Structural Phase Behaviour
Via Monte Carlo Techniques

Andrew N Jackson

Doctor of Philosophy

The University of Edinburgh

August 2001

Abstract

There are few reliable computational techniques applicable to the problem of structural phase behaviour. This is starkly emphasised by the fact that there are still a number of unanswered questions concerning the solid state of some of the simplest models of matter. To determine the phase behaviour of a given system we invoke the machinery of statistical physics, which identifies the equilibrium phase as that which minimises the free-energy. This type of problem can only be dealt with fully via numerical simulation, as any less direct approach will involve making some uncontrolled approximation. In particular, a numerical simulation can be used to evaluate the free-energy *difference* between two phases *if* the simulation is free to visit them both. However, it has proven very difficult to find an algorithm which is capable of efficiently exploring two different phases, particularly when one or both of them is a crystalline solid.

This thesis builds on previous work [1] (Physical Review Letters **79** p.3002), exploring a new Monte Carlo approach to this class of problem. This new simulation technique uses a global coordinate transformation to switch between two different crystalline structures. Generally, this 'lattice switch' is found to be extremely unlikely to succeed in a normal Monte Carlo simulation. To overcome this, extended-sampling techniques are used to encourage the simulation to visit 'gateway' microstates where the switch will be successful. After compensating for this bias in the sampling, the free-energy difference between the two structures can be evaluated directly from their relative probabilities. As concrete examples on which to base the research, the lattice-switch Monte Carlo method is used to determine the free-energy difference between the face-centred cubic (*fcc*) and hexagonal close-packed (*hcp*) phases of two generic model systems — the hard-sphere and Lennard-Jones potentials.

The structural phase behaviour of the hard-sphere solid is determined at densities near melting and in the close-packed limit. The factors controlling the efficiency of the lattice-switch approach are explored, as is the character of the 'gateway' microstates. The face-centred cubic structure is identified as the thermodynamically stable phase, and the free-energy difference between the two structures is determined with high precision. These results are shown to be in complete agreement with the results of other authors in the field (published during the course of this work), some of whom adopted the lattice-switch method for their calculations. Also, the results are favourably compared against the experimentally

observed structural phase behaviour of sterically-stabilised colloidal dispersions, which are believed to behave like systems of hard spheres.

The logical extension of the hard sphere work is to generalise the lattice-switch technique to deal with 'softer' systems, such as the Lennard-Jones solid. The results in the literature for the structural phase behaviour of this relatively simple system are found to be completely inconsistent. A number of different approaches to this problem are explored, leading to the conclusion that these inconsistencies arise from the way in which the potential is truncated. Using results for the ground-state energies and from the harmonic approximation, we develop a new truncation scheme which allows this system to be simulated accurately and efficiently. Lattice-switch Monte Carlo is then used to determine the *fcc–hcp* phase boundary of the Lennard-Jones solid in its entirety. These results are compared against the experimental results for the Lennard-Jones potential's closest physical analogue, the rare-gas solids. While some of the published rare-gas observations are in approximate agreement with the lattice-switch results, these findings contradict the widely held belief that *fcc* is the equilibrium structure of the heavier rare-gas solids for all pressures and temperatures. The possible reasons for this disagreement are discussed. Finally, we examine the pros and cons of the lattice-switch technique, and explore ways in which it can be extended to cover an even wider range of structures and interactions.

An electronic (PDF) version of this thesis can be found at:
http://www.ph.ed.ac.uk/cmatter/links/anj-thesis/.

It sounded to him like a riddle, and he was never much good at riddles, being a Bear of Very Little Brain. So he sang *Cottleston Pie* instead:

Cottleston, Cottleston, Cottleston Pie,
A fly can't bird, but a bird can fly.
Ask me a riddle, and I reply:
"*Cottleston, Cottleston, Cottleston Pie.*"

Cottleston, Cottleston, Cottleston Pie,
A fish can't whistle, and neither can I.
Ask me a riddle, and I reply:
"*Cottleston, Cottleston, Cottleston Pie.*"

Cottleston, Cottleston, Cottleston Pie,
Why does a chicken, I don't know why.
Ask me a riddle, and I reply:
"*Cottleston, Cottleston, Cottleston Pie.*"

Edward Bear.

(Winnie-the-Pooh, A.A. Milne)

Declaration

This thesis has been composed by myself and has not been submitted in any previous application for a degree. The work reported herein was executed by me, unless otherwise stated. Elements of this work appear in references [2] and [3].

