithm was used for all of the energy minimisations, for reasons discussed in Chapter 5 and Appendix A.

The previous paragraphs describe the methods used to find the jamming volume, V_J, associated to a given initial state (in this case a HS fluid) and protocol. If this experiment is repeated many times for different protocols, or variations of the same class of protocols, the distribution of V_J and possibly a relative compactivity can be measured. To be able to measure the compactivity X the distribution of jamming volumes must be measured for different protocols. This is achieved by varying the density of the HS cores of the particles defined by the diameter d_{ij}^{HD}.The results of these simulations are presented in the following section.

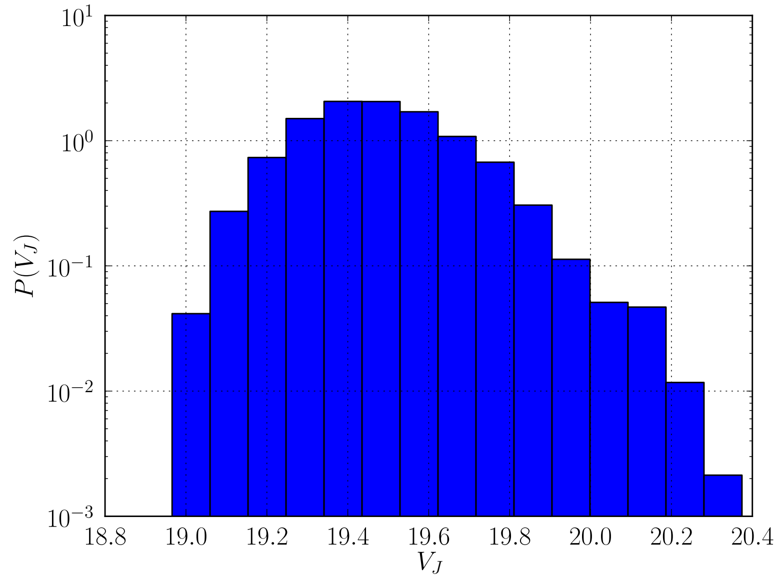


Figure 7.1: Distribution of jamming volumes V_J for a system of 16 particles with initial volume fraction $\phi_{\mathrm{HD}} = 0.71$ interacting via the HS+WCA potential with $d_{ij}^{S}/d_{ij}^{\mathrm{HD}} = 1.12$. Protocol Π_{I}.

7.3 Results and discussion

In this Section the results of the simulations described in Section 7.2 are presented and the validity of the compactivity described in Section 7.1, in particular the overlapping-distribution relation described in Equation (7.23), is tested for the protocols that were used in this work.

Figure 7.1 shows a characteristic jamming density distribution found using protocol Π_{I}. Figure 7.2 shows jamming density distributions for variants of protocol Π_{II} starting from different initial hard-disk volume fractions.

The ratio Equation (7.23) has been plotted for protocol Π_{I} in Figure 7.3 and for protocol Π_{II} in Figure 7.4.

As can be seen from these Figures, the ratio is not well represented by a straight line. Even if we take into account the ratios between many different variants of protocol Π_{II}, as shown in Figure 7.4, we cannot even imagine a straight line describing such data. It is important to stress the fact that in Figure 7.4 the ratio between all of the possible combinations of different initial states such that if the ratio is $P(V_J^{(i)})/P(V_J^{(j)})$ then $V_J^{(j)} > V_J^{(i)}$ have been accounted for so if there were a linear trend it should be visible in the Figure. Hence for the protocols tested here Equation (7.22) does not hold.

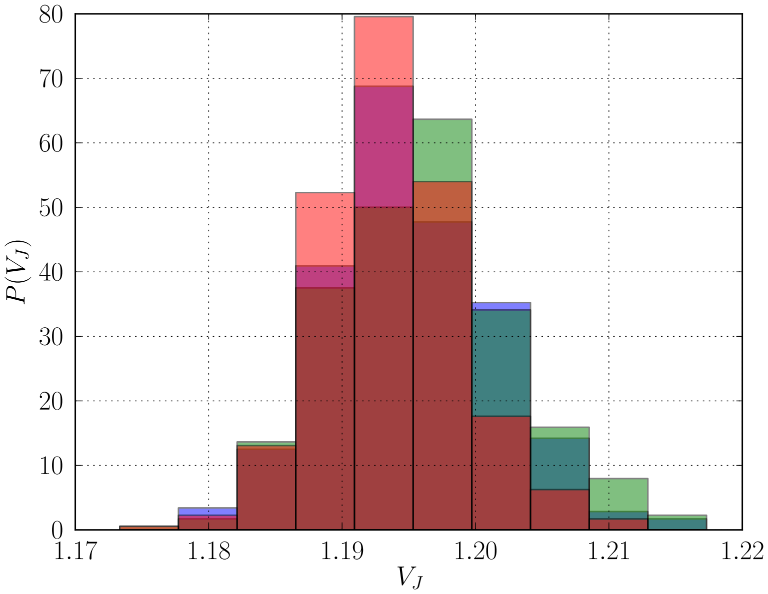


Figure 7.2: Distribution of jamming volumes V_J for a system of 64 particles with different initial volume fractions ϕ_{HD} interacting via the HS+WCA potential with $d_{ij}^{S}/d_{ij}^{\mathrm{HD}} = 1.12$. Protocol Π_{II}.

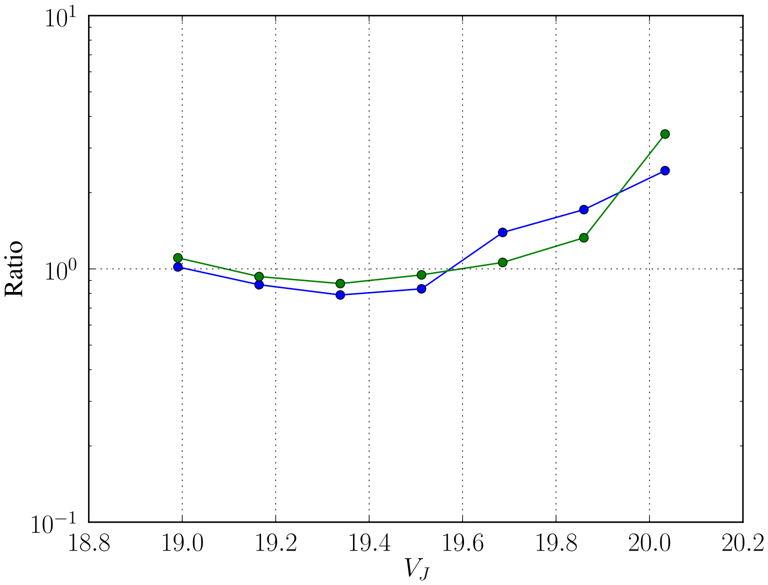


Figure 7.3: Ratio of the jamming distributions, Equation (7.23), for a system of 16 particles with initial volume fraction $\phi_{\mathrm{HD}} = 0.71$ and $\phi_{\mathrm{HD}} = 0.62$ interacting via the HS+WCA potential with $d_{ij}^{S}/d_{ij}^{\mathrm{HD}} = 1.12$. The reference state was prepared at $\phi = 0.448$. Protocol Π_{I}.