A. N. Jackson

August 2001

For my father, who taught me to look up.

Acknowledgements

To Alastair Bruce, for unerring patience in the face of ignorance, for listening, for making sense and for making time, thank you. To my other close collaborators, Graeme Ackland and Nigel Wilding, thank you for putting up with my questions and helping me find my way through it all. To anyone in the group who has ever been forced to sit patiently through my more incoherent ramblings, cheers.

To my family and my friends, thank you for putting up with my occasional bad moods, my apparent inability to respond to ’phone calls, letters, emails and large piles of washing-up, and for your support, advice and general good humour.

To everyone who offered to proof-read this thing: If I took you up on the offer, then thank you, thank you, thank you. If I didn’t, consider yourself lucky.

I’d also like to thank my CD collection, but that would probably be pushing it.

Anj.

Table Of Acronyms

Acronym	Description	Section	Page
BHM	Broad Histogram Method	3.6.2	51
BRW	Bounded Random Walk	3.2.1	37
BTP	Blocked Transition Probability weight-function estimator	3.6.2	49
BVS	Blocked Visited States weight-function estimator	3.6.2	46
EB	Estimated By	3.1.3	30
FVM	Fixed aspect-ratio Volume Moves	3.2.2	39
GDI	Gibbs-Duhem Integration	5.6	104
IM	Integration Methods	4.3	84
LJ	Lennard-Jones	6.5.5	140
LS	Lattice Switch	17	60
LSMC	Lattice-Switch Monte Carlo	4.10	68
MB	Mobile Barrier weight-function estimator	35	142
MBTP	Mobile Barrier Transition Probability weight-function estimator	3.6.2	49
MBVS	Mobile Barrier Visited States weight-function estimator	3.6.2	46
MC	Monte Carlo	3	15
MCMC	MultiCanonical Monte Carlo	3.6	43
MCS	Monte Carlo Sweep	3.2.1	36
MD	Molecular Dynamics	3	14
MI	Multiple-Initialisation weight-function estimator	35	142
MITP	Multiple-Initialisation Transition Probability weight-function estimator	3.6.2	49
MTPM	Multiple-Initialisation Visited States weight-function estimator	3.6.2	47
NAG	Numerical Algorithms Group	6.4.1	126
NLM	Non-Local Mapping	4.3.1	60
NPT	Constant N, P & T (isobaric-isothermal) ensemble	5.3.3	102
NR	Newton-Raphson algorithm	5.3	100
NVT	Constant N, V, T (canonical) ensemble	5.3.1	100
PC	Predictor-Corrector	5.6.2	106
PCLS	Predictor-Corrector Lattice-Switch method	6.6.3	150
PEC	Predict-Evalute-Correct PC algorithm	25	106
PECE	Predict-Evaluate-Correct-Evaluate PC algorithm	5.4	105
PW	Present Work (this thesis)	4.3	84
RGS	Rare Gas Solids	6.7	156
RNG	Random Number Generator	3.3	40
RW	Random Walk	3.2.1	36
SHE	Single Histogram Extrapolation	5.2	95
SM	Shear Method	4.3	84
SOC	Single-Occupancy Cell model	6	20

continued overleaf...

...continued from previous page

Acronym	Description	Section	Page
SS	‘Strong’ Sampling	3.6.2	50
SSM	Site-to-Site Mapping	4.3.1	60
TH	Top-Hat algorithm	3.2.1	37
TP	Transition Probability weight-function estimator	3.6.2	48
TSSM	Transformed Site-to-Site Mapping	4.3.1	60
UVM	Unconstrained Volume Moves	3.2.2	39
VS	Visited States	3.6.2	46

Contents

Chapter 1

Introduction

The precise microscopic laws of nature are not the ‘end of physics’, rather they are the starting point from which the physical universe may be explained. The emergent behaviour of large systems of interacting particles cannot be predicted from microscopic considerations alone. To explain why the proton is stable, why tables do not melt, or how the human brain works, we must consider the collective behaviour of the large number of elements involved. This is the central idea behind many-body physics and in particular, equilibrium statistical mechanics.

More specifically, the work presented here falls into the ‘why tables do not melt’ category (otherwise known as condensed matter physics) and can be summarised by a single, deceptively simple question:

Can we predict the crystalline structure of a given material?

While the prediction of liquid-gas coexistence behaviour can often be performed successfully (at least for our simpler physical models), the prediction of the solid-fluid or solid-solid phase behaviour remains somewhat more challenging. The reason why the solid state presents such a problem can be understood by considering the two timescales one must confront when attempting to model this type of system.

The first of these timescales concerns the limitations of the computational techniques that are applied to this type of problem. For the accurate quantitative prediction of the phase behaviour of a given model system, we cannot rely on uncontrolled approximations, and so we have no choice but to employ numerical techniques; we must *simulate* the system’s behaviour, not *approximate* it. However, even using the largest computational resources currently available, and the simplest of physical models, it is only possible to simulate systems of a few thousands of particles for a few microseconds of real time.

Secondly, the physical dynamics of phase transitions involving solids are, broadly speaking, intrinsically slower than in the case of fluid/liquid/gas transitions. Solidification occurs at low temperatures,

and this, combined with the greater energy barriers involved in melting an ordered structure into a disordered phase, means that a longer timescale is required to observe a complete freezing/melting process than a comparable condensation/vaporisation process. The degree of hysteresis involved in first-order *structural* phase transitions is greater still, as a consequence of the slow physical restructuring processes involved in moving between different crystalline phases.

Therefore, if we perform physically realistic simulations of phase transitions, we will require significantly greater computational resources if a solid phase is involved. In general, the resources required to study this kind of phase transition are simply too large, rendering the problem effectively intractable. There are techniques capable of circumnavigating this problem, but they are few in number. In fact, the general lack of reliable theoretical techniques capable of dealing with this type of calculation has been described as a "continuing scandal" within condensed matter physics [4]. The truth of this statement is clearly emphasised by the fact that the structural phase behaviour of two of the simplest models of matter, the hard-sphere and Lennard-Jones potentials, has only now become computationally tractable.

The aim of this research is to develop a new computational technique capable of predicting solid-state phase behaviour. The central idea is that of a 'lattice-switch', a non-physical mechanism by which a simulation can leap between different structures in order to compare them. This idea was in existence before the work presented here began, and some initial investigations had been performed in the context of the hard-sphere solid [1]. This doctoral thesis is concerned with understanding the factors controlling the efficiency of this technique, broadening the range of results taken using it, and with its development into a widely applicable tool for the computational condensed-matter physicist.

In chapter 2, we present the two models that form the focus of this thesis (the hard-sphere and Lennard-Jones potentials), and outline the essential elements of the physical theory (statistical mechanics) upon which the work is based. Then, in chapter 3, we explore the (mostly standard) numerical and theoretical techniques by which some of the prescriptions offered by statistical mechanics may be implemented. This exploration leads to a discussion of the kind of mechanisms that a non-physical simulation can exploit. In chapter 4, these ideas will be developed into a practical computational technique in the context of the solid-state phase behaviour of a system of hard-spheres. The remainder of the thesis is concerned with developing the lattice-switch technique to deal with soft-potentials. This generalisation requires us to do more than simply measure the free-energy difference between the structures at a given temperature and pressure; we must also be able to extrapolate our results over a range of temperatures and pressures, so that the coexistence curve of the system can be determined. Chapter 5 will discuss a number of techniques relevant to this extrapolation, mostly drawn from the literature, but also including a few new ideas based upon the lattice-switch approach. These different methods and ideas will then be

drawn together in chapter 6, where the generalised lattice switch will be presented, and the structural phase behaviour of the Lennard-Jones solid will be explored. Finally, in chapter 7 we will discuss the successes and limitations of this work, and outline ways in which the work presented here may be developed further.

Chapter 2

Background

2.1 Models

This work will concentrate on a single class of model systems, that of pair-wise additive potentials. We choose to define our models via a configurational energy of the form,

$$E = \sum_{\langle i,j \rangle} \phi(r_{ij}), \tag{2.1}$$

meaning that we have a system of particles interacting through a simple pair-potential $\phi(r_{ij})$, where r_{ij} is the distance between a pair of particles i and j, and the total energy E of the system is the sum of all the individual $\langle i, j \rangle$ pair-wise interactions. Furthermore, this work will focus on the *classical* phase behaviour of two particular pair-wise interactions, the hard-sphere and Lennard-Jones potentials.