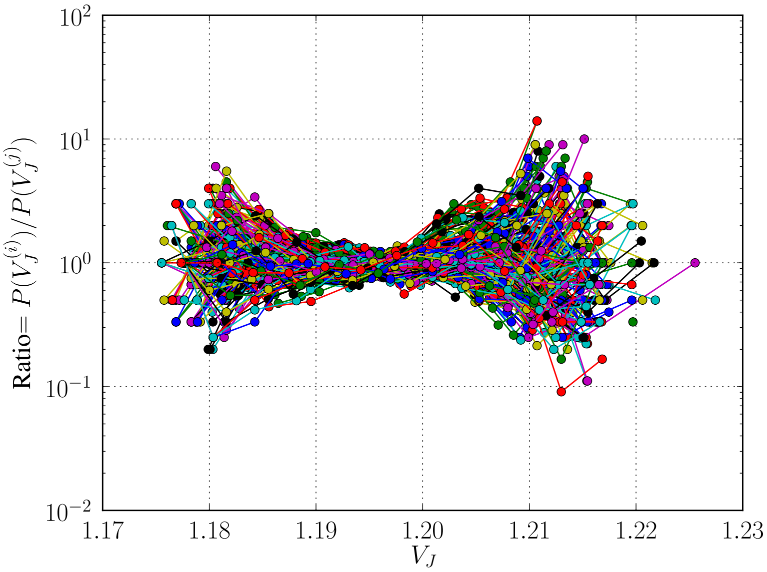


Figure 7.4: Ratio of the jamming distributions, Equation (7.23) for a system of 64 particles with different initial volume fractions ϕ_{HD} interacting via the HS+WCA potential with $d_{ij}^{S}/d_{ij}^{\mathrm{HD}} = 1.12$. All of the combinations of jamming distributions found using Protocol Π_{II} are shown.

The tapping protocols described in [68] are another example where Equation (7.22) does not hold.

An ensemble can be constructed such that the compactivity is a meaningful quantity related to the volume of the parent system. In this ensemble, the volume at which a system of N particles becomes jammed due to a certain protocol, $V_J = N/\phi_J$, acts as an order parameter that is related to the coordinates of the original system at volume V. We can introduce a field $Y \equiv 1/X$ that is conjugate to V_J and a constant NVY (equivalent to $NVTX$) ensemble can be introduced:

$$\mathcal{P}(V_J; Y, \Pi) = \frac{e^{S_J(N,V_J;\Pi)} e^{-YV_J(\mathbf{r}^N)}}{\int dV_J \, e^{S_J(N,V_J;\Pi)} e^{-YV_J(\mathbf{r}^N)}}. \tag{7.24}$$

This ensemble is unphysical because V_J can only be controlled in simulation, where particle overlap can easily be adjusted. For the most likely value of V_J the inverse compactivity of the parent system is equal to the imposed value of Y:

$$\left(\frac{\partial S_J(N, V_J; \Pi)}{\partial V_J}\right)_{\mathrm{max}} = Y. \tag{7.25}$$

We can imagine that for large N, V_J is a sharply peaked function of V and that there is effectively a one-to-one relation between V and V_J. We can then apply the chain rule and for $N \to \infty$ write:

$$\left(\frac{\partial S_J(N, V; \Pi)}{\partial V}\right)\left(\frac{\partial V}{\partial V_J}\right) = Y. \tag{7.26}$$

The first term is $-P/T$ and the second term relates V, the volume (or density) of a certain initial configuration (hard sphere fluid in the case of protocol Π_{II} defined in Section 7.2) to V_J, the volume (or density) at which this initial configuration becomes jammed. V_J is protocol dependent. We now make the conjecture that we can use the same expression to compute the compactivity as a function of V_J for a small system:

$$\left(\frac{\partial S_J(N, V_J; \Pi)}{\partial V_J}\right) = -\frac{P(N, V)}{T}\left(\frac{\partial V}{\partial V_J}\right)_{N\to\infty}. \tag{7.27}$$

If this equation can be applied to small systems it will allow us to predict the shape of the distribution of jamming densities. Both terms on the right hand side are functions of N. The first term is only weakly N-dependent. The work of O'Hern *et al.* [67] suggest that the second term has a $1/N^{d\gamma}$ correction where d is the dimensionality and γ is a scaling exponent (they found $\gamma \approx 0.71$). The relation between V and V_J is given by the protocol Π and in general is unknown, although a preliminary study of one of the protocols used in this work is presented below. We know from [90] that very slow compression leads to high jamming densities. As the compression becomes very slow, the jamming density no longer depends on the initial density and in this case Equation (7.27) is not valid. It is worth noting that slow compression is fundamentally different from the instantaneous minimisation protocols used by [67, 17]. Slow compression ensures a certain memory loss where the initial state is forgotten. Instantaneous minimisation, or fast protocols, retain some information about the parent structure.

The difference between slow and fast protocols can be seen by defining scaled particle coordinates $\mathbf{s}_i = \mathbf{r}_i/L$. The coordinates of jammed structures obtained by a fast protocol or instantaneous minimisation are close to the starting coordinates:

$$\frac{1}{N}\sum_{i=1}^{N}(\mathbf{s}_i - \mathbf{s}_i^{\mathrm{jammed}})^2 \to 0 \quad \text{as} \quad N \to \infty \tag{7.28}$$

On the other hand, for a slow compression or tapping protocol the initial and final coordinates are uncorrelated:

$$\frac{1}{N}\sum_{i=1}^{N}(\mathbf{s}_i - \mathbf{s}_i^{\text{jammed}})^2 \to \mathcal{O}(1) \quad \text{as} \quad N \to \infty. \tag{7.29}$$

We do not expect Equation (7.27) to be valid for slow protocols beause there is no relation between the initial and jammed configurations. We show some preliminary results of how they are related for a fast protocol in what follows.

We have studied the relation between the volume fraction of the parent hard disk fluid, $\phi_{\text{HD}} = \phi_U = N/V$ and the volume fraction of the jammed phase $\phi_J = N/V_J$, namely the second term on the right hand side of Equation (7.27), by applying protocol Π_{II} described in Section 7.2 to 400 different parent structures for each ϕ_U for a system of $N = 64$ particles interacting via the HS+WCA potential described in Section 6.1.1 with $d_{ij}^S/d_{ij}^{\text{HD}} = 1.12$. The resulting curves are presented in Figure 7.5 for two values of the compression rate, $\Delta D = 0.99$ and $\Delta D = 0.999$. The green curve corresponds to slower compression which is why it reaches higher density states. We expect to see an increase in ϕ_J as ϕ_U is increased as the system starts from a higher density and although for values of ϕ_U lower than about 0.7 we see decreasing values of ϕ_J the curve shows the expected tendency for higher values of ϕ_U.

7.4 Conclusions

The distribution of jamming densities produced by different protocols and their relation with the density of the initial configurations, or parent system has been studied in this Chapter. We have shown that the distribution of jamming volumes that are generated by similar jamming protocols that start from different initial densities cannot be characterised by an Edwards'-style compactivity. This has also been found for other protocols [68].

It is possible to construct an ensemble where compactivity is well defined and for fast compression there is a one-to-one relation between the volume of the parent system and that of the jammed system in the thermodynamic limit.

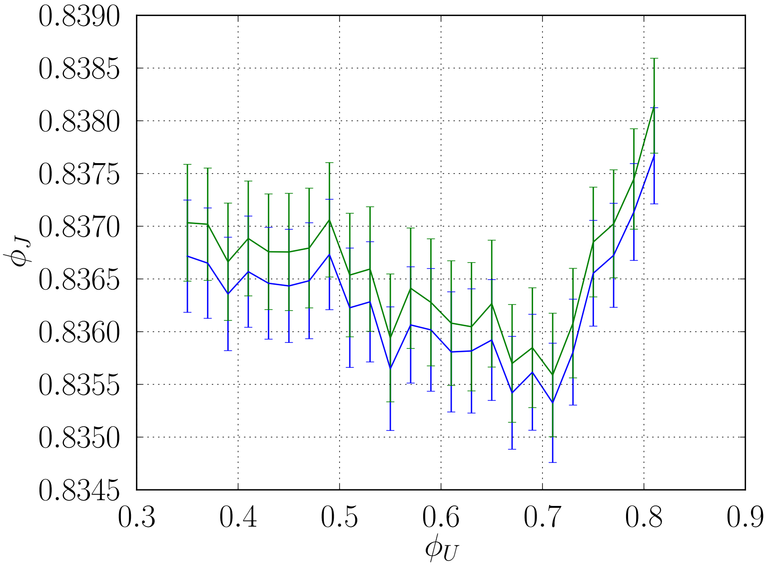


Figure 7.5: Jamming density ϕ_J as a function of the parent fluid density ϕ_U. Protocol Π_{II}. Each point is an average over 400 distinct parent configurations for a system of 64 particles interacting via the HS+WCA potential with $d_{ij}^S/d_{ij}^{\mathrm{HD}} = 1.12$. Two different compression rates are shown: $\Delta D = 0.99$ (blue) and $\Delta D = 0.999$ (green). The error bars are the standard deviation of the 400 points.

Chapter 8

General conclusions and outlook

In this thesis we have described computer simulations that allow us to evaluate the number of distinct packings of a granular material or, more generally, the so-called ‘granular entropy’ for system sizes going well beyond what was previously possible. Monte Carlo simulations were used to probe the energy landscape of jammed systems of disks interacting via a repulsive, finite-range potential. In these simulations we make use of a soft-sphere model with a hard core that approaches the hard-sphere model as the width of the soft shell is decreased. We have shown that the distribution of jamming densities produced by the protocols tested here cannot be characterised by an Edwards’-style compactivity although it is possible to construct an ensemble where compactivity is well defined.

To compute the packing entropy, we used and developed Monte Carlo techniques to determine the volumes of the basins of attraction of the potential energy minima at different system sizes.

Such Monte Carlo simulations require energy minimisation after every trial move to make sure that all accepted moves keep the system within the same basin of attraction. Hence efficient energy minimisation is a point of paramount significance in this work. A first objective was to find a suitable minimisation algorithm. We found that whereas some minimisation algorithms produce well-defined basins, others produce basins with complex boundaries or basins consisting of disconnected parts. The Fast Inertial Relaxation Engine (FIRE) algorithm was chosen because it produces compact basins at a reasonable computational cost compared to more efficient algorithms like Limited memory BFGS (L-BFGS) and Conjugate Gradient (CG) used in previous studies.

Once the minimisation algorithm was chosen a numerical approach was used to compute the number of ways in which N particles can pack into a given volume V. This technique extends the existing methods in such a way that it can be applied to much larger systems than before (over 100 particles instead of 16). Many of the caveats of previous methods were addressed. Using this novel approach, the system size dependence of the number of distinct packings of a system of poly-disperse soft disks was studied. Our simulations enable us to validate a more than 20 years old conjecture due to Edwards. We show that, even though granular particles are distinguishable, we have to deal with the Gibbs paradox. The simulations presented here provide strong evidence that the packing entropy, when properly defined, is extensive. This extensivity has often been assumed but, thus far, direct evidence was lacking. We found that, as different packings are created with unequal probabilities, the packing entropy should not be expressed as the logarithm of the number of packings but as $S = -\sum_i p_i \ln p_i$. We have shown that we can compute this quantity reliably and that, it too, is extensive. Our finding that Edwards' entropy, once properly defined, is indeed extensive is highly significant. Extensivity of the granular entropy is absolutely crucial if this quantity is to be used to define quantities such as the compactivity and angoricity [59] that play a central role in Edwards' theory of granular transport. Now that we can compute the entropy, we can start to test these theories. Of course, we also should test if other protocols to generate jammed structures also find an extensive granular entropy.

The distribution of jamming densities produced by different protocols and their relation with the density of the initial configurations, or parent system has been studied. We found that the distribution of jamming volumes that are generated by starting from different initial densities cannot be characterised by an Edwards'-style compactivity. It is possible to construct an ensemble where compactivity is well defined and for fast compression there is a one-to-one relation between the volume of the parent system and that of the jammed system in the thermodynamic limit. This allows us to hypothesise a relation between the thermodynamic pressure and compactivity.

Future directions of this work include the implementation of different, possibly more efficient, ways of sampling the basins of attraction such as the so-called *Nested Sampling* technique developed by Skilling [80]. It would also be interesting to explore even bigger systems and go to higher dimensions. To this end real experimental configurations could be used as starting, parent, systems. We should note that experimental colloidal systems are inherently poly-disperse but as we have shown, Gibbs paradox still applies. Another interesting aspect that has not been treated here is the dependence of the entropy on the volume fraction of the jammed states and how it behaves as the jamming transition is approached. For the study of jamming density distributions it would be advantageous to test if different protocols can be described by an Edwards'-style compactivity and to study the system size dependence of the jamming density.

Appendix A

Minimisation algorithms

In this Chapter various minimisation algorithms are described in detail. All of these methods are tested in Chapter 5 to see how they define basins of attraction on an energy landscape.

There are many iterative schemes for minimising functions that, starting from the initial point $\mathbf{x}_0$, are of the general form:

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \sigma_i \mathbf{w}_i. \tag{A.1}$$

Iterating over i until convergence is reached. Different methods have different ways of determining the search direction $\mathbf{w}_i$ and the step size σ_i. The search direction is usually calculated using values of the function and its gradient. The step size is determined either by scaling the search direction $\mathbf{w}_i$ by a constant value $\sigma_i = \sigma$ in Equation A.1 or determining σ_i for each iteration using a Line Search (LS) algorithm described in the next Section.

In this Chapter some of the most popular minimisation algorithms will be described along with some modifications made to the Fast Inertial Relaxation Engine (FIRE) algorithm to improve its behaviour for the systems considered in this work. In particular the modifications are made to be able to minimise a potential that involves hard-core interactions, like the one used in this work, described in Section 6.1.1.

A.1 Line search

In the previous Section a general form was given for the iterations of a minimisation method (Equation A.1). The methods described in the following Sections of this Chapter will give us the search direction $\mathbf{w}_i$ but we will need to determine how far we move along this direction, σ_i. The methods used to determine this length are called LS. Most of these methods rely on the fact that $\mathbf{w}_i$ is a *descent direction* [66] and choose a step size such that the function decreases by a certain amount.

When implementing a LS routine it is important to take into account that there is a trade-off between its efficiency at reducing the function and its computational cost. In most cases the best choice is an *inexact* LS where various different values for σ_i are considered in a certain interval which is determined by bracketing. The routine stops as soon as certain conditions are satisfied [66].

The Wolfe conditions [107, 108] are commonly used to determine whether a step size is accepted when performing inexact LS. These conditions are a set of inequalities that specify a certain decrease of the objective function f and also a condition involving its curvature. For a detailed description see [66].

A.2 Steepest descent

Steepest Descent (SD), or gradient descent, is a first-order optimisation algorithm used to find a local minimum of a function. It is also the slowest. If we want to find the minimum of a function $f(\mathbf{x})$ first a step scaling factor σ is chosen (can be varied during the minimisation in order to reach the minimum in fewer steps). Steps are taken downhill in the opposite direction of the gradient:

$$\mathbf{x}_{i+1} = \mathbf{x}_i - \sigma \nabla f(\mathbf{x}_i) \tag{A.2}$$

until the convergence criteria are satisfied.