2.1.1 The hard-sphere potential

The hard-sphere potential (figure 2.1(a)) is of enduring interest due to the wealth of behaviour it demonstrates despite its simplicity. It is defined as

$$\phi(r_{ij}) = \infty \qquad \text{if} \quad r_{ij} < \sigma, \tag{2.2}$$

$$= 0 \qquad \text{otherwise}, \tag{2.3}$$

where σ is the diameter of the spheres. The interaction is purely repulsive, with the overlapping of particles prevented by an infinite energy barrier. As all accessible configurations have the same energy, temperature plays no role in the system's equilibrium phase behaviour, which is determined purely by the density. Even so, the hard-sphere system has been shown to possess a freezing/melting transition, and can also display glassy behaviour (see §4.1 for a more detailed discussion).

The question is, when the hard-sphere system freezes, what structure will it form? To the physicist it is intuitively clear that the most efficient way to pack together a number of spheres is by stacking hexagonal layers of these one on top of another, in much the same way one might arrange cannonballs.[1] Of these, however, there are any number of possible stacking patterns (for example hexagonal close-packed or face-centred cubic). As all these stackings have precisely the same density, the differences in the relative stabilities of the structures are determined entirely by many-body effects, and tend to be very small and difficult to determine. This a long standing problem in statistical physics [7], and will be attacked in chapter 4.

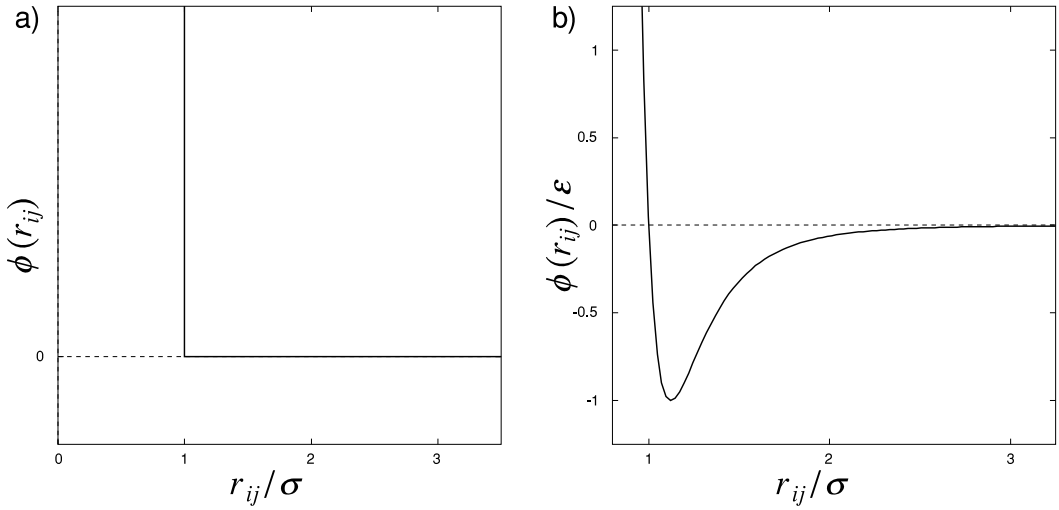


Figure 2.1: (a) The hard-sphere potential. Particles are prevented from overlapping by an infinite energy barrier for $r_{ij} < \sigma$, where σ is the diameter of the spheres. (b) The Lennard-Jones potential, which is composed of a steep repulsive term and a softer attractive term, leading to a well of depth ϵ.

2.1.2 The Lennard-Jones potential

This potential is characterised by a steeply repulsive core, describing the mutual repulsion of electrons as atoms are pushed together, and a weaker van der Waals attraction. When combined these opposing forces produce a minimum in the potential where, at low pressure, one might expect neighbouring particles to reside (see figure 2.1(b)). It has the form,

$$\phi(r_{ij}) = 4\epsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right]. \tag{2.4}$$

It is a 'soft' potential, because in contrast to hard-core potentials (such as the hard spheres introduced above) there is no finite range of r_{ij} for which the potential is infinite. As this system has both an energy scale (ϵ) *and* a length scale (σ), the properties of a system of Lennard-Jones particles will depend on

[1] To the mathematician this piece of guesswork, known as Kepler's conjecture [5], presents a formidable technical challenge and has only recently been proven [6].

both the density and the temperature. The Lennard-Jones system is also expected to favour close-packed structures, but the precise balance between the face-centred cubic and hexagonal close-packed phases remains unclear. This system's structural phase behaviour will form the subject matter of chapter 6.

2.1.3 What do we want to know?

Quite simply, we want to take our microscopic models of particle interactions and use them to predict the structure and properties of our model systems in the thermodynamic limit (i.e. as the number of particles, N, tends towards $O(10^{23}) \sim \infty$). The definition of the model indicates the macroscopic conditions that the system will be sensitive to, and we wish to predict how the properties of the system change as those macroscopic conditions are altered.

2.2 Phase behaviour

Despite the gigantic number of degrees of freedom that a macroscopic system possesses, it is found that real materials have well defined bulk properties (for example the observed density or the structure) that only depend on a few macroscopic control parameters, such as temperature and pressure. The different phases of a system are identified by those macroscopic bulk properties, and a *phase transition* simply refers to a point in the macroscopic-constraint space where a small change in those control parameters causes the properties of the system to change sharply.

It is generally possible to characterise a phase transition by the use of an *order parameter*, often defined to be zero in one phase and finite in the other. As well as identifying which side of the phase boundary a given system is on for a given set of macroscopic constraints, it also allows the phase transition to be classified as either first-order or continuous. First-order transitions are characterised by step-discontinuity in the order parameter, accompanied by a latent heat of transformation. For example, as we increase the pressure of gas, we may find that the density will jump from a low to a high value as the gas condenses into a liquid, while simultaneously warming its surroundings. In the case of continuous transitions, the step discontinuities are present in the *derivatives* of the order parameter, and no latent heat of transformation is observed.

The overall phase behaviour of a system can be described using a phase diagram, a schematic example of which is presented in figure 2.2. The lines indicate the location of first-order phase transitions, and the regions between them indicate which phase one would observe for any pair of values of the pressure and temperature. The liquid-gas coexistence curve of figure 2.2 terminates in a critical point, where a continuous transition from gas-liquid coexistence to a fluid state occurs. Continuous phase transitions

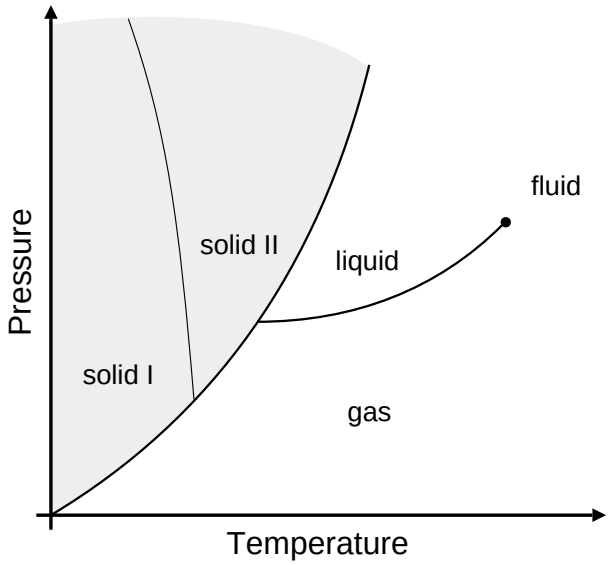


Figure 2.2: Schematic phase diagram, showing the kinds of phase transitions one may typically observe in a simple one-component system.

such as this present their own array of subtle challenges due to the infinite range of fluctuations which occur at the critical point [8]. Structural phase transitions can also be continuous, for example in the "soft-mode" instability of perovskites [9, p. 556]. This work will, however, be concerned with determining the relative stabilities of different crystalline solid states and with the locations of any *first-order* transitions between them.

2.3 Statistical physics of phase transitions

In order to use our microscopic laws to predict the macroscopic behaviour, we invoke the machinery of classical statistical mechanics. This potentially allows the bulk properties of our systems to be predicted from a knowledge of the microscopic interactions. A full exposition of this discipline is outwith the scope of this work, and only selected highlights of particular relevance will be described here. For a fuller treatment the reader is directed to reference [10].

We begin by attaching a probability, expressed in terms of a weighting factor, to each of the possible microstates. Each microstate embodies a unique position in configuration space, as defined by the set $\{\vec{q}\}$ where each $\vec{q}_i$ refers to the value of the ith degree of freedom. To each microstate we associate a certain microstate weighting, $W(\{\vec{q}\})$, some possible forms of which will be specified below. Figure 2.3 illustrates this idea schematically in a two-dimensional configuration space, the darker regions indicating the microstates with highest weight.