A.3 Conjugate gradient

Conjugate Gradient (CG), created by Hestenes and Stiefel in 1952 [37], refers to direct and iterative methods that are used to solve large systems of linear equations and can be adapted to solve optimisation problems. The former are known as *linear* CG and the latter *nonlinear* CG. They rely on the properties of a conjugate set of vectors. The set of nonzero vectors $\{\mathbf{p}_0, \mathbf{p}_1, \ldots, \mathbf{p}_n\}$ is conjugate if $\mathbf{p}_i^T A \mathbf{p}_j = 0$ for all $i \neq j$ where A is a symmetric positive definite $n \times n$ matrix. These vectors are also linearly independent.

Linear CG is used to solve $A\mathbf{x} = \mathbf{b}$ which is equivalent to solving the problem

$$\min \varphi(\mathbf{x}) \equiv \frac{1}{2}\mathbf{x}^T A \mathbf{x} - \mathbf{b} \tag{A.3}$$

and the gradient of φ is the residual of the linear system

$$\nabla \varphi(\mathbf{x}) = A\mathbf{x} - \mathbf{b} \equiv r(\mathbf{x}) \tag{A.4}$$

and $r_i = A\mathbf{x}_i - \mathbf{b}$ when $\mathbf{x} = \mathbf{x}_i$. The idea is that it is possible to minimise $\varphi(\mathbf{x})$ in n steps by minimising along the directions defined by vectors in a conjugate set. The algorithm is capable of generating a new search direction $\mathbf{p}_i$ that is conjugate to the previous directions without explicitly knowing all of them. In fact only the last value is needed and $\mathbf{p}_i$ is a linear combination of the previous search direction $\mathbf{p}_{i-1}$ and the residual $\mathbf{r}_i$. The first step is taken in the SD direction as no previous direction exists. Only a few vectors are stored at the same time keeping memory usage to a minimum. The details of this algorithm are not presented, for details consult [66].

Linear CG has been shown to minimise the quadratic function φ (Equation A.3). The method can be generalised in order to minimise any general nonlinear function f. This is called nonlinear CG and one of the most popular implementations was created by Fletcher and Reeves in 1964 [29]. Given the initial coordinates $\mathbf{x}_0$, $f_0 = f(\mathbf{x}_0)$, $\nabla f_0 = \nabla f(\mathbf{x}_0)$ and the initial search direction (SD) $\mathbf{p}_0 = -\nabla f_0$, each iteration of the algorithm consists of the following steps [66]:

1. Check for convergence.
2. Compute σ_i, the length of the step, via LS (see Section A.1).
3. Set $\mathbf{x}_{i+1} = \mathbf{x}_i + \sigma_i \mathbf{p}_i$ and evaluate ∇f_{i+1}.

4. Calculate

$$\beta_{i+1} = \frac{\nabla f_{i+1}^T \nabla f_{i+1}}{\nabla f_i^T \nabla f_i} \tag{A.5}$$

$$\mathbf{p}_{i+1} = -\nabla f_{i+1} + \beta_{i+1}\mathbf{p}_i. \tag{A.6}$$

As can be seen from these equations, no matrix operations are required and only the function and its gradient are evaluated. The choice of a LS routine can have a significant impact on the resulting minimum as can be seen in Chapter 5.

A.4 BFGS and L-BFGS

Broyden, Fletcher, Goldfarb and Shanno (BFGS), a quasi-Newtonian method named after its developers [12, 28, 33, 79], uses an approximate Hessian to determine the search direction to minimise functions. Quasi-Newton methods are based on Newton's method for finding zeros of functions. Newton's method uses the first and second derivatives of the function and assumes that the function can be approximated by a quadratic near the zero. Quasi-Newtonian methods do not compute the Hessian directly, instead it is constructed from the values of the gradient. A quasi-Newtonian method looks like

$$\mathbf{x}_{i+1} = \mathbf{x}_i - \sigma_i B^{-1}[f(\mathbf{x}_i)]\nabla f(\mathbf{x}_i) \tag{A.7}$$

where B is the approximate Hessian and σ_i is the step size. Different quasi-Newtonian methods use different approximations of the Hessian ($\mathbf{w}_i = -B^{-1}[f(\mathbf{x}_i)]\nabla f(\mathbf{x}_i)$ in Equation (A.1)).

The "L" in Limited memory BFGS (L-BFGS) (developed by Nocedal and Liu [53]) comes from "limited memory" and it means that the algorithm was made more efficient by using the values of the coordinates and gradient from previous iterations to construct the Hessian, which is never explicitly calculated or stored. This is well suited for high dimensional problems where calculating the Hessian at every iteration would be very expensive. For a deeper view of this method please refer to [53, 66].

A.5 FIRE

The name FIRE comes from *fast inertial relaxation engine* as this method relies on inertia to reach the minimum on a landscape defined by an objective function. The method is based on molecular dynamics with modifications to the velocity and adaptive time steps [9]. In the following Sections the original FIRE algorithm is described along with a modified FIRE algorithm designed to improve the definition of the boundaries of the basins of attraction and to ensure that the energy always decreases during the minimisation.

A.5.1 Original FIRE

Given the initial values for the time step Δt, $\alpha = \alpha_{\text{start}}$, the velocity $\mathbf{v} = 0$ and $\mathbf{x}$ the initial positions, each iteration follows consists of the following steps:

1. Calculate $\mathbf{x}$, $\mathbf{F} = -\nabla E(\mathbf{x})$.
2. Check for convergence.
3. Calculate $P = \mathbf{F} \cdot \mathbf{v}$.
4. Set $\mathbf{v} = (1-\alpha)\mathbf{v} + \alpha \hat{\mathbf{F}} |\mathbf{v}|$.
5. If $P > 0$ and this has been the case for at least N_{min} steps then increase the time step to $\Delta t = \min(\Delta t f_{\text{inc}}, \Delta t_{\text{max}})$ and decrease α so that $\alpha = \alpha f_{\alpha}$.
6. If $P \leq 0$, the energy has increased since the previous step so the velocity is set to zero $\mathbf{v} = 0$, the time step is decreased $\Delta t = \Delta t f_{\text{dec}}$ and $\alpha = \alpha_{\text{start}}$.
7. Calculate $\mathbf{v}$ using an integrator of your choice[1].
8. Update positions.

where E is the function to be minimised. This process is repeated until convergence is reached and the minimum is found. The parameters used here

---

[1] The Euler method was used in this work. Velocity Verlet was tried but did not significantly improve the general performance of the algorithm.