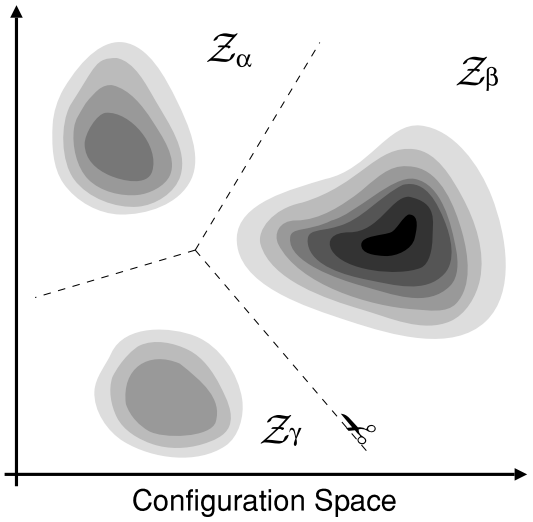


Figure 2.3: Schematic of configuration space, where the darker regions correspond to the microstate configurations with the greatest probability. This figure shows three phases (α, β and γ) which are separated by regions of unlikely configurations, and of which β has the greatest total weight.

The value of $W(\vec{q})$ depends on both the positions of the constituent particles ($\{\vec{r}\}$) and on the nature of the macroscopic constraints. For example in the canonical (NVT) ensemble, where the number of particles, the density and the temperature is fixed, the weight is defined by the configurational energy of the particles,

$$W(\{\vec{q}\}) = W_{NVT}(\{\vec{r}\}) = \exp\left[-\frac{E(\{\vec{r}\})}{kT}\right], \tag{2.5}$$

where k is Boltzmann's constant and T is the temperature. If one moves to the isobaric-isothermal (NPT) ensemble, where pressure, temperature and number are fixed (but the volume V can fluctuate), it can be shown that

$$W(\{\vec{q}\}) = W_{NPT}(\{\vec{r}\}, V) = \exp\left[-\frac{E(\{\vec{r}\})}{kT} - \frac{\mathrm{p}V}{kT}\right]. \tag{2.6}$$

Clearly, the probability of a system being in any individual microstate can be determined as the weight of that microstate divided by the total weights of all the possible microstates. That is,

$$P(\{\vec{q}\}) = \frac{1}{\mathcal{Z}} W(\{\vec{q}\}), \tag{2.7}$$

where the constant of normalisation ($\mathcal{Z}$) is the partition function of the entire system, defined as

$$\mathcal{Z} = \prod_i \left[\int \mathrm{d}\vec{q}_i\right] W(\{\vec{q}\}). \tag{2.8}$$

In order to determine the preferred phase, we must construct a 'total phase probability' from our microstates. This can be achieved by breaking the system's configuration space up into sections, each

corresponding to a different phase. This is analogous to cutting along the dotted lines of figure 2.3, and re-expressing the total partition function as no more that the sum of its (now separated) parts,

$$\mathcal{Z} = \mathcal{Z}_\alpha + \mathcal{Z}_\beta + \mathcal{Z}_\gamma. \tag{2.9}$$

The process of dividing up the configuration space is based on restricting the summation of $\mathcal{Z}$ to include only those configurations 'belonging to' a given phase, and is written as

$$\mathcal{Z}_\alpha = \prod_i \left[\int_\alpha d\vec{q}_i \right] W(\{\vec{q}\}). \tag{2.10}$$

Thus, the probability of any given phase can be written as

$$P_\alpha = \frac{\mathcal{Z}_\alpha}{\mathcal{Z}}. \tag{2.11}$$

This process is rather like taking the pieces cut from figure 2.3, and then comparing the amounts of laser-printer toner on each. The most likely phase would be determined from the piece of paper with the most toner on it.

These 'total phase probabilities' are easier to deal with when cast in terms of thermodynamic potentials. In the canonical ensemble, we construct the Helmholtz free-energy,

$$F = -k_B T \ \ln \mathcal{Z}, \tag{2.12}$$

and express the *relative* probabilities of two phases as a *difference* in free-energies,

$$\Delta F = F_\alpha - F_\beta = -kT \ \ln \left[\frac{\mathcal{Z}_\alpha}{\mathcal{Z}_\beta} \right] = -kT \ \ln \left[\frac{P_\alpha}{P_\beta} \right]. \tag{2.13}$$