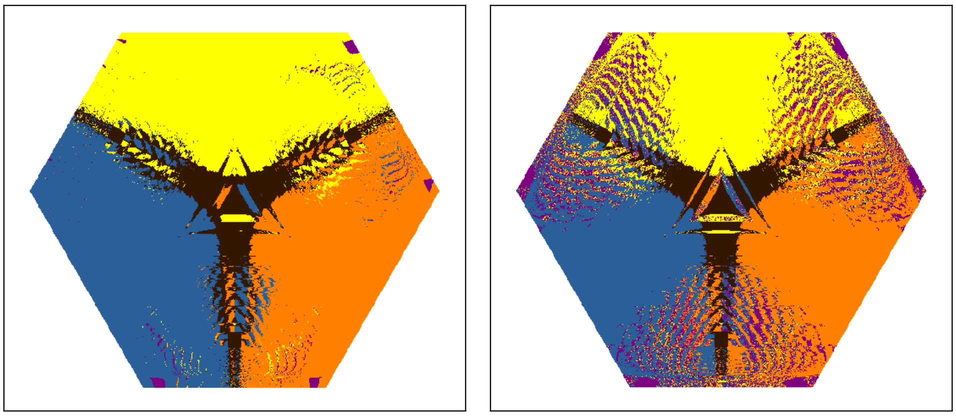
are: $\alpha_{\text{start}} = 0.1$, $N_{\text{min}} = 5$, $f_{\text{inc}} = 1.1$, $f_{\text{dec}} = 0.5$, $f_\alpha = 0.99$. These parameters have been tested and produce reliable and fast results [9]. The only remaining parameter to fix is Δt_{max}. We show in Chapter 5 and [4] how this choice can greatly affect the topology of the basins of attraction found using this algorithm. In the next Section a modification to FIRE that prevents uphill motion on the energy landscape is introduced.

A.5.2 Modified FIRE

Looking closely at the algorithm described in the previous Section, it is clear from step 5 that one uphill step is taken ($P > 0$) before the parameters are reset. This can cause problems when combined with a large maximum step size because large energies produced by particles being too close together can cause the system to jump into another basin of attraction. This effect can be seen in Chapter 5 and Figure 5.4.

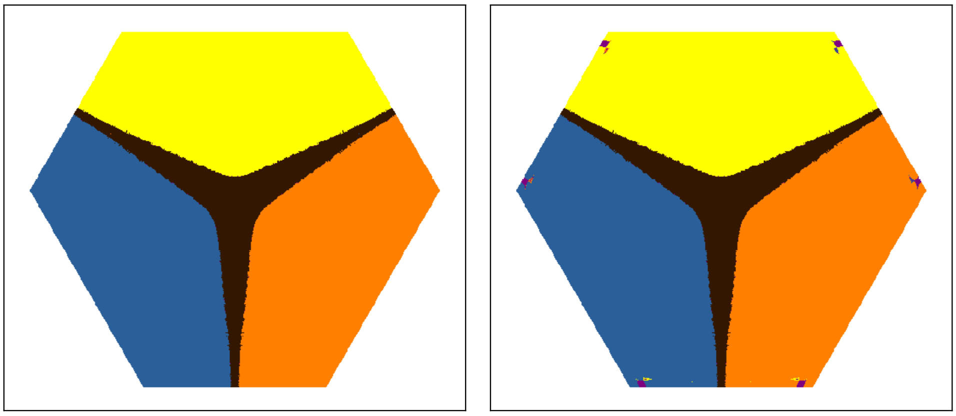
To address this issue a modification to FIRE has been implemented that assures that all steps taken always go downhill on the energy landscape. It is essentially the same algorithm except that the step is not taken if an uphill step is proposed. The steps taken at each iteration are the same as before except that steps 7. and 8. are only performed if $P > 0$ in step 5..

Figure A.1 shows a comparison of the basins of attraction described by the original FIRE algorithm (Section A.5.1) and the modified version described in this Section using the framework described in Chapter 5. The top row corresponds to the original FIRE algorithm and the bottom row to the modified FIRE algorithm for different maximum step sizes $\Delta = 0.5$ and $\Delta = 1.0$. It is clear that the modified FIRE algorithm proposed here is more stable than the original algorithm in the sense that the basins of attraction that it produces are consistent with those produced by SD (Figure 5.2). It is also worth mentioning that when using the Hard Sphere core plus Weeks-Chandler-Andersen (HS+WCA) potential (described in Section 6.1.1) where the particles have a hard core that cannot overlap, the original FIRE cannot be used due to the fact that an uphill move could produce overlapping cores. Overlaps of the hard cores are tested at the same time as the $P > 0$ condition is tested and both conditions must return false for the step to be taken.



FIRE, $\Delta = 0.5$

FIRE, $\Delta = 1.0$



modified FIRE, $\Delta = 0.5$

modified FIRE, $\Delta = 1.0$

Figure A.1: Comparison between defining basins of attraction with FIRE (top row, described in Section A.5.1) and with modified FIRE (bottom row, Section A.5.2) using the methods described in Chapter 5. This Figure shows results for two different values of the maximum step size Δ.

Appendix B

Centre of mass constraint in free energy calculation.

For a given spring constant k, we numerically measure the probability p_k to be in a basin:

$$p_k = C_k \int_{-\infty}^{\infty} \exp\left[\frac{-ku^2}{2}\right] \chi_i \, d\mathbf{u} \tag{B.1}$$

where

$$\chi_i = \begin{cases} 1 & \text{inside basin } i \\ 0 & \text{outside basin } i \end{cases}$$

restricts the integral to a certain basin i. Finally C_k is a normalisation constant associated with the Gaussian "umbrella" and reads

$$C_k = \left(\frac{k}{2\pi}\right)^{dN/2}.$$

For our volume calculation, presented in Chapter 4, we are interested in the biased volume of the reference state:

$$v_k = \int_{-\infty}^{\infty} \exp\left[\frac{-ku^2}{2}\right] \chi_i \, d\mathbf{u} \tag{B.2}$$

that we can easily estimate thanks to the probability p_k that has already been measured:

$$-\ln v_k = -\ln p_k - \frac{dN}{2} \ln \frac{2\pi}{k} \tag{B.3}$$

This is the quantity that we compute for the reference state. The only problem here is that, for efficiency reasons, described in Chapter 3, we perform simulations at fixed centre of mass and therefore Equation (B.3) does not quite hold and a correction must be applied.

Since we work at fixed centre of mass, what we measure is the conditional probability for the configuration vector sampled with a Gaussian probability, to be in a certain basin i knowing that the centre of mass is fixed at the centre of the box (assumed to be the origin of the coordinate system). This reads

$$p_k(\text{basin } i | \vec{R}_{\text{CM}} = \vec{0}) = \frac{p_k(\text{basin } i \cap \vec{R}_{\text{CM}} = \vec{0})}{p_k(\vec{R}_{\text{CM}} = \vec{0})} \tag{B.4}$$

Now by definition we can write

$$\begin{aligned} p_k(\text{basin } i \cap \vec{R}_{\text{CM}} = \vec{0}) &= C_k \int_{-\infty}^{\infty} \exp\left[\frac{-ku^2}{2}\right] \chi_i \delta(\vec{R}_{\text{CM}})\, d\mathbf{u} \\ &= C_k v_k^{\text{CM}}. \end{aligned} \tag{B.5}$$

Hence the reference biased volume that we can measure is v_k^{CM}. The probability for the centre of mass to be at the origin can easily be estimated if we remember that

$$\vec{R}_{\text{CM}} = \frac{1}{N} \sum_{i=1}^{N} \vec{x}_i$$

and because the Gaussian probability is unrelated to the basin i –and the particle interactions– the random variables $\vec{x}_i$ are all independent. The distribution of $\vec{R}_{\text{CM}}$, which can easily be shown to be Gaussian, is given by:

$$p_k(\vec{R}_{\text{CM}}) = \left(\frac{Nk}{2\pi}\right)^{d/2} \exp\left[-\frac{1}{2} Nk R_{\text{CM}}^2\right]. \tag{B.6}$$

Putting all of these results together gives us:

$$\begin{aligned} -\ln v_k^{\text{CM}} &= -\ln p_k(\text{basin } i | \vec{R}_{\text{CM}} = \vec{0}) - \frac{dN}{2} \ln \frac{2\pi}{k} + \frac{d}{2} \ln \frac{2\pi}{kN} && \text{(B.7)} \\ &= -\ln v_k + \frac{d}{2} \ln \frac{2\pi}{kN} && \text{(B.8)} \end{aligned}$$

where the last term on the right hand side is the correction needed when considering a system at fixed centre of mass.

The idea of thermodynamic integration, described in Chapter 3 and further explored in Chapter 4, consists in writing:

$$\ln v^{\mathrm{CM}} - \ln v_k^{\mathrm{CM}} = \int_0^k dx \, \frac{\partial \ln v_x^{\mathrm{CM}}}{\partial x} \equiv \frac{1}{2} \int_0^k dx \, \langle u^2 \rangle_{x, \vec{R}_{\mathrm{CM}} = \vec{0}} \tag{B.9}$$

Once we have calculated the reference volume at fixed centre of mass, we can then calculate the volume of the basin without any Gaussian bias for a fixed centre of mass at the origin. The actual volume of basin i can then be obtained by integrating over all possible positions of the centre of mass:

$$v = \int_{\text{basin } i} d\vec{R} \, v(\vec{R}_{\mathrm{CM}} = \vec{R}). \tag{B.10}$$

We then use the fact that in our model, the box does not contain any hard walls but rather has Periodic Boundary Conditions (PBC) so we must enforce invariance by translation (or equivalently conservation of momentum). This implies that $v(\vec{R}_{\mathrm{CM}} = \vec{R}) = v(\vec{R}_{\mathrm{CM}} = \vec{0})$ and the volume of the basin is $v = v(\vec{R}_{\mathrm{CM}} = \vec{0}) V_{\mathrm{box}}$. Since we work in a unit box, it gives finally:

$$-\ln v = -\ln p_k(\vec{R}_{\mathrm{CM}} = \vec{0}) - \frac{dN}{2} \ln \frac{2\pi}{k} + \frac{d}{2} \ln \frac{2\pi}{Nk} - \frac{1}{2} \int_0^k dx \, \langle u^2 \rangle_{x, \vec{R}_{\mathrm{CM}} = \vec{0}} \tag{B.11}$$

Bibliography

[1] Abramowitz, M., and Stegun, I. *Handbook of Mathematical Functions*, first ed. Dover Publications, 1965.

[2] Adams, D. J., and Matheson, A. J. Computation of Dense Random Packings of Hard Spheres. *J. Chem. Phys. 56*, 5 (1972), 1989.

[3] Alexander, S. Amorphous solids: their structure, lattice dynamics and elasticity. *Phys. Rep. 296*, 24 (1998), 65.

[4] Asenjo, D., Stevenson, J. D., Wales, D. J., and Frenkel, D. Visualizing basins of attraction with different minimizers. *J. Phys. Chem. B* (2012).

[5] Ashwin, S. S., Blawzdziewicz, J., O'Hern, C. S., and Shattuck, M. D. Calculations of the structure of basin volumes for mechanically stable packings. *Phys. Rev. E 85* (Jun 2012), 061307.

[6] Axilrod, B. M., and Teller, E. Interaction of the van der waals type between three atoms. *J. Chem. Phys. 11*, 6 (1943), 299–300.

[7] Barrat, A., Kurchan, J., Loreto, V., and Sellitto, M. Edwards' measures for powders and glasses. *Phys. Rev. Lett. 85*, 24 (2000), 5034.

[8] Bernal, J. D., and Mason, J. Packing of Spheres: Co-ordination of Randomly Packed Spheres. *Nature 188*, 4754 (Dec. 1960), 910–911.

[9] Bitzek, E., Koskinen, P., Gähler, F., Moseler, M., and Gumbsch, P. Structural relaxation made simple. *Phys. Rev. Lett. 97* (Oct 2006), 170201.

[10] Blumenfeld, R., and Edwards, S. F. On granular stress statistics: compactivity, angoricity, and some open issues. *J. Phys. Chem. B 113*, 12 (2009), 3981–3987.

[11] Brown, E., Rodenberg, N., Amend, J., Mozeika, A., Steltz, E., Zakin, M. R., Lipson, H., and Jaeger, H. M. Universal robotic gripper based on the jamming of granular material. *Proc. Natl. Acad. Sci. USA 107*, 44 (2010), 18809–18814.

[12] Broyden, C. G. The convergence of a class of double-rank minimization algorithms 1. general considerations. *IMA J. Appl. Math. 6*, 1 (1970), 76–90.

[13] Brujić, J., Song, C., Wang, P., Briscoe, C., Marty, G., and Makse, H. A. Measuring the coordination number and entropy of a 3d jammed emulsion packing by confocal microscopy. *Phy. Rev. Lett. 98*, 24 (2007), 248001.

[14] Byrd, R., Lu, P., Nocedal, J., and Zhu, C. A limited memory algorithm for bound constrained optimization. *SIAM J. Sci. Comput. 16*, 5 (1995), 1190–1208.

[15] Camia, F. A note on Edwards' hypothesis for zero-temperature Ising dynamics. *Eur. Phys. J. B 47* (Sept. 2005), 239–243.

[16] Casselman, B. `http://www.math.sunysb.edu/~tony/whatsnew/column/pennies-1200/cass1.html`, 2000.

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

[18] de Souza, V. K., and Wales, D. J. Energy landscapes for diffusion: Analysis of cage-breaking processes. *J. Chem. Phys. 129* (2008), 164507.

[19] Doliwa, B., and Heuer, A. Energy barriers and activated dynamics in a supercooled lennard-jones liquid. *Phys. Rev. E 67* (2003), 031506.

[20] DONEV, A., TORQUATO, S., STILLINGER, F. H., AND CONNELLY, R. A linear programming algorithm to test for jamming in hard-sphere packings. *Journal of Computational Physics 197*, 1 (2004), 139–166.

[21] DOYE, J. P. K., AND WALES, D. J. Systematic investigation of the structures and rearrangements of 6- atom clusters bound by a model anisotropic potential. *J. Chem. Soc., Faraday Trans. 88* (1992), 3295–3304.

[22] DOYE, J. P. K., AND WALES, D. J. An order-parameter approach to coexistence in atomic clusters. *J. Chem. Phys. 102* (1995), 9673–9688.

[23] EARL, D. J., AND DEEM, M. W. Parallel tempering: Theory, applications, and new perspectives. *Phys. Chem. Chem. Phys. 7*, 23 (2005), 3910.

[24] EDWARDS, S., AND OAKESHOTT, R. Theory of powders. *Physica A 157*, 3 (June 1989), 1080–1090.

[25] EDWARDS, S. F. The flow of powders and of liquids of high viscosity. *J. Phys.: Condens. Matter 2*, S (Dec. 1990), SA63–SA68.

[26] EFRON, B., AND TIBSHIRANI, R. *An Introduction to the Bootstrap*. Monographs on statistics and applied probabilities. Chapman & Hall/CRC, 1993.

[27] FEJES TÓTH, L. *Lagerungen in der Ebene, auf der Kugel und im Raum*. Die Grundlehren der Mathematischen Wissenschaften in Einzeldarstellungen mit besonderer Berücksichtigung der Anwendungsgebiete, Band LXV. Springer-Verlag, Berlin, 1953.

[28] FLETCHER, R. A new approach to variable metric algorithms. *Comput. J. 13*, 3 (1970), 317–322.

[29] FLETCHER, R., AND REEVES, C. M. Function minimization by conjugate gradients. *Comput. J. 7*, 2 (1964), 149–154.