In the case of hard-spheres, the energy of the system can be only zero or infinity, and so in the canonical ensemble the microstate weight (eqn. 2.5), can be only 1 or 0 respectively. Therefore in this case, our partition function is simply a count of the number of ways of arranging the spheres (Ω) and the free-energy becomes the familiar Boltzmann entropy,

$$S = k \ \ln \Omega. \tag{2.14}$$

In the case of the isobaric-isothermal ensemble, the thermodynamic potential is the Gibbs free-energy, and has the same form as in the case of the canonical ensemble,

$$G = -kT \ \ln \mathcal{Z}, \tag{2.15}$$

where the partition function now includes a sum over all possible volumes of the system,

$$\mathcal{Z} = \prod_i \left[\int \mathrm{d}\vec{r}_i \right] \int \mathrm{d}V \ W_{NPT}(\{\vec{r}\}, V). \tag{2.16}$$

The entropy and the Helmholtz and Gibbs free-energies are essentially quantifying the same thing, the number of ways of constructing the system, with the only difference being in the relative weighting of different microstates. For this reason, the term free-energy (and its symbol, F) will sometimes be used in this thesis to describe the appropriate thermodynamic potential for a given ensemble, whatever it may be.

It can be shown that as $N \to \infty$ the value of $f = F/N$ (the intensive free-energy density) will approach a steady value. This means that while the intensive Δf stays constant, the extensive ΔF must grow as N increases, so in this limit, the most likely phase will very rapidly grow to dominate over the diminishing probabilities of the less likely phases (eqn. 2.13). Therefore, the phase with the greatest probability and (by construction) the lowest free-energy at *finite* N will become the equilibrium phase in the thermodynamic limit (i.e. as $N \to \infty$). Phase boundaries will be identified as lines for which $\Delta f = 0$, i.e where the two phases are equally likely.

Given that we can examine structural phase behaviour using either the NVT or NPT ensemble, we must understand the consequences of making this choice. For a given phase, in the thermodynamic limit, the fluctuations of extensive variables become vanishingly small, and the bulk properties become sharply defined. For example, in the NPT ensemble, the density observed at a given pressure, which forms a Gaussian distribution for finite N, tends to a single value as $N \to \infty$. Thus, one would expect that, if one carried out a NVT experiment at this density, one would recover the same pressure as in the NPT case. In fact, it is generally true that for a single phase, the different ensembles should become equivalent in the thermodynamic limit [10, p. 227]. From a practical point of view, the equivalence of ensembles also means that one is free to choose whichever ensemble is most convenient for any given calculation. Note, however, that for coexisting phases this issue should be dealt with more carefully. For a given pressure, say, two *coexisting* phases can (and usually will) have different densities. Therefore, an analysis of coexistence which uses the same (constant) density for *both* of the phases is *never* equivalent to the constant pressure scenario. When the density difference is small, this may not cause a problem, but one must endeavour to test the validity of such an approximation if it is deemed necessary.

2.4 Techniques & solutions

We now narrow our focus onto the specific problem of solid-state phase behaviour. For a given system of interest, the usual interplay between theory and experiment should produce a list of likely candidates for the solid-state structure. Our quest is thus to determine which of these structures yields the lowest free-energy, as a function of the macroscopic constraints.

2.4.1 Pen & paper theory

Ground-state energies

When wishing to compare a number of different candidate structures without recourse to numerical simulation, one might proceed as follows. Firstly, one may attempt to evaluate the energy of each of the structures when all particles are sitting exactly on their lattice sites (i.e. at zero Kelvin). If this ground-state energy difference is large (compared with kT in the range of temperatures of interest) one may decide to consider this as sufficient. For the pair-potentials considered here, the ground-state energy is trivial to evaluate: For hard-spheres, *all* valid configurations have the same energy, and for the Lennard-Jones interaction one can calculate the sum of the pair-potentials over each lattice one wishes to consider (see §6.3, and also p.400 of [9]).

Unfortunately, while this approach *can* yield useful results, it clearly offers no way to distinguish the structures for the hard-sphere case. Furthermore, for the Lennard-Jones system the ground-state energy difference between *fcc* and *hcp* is so small that one cannot expect it to dominate the phase behaviour in the finite-temperature regime. Therefore in both cases the full free-energy must be considered, *including* the entropic contribution.

Dynamics & oscillations

Much has been learned about crystalline solids from the Einstein and Debye models. In the Einstein model [9, p.462], one assumes that the system can be represented as a set of independent harmonic oscillators with identical frequencies. In the case of the Debye model [9, p.458], a simple distribution of frequencies is constructed by considering the elastic behaviour of a continuous