[30] FRANZ, S., AND VIRASORO, M. A. Quasi-equilibrium interpretation of ageing dynamics. *J. Phys. A 33*, 5 (2000), 891.

[31] FRENKEL, D., AND LADD, A. J. C. New monte carlo method to compute the free energy of arbitrary solids. application to the fcc and hcp phases of hard spheres. *J. Chem. Phys. 81*, 7 (Oct. 1984), 3188–3193.

[32] FRENKEL, D., AND SMIT, B. *Understanding Molecular Simulation: From Algorithms to Applications*, second ed. Academic Press, Nov. 2001.

[33] GOLDFARB, D. A family of variable-metric methods derived by variational means. *Math. Comp. 24*, 109 (1970), pp. 23–26.

[34] GREBOGI, C., OTT, E., AND YORKE, J. A. Chaos, strange attractors, and fractal basin boundaries in nonlinear dynamics. *Science 238*, 4827 (1987), 632–638.

[35] HALES, T. C. A proof of the Kepler conjecture. *Annals of Mathematics 162*, 3 (2005), 1063–1183.

[36] HENKES, S., VAN HECKE, M., AND VAN SAARLOOS, W. Critical jamming of frictional grains in the generalized isostaticity picture. *Europhys. Lett. 90*, 1 (2010), 14003.

[37] HESTENES, M. R., AND STIEFEL, E. Methods of Conjugate Gradients for Solving Linear Systems. *J. Res. Nat. Bur. Stand. 49*, 6 (Dec. 1952), 409–436.

[38] JAEGER, H. M., NAGEL, S. R., AND BEHRINGER, R. P. The Physics of Granular Materials. *Phys. Today 49*, 4 (1996), 32.

[39] JAYNES, E. The gibbs' paradox. In *Maximum Entropy and Bayesian Methods* (1992), C. Smith, G. Erickson, and P. Neudorfer, Eds., Kluwer Academic, pp. 1–22.

[40] JONES, E., OLIPHANT, T., PETERSON, P., ET AL. SciPy: Open source scientific tools for Python, 2001–.

[41] JORJADZE, I., PONTANI, L.-L., NEWHALL, K. A., AND BRUJIĆ. Attractive emulsion droplets probe the phase diagram of jammed granular matter. *Proc. Natl. Acad. Sci. USA 108* (2011), 4286–4291.

[42] KAMIEN, R., AND LIU, A. Why is Random Close Packing Reproducible? *Phys. Rev. Lett. 99*, 15 (Oct. 2007), 155501.

[43] KIRKPATRICK, T. R., THIRUMALAI, D., AND WOLYNES, P. G. Scaling concepts for the dynamics of viscous liquids near an ide al glassy state. *Phys. Rev. A 40* (1989), 1045.

[44] KIRKWOOD, J. G. Statistical mechanics of fluid mixtures. *J. Chem. Phys. 3*, 5 (1935), 300–313.

[45] KNIGHT, J., FANDRICH, C., LAU, C., JAEGER, H., AND NAGEL, S. Density relaxation in a vibrated granular material. *Phys. Rev. E 51*, 5 (May 1995), 3957–3963.

[46] KOB, W., AND ANDERSEN, H. C. Kinetic lattice-gas model of cage effects in high-density liquids and a test of mode-coupling theory of the ideal-glass transition. *Phys. Rev. E 48* (Dec 1993), 4364–4377.

[47] KRZAKALA, F., AND KURCHAN, J. Landscape analysis of constraint satisfaction problems. *Phys. Rev. E 76* (Aug 2007), 021122.

[48] LECHENAULT, F., DA CRUZ, F., DAUCHOT, O., AND BERTIN, E. Free volume distributions and compactivity measurement in a bidimensional granular packing. *J. Stat. Mech. Theor. Exp. 2006*, 07 (2006), P07009.

[49] LENNARD-JONES, J. E. On the determination of molecular fields. ii. from the equation of state of a gas. *Proc. R. Soc. A 106*, 738 (1924), 463–477.

[50] LEVENBERG, K. A method for the solution of certain problems in least squares. *Q. Appl. Math. 2* (1944), 164–168.

[51] LI, Z., AND SCHERAGA, H. A. Monte carlo-minimization approach to the multiple-minima problem in protein folding. *Proc. Natl. Acad. Sci. USA 84* (1987), 6611.

[52] LIU, A. J., AND NAGEL, S. R. Nonlinear dynamics: Jamming is not just cool any more. *Nature 396*, 6706 (Nov. 1998), 21–22.

[53] Liu, D. C., and Nocedal, J. On the limited memory bfgs method for large scale optimization. *Math. Prog. 45* (1989), 503–528. 10.1007/BF01589116.

[54] Lòpez de Haro, M., Santos, A., and Yuste, S. A student oriented derivation of a reliable equation of state for a hard-disk fluid. *Eur. J. Phys. 19* (1998), 281–286.

[55] Lubachevsky, B. D., Stillinger, F. H., and Pinson, E. N. Disks vs. spheres: Contrasting properties of random packings. *J. Stat. Phys. 64*, 3-4 (Aug. 1991), 501–524.

[56] Makse, H. A., and Kurchan, J. Testing the thermodynamic approach to granular matter with a numerical model of a decisive experiment. *Nature 415*, 6872 (2002), 614–617.

[57] Mari, R., Krzakala, F., and Kurchan, J. Jamming versus glass transitions. *Phys. Rev. Lett. 103* (Jul 2009), 025701.

[58] Marquardt, D. W. An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. Ind. Appl. Math. 11*, 2 (1963), 431–441.

[59] Mehta, A. *Granular physics*. Cambridge University Press, 2007.

[60] Mehta, D., Hauenstein, J. D., and Wales, D. J. Communication: Certifying the potential energy landscape. *J. Chem. Phys. 138*, 17 (2013), 171101.

[61] 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*, 6 (1953), 1087.

[62] Mezey, P. G. Catchment region partitioning of energy hypersurfaces, i. *Theo. Chim. Acta 58* (1981), 309.

[63] Morales, J. L., and Nocedal, J. Remark on "Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound constrained optimization". *ACM Trans. Math. Software 38*, 1 (Nov. 2011), 7:1–7:4.

[64] More, J. J., and Thuente, D. J. Line search algorithms with guaranteed sufficient decrease. *Acm T. Math. Software 20*, 3 (1994), 286–307.

[65] MURRELL, J. N., AND LAIDLER, K. J. Symmetries of activated complexes. *Trans. Faraday. Soc. 64* (1968), 371–377.

[66] NOCEDAL, J., AND WRIGHT, S. *Numerical optimization.* Springer verlag, 1999.

[67] O'HERN, C., SILBERT, L., LIU, A., AND NAGEL, S. Jamming at zero temperature and zero applied stress: The epitome of disorder. *Phys. Rev. E 68*, 1 (July 2003), 011306.

[68] PAILLUSSON, F., AND FRENKEL, D. Probing ergodicity in granular matter. *Phys. Rev. Lett. 109* (Nov 2012), 208001.

[69] PARISI, G., AND ZAMPONI, F. Mean-field theory of hard sphere glasses and jamming. *Rev. Mod. Phys. 82*, 1 (2010), 789.

[70] PAVLOVITCH, A., JULLIEN, R., AND MEAKIN, P. Geometrical properties of a random packing of hard spheres. *Physica A 176*, 2 (Aug. 1991), 206–219.

[71] PECHUKAS, P. On simple saddle points of a potential surface, the conservation of nuclear symmetry along paths of steepest descent, and the symmetry of transition states. *J. Chem. Phys. 64* (1976), 1516.

[72] PHILIPPE, P., AND BIDEAU, D. Compaction dynamics of a granular medium under vertical tapping. *Europhys. Lett. 60*, 5 (Dec. 2002), 677–683.

[73] POLSON, J. M., AND FRENKEL, D. Calculation of solid-fluid phase equilibria for systems of chain molecules. *J. Chem. Phys. 109*, 1 (July 1998), 318–328.

[74] PRESS, W., TEUKOLSKY, S., VETTERLING, W., AND FLANNERY, B. *Numerical Recipes in C: The Art of Scientific Computing*, second ed. Cambridge University Press, 1992.

[75] PUCKETT, J. G., AND DANIELS, K. E. Equilibrating temperaturelike variables in jammed granular subsystems. *Phys. Rev. Lett. 110* (Jan 2013), 058001.

[76] RICHARD, P., NICODEMI, M., DELANNAY, R., RIBIÈRE, P., AND BIDEAU, D. Slow relaxation and compaction of granular systems. *Nat. Mater. 4*, 2 (Feb. 2005), 121–8.

[77] SANTOS, A., YUSTE, S., AND LÒPEZ DE HARO, M. Equation of state of a multicomponent d-dimensional hard-sphere fluid. *Mol. Phys. 96* (1999), 1–5.

[78] SCOTT, G. D. Packing of Spheres: Packing of Equal Spheres. *Nature 188*, 4754 (Dec. 1960), 908–909.

[79] SHANNO, D. F. Conditioning of quasi-newton methods for function minimization. *Math. Comp. 24*, 111 (1970), pp. 647–656.

[80] SKILLING, J. Nested Sampling for General Bayesian computation. *Bayesian Analysis 1*, 4 (2006), 833–859.

[81] SONG, C., WANG, P., AND MAKSE, H. A. Experimental measurement of an effective temperature for jammed granular materials. *Proc. Natl. Acad. Sci. USA 102*, 7 (2005), 2299–2304.

[82] SONG, C., WANG, P., AND MAKSE, H. A. A phase diagram for jammed matter. *Nature 453*, 7195 (2008), 629–632.

[83] STILLINGER, F. Exponential multiplicity of inherent structures. *Phys. Rev. E 59* (1999), 1063.

[84] STILLINGER, F., AND WEBER, T. Dynamics of structural transitions in liquids. *Phys. Rev. A 28*, 4 (Oct. 1983), 2408–2416.

[85] STILLINGER, F. H., ET AL. A topographic view of supercooled liquids and glass formation. *Science 267*, 5206 (1995), 1935–1939.

[86] STRODEL, B., AND WALES, D. J. Free energy surfaces from an extended harmonic superposition approach and kinetics for alanine dipeptide. *Chem. Phys. Lett. 466* (2008), 105–115.

[87] SWENDSEN, R. H., AND WANG, J.-S. Replica Monte Carlo Simulation of Spin-Glasses. *Phys. Rev. Lett. 57*, 21 (Nov. 1986), 2607–2609.

[88] SZPIRO, G. Mathematics: does the proof stack up? *Nature 424*, 6944 (July 2003), 12–3.

[89] TORQUATO, S., TRUSKETT, T., AND DEBENEDETTI, P. Is Random Close Packing of Spheres Well Defined? *Phys. Rev. Lett. 84*, 10 (Mar. 2000), 2064–2067.

[90] TORQUATO, S., TRUSKETT, T. M., AND DEBENEDETTI, P. G. Is random close packing of spheres well defined? *Phys. Rev. Lett. 84* (Mar 2000), 2064–2067.

[91] UPPENBRINK, J., AND WALES, D. J. When do gradient optimizations converge to saddle-points? *Chem. Phys. Lett. 190* (1992), 447–452.

[92] VAN HECKE, M. Jamming of soft particles: geometry, mechanics, scaling and isostaticity. *J. Phys.: Condens. Matter 22*, 3 (2010), 33101.

[93] VAN KAMPEN, N. G. The gibbs paradox. In *Essays in Theoretical Physics* (1984), vol. 1, Pergamon Press, Oxford, UK, p. 303.

[94] WALES, D. J. Optim: A program for optimising geometries and calculating pathways.

[95] WALES, D. J. Structural and topological consequences of anisotropic interactions in clusters. *J. Chem. Soc., Faraday Trans. 86* (1990), 3505–3517.

[96] WALES, D. J. Basins of attraction for stationary points on a potential-energy surface. *J. Chem. Soc., Faraday Trans. 88*, 5 (1992), 653.

[97] WALES, D. J. Coexistence in small inert-gas clusters. *Mol. Phys. 78* (1993), 151–171.

[98] WALES, D. J. Locating stationary points for clusters in cartesian coordinates. *J. Chem. Soc., Faraday Trans. 89*, 9 (1993), 1305.

[99] WALES, D. J. *Energy Landscapes: With Applications to Clusters, Biomolecules and Glasses*. Cambridge University Press, 2003.

[100] WALES, D. J., AND BOGDAN, T. V. Potential energy and free energy landscapes. *J. Phys. Chem. B 110* (2006), 20765–20776.

[101] WALES, D. J., AND DOYE, J. P. K. Global optimization by basin-hopping and the lowest energy structures of lennard-jones clusters containing up to 110 atoms. *J. Phys. Chem. A 101* (1997), 5111–5116.

[102] WALES, D. J., ET AL. Gmin: A program for finding global minima and calculating thermodynamic properties from basin-sampling. `http://www-wales.ch.cam.ac.uk/GMIN/`.

[103] WEAIRE, D., AND ASTE, T. *The Pursuit of Perfect Packing*, second ed. Taylor & Francis, 2008.

[104] WEEKS, J. D., CHANDLER, D., AND ANDERSEN, H. C. Role of repulsive forces in determining the equilibrium structure of simple liquids. *J. Chem. Phys. 54*, 12 (1971), 5237–5247.

[105] WIKIPEDIA. Generalized normal distribution — wikipedia, the free encyclopedia, 2013. [Online; accessed 13-June-2013].

[106] WIKIPEDIA. Sphere packing — wikipedia, the free encyclopedia, 2013. [Online; accessed 13-June-2013].

[107] WOLFE, P. Convergence conditions for ascent methods. *SIAM Rev. 11*, 2 (1969), 226–235.

[108] WOLFE, P. Convergence conditions for ascent methods. ii: Some corrections. *SIAM Rev. 13*, 2 (1971), 185–188.

[109] XU, N., BLAWZDZIEWICZ, J., AND O'HERN, C. Random close packing revisited: Ways to pack frictionless disks. *Phys. Rev. E 71*, 6 (June 2005), 61306.

[110] XU, N., FRENKEL, D., AND LIU, A. J. Direct determination of the size of basins of attraction of jammed solids. *Phys. Rev. Lett. 106* (Jun 2011), 245502.

[111] ZAMPONI, F. Some recent theoretical results on amorphous packings of hard spheres. *Philos. Mag. 87*, 3-5 (2007), 485–495.

[112] ZHANG, H., AND MAKSE, H. Jamming transition in emulsions and granular materials. *Phy. Rev. E 72*, 1 (2005), 011301.

[113] ZHU, C., BYRD, R. H., LU, P., AND NOCEDAL, J. Algorithm 778. L-BFGS-B: Fortran subroutines for Large-Scale bound constrained optimization. *ACM Trans. Math. Software 23*, 4 (Dec. 1997), 550–560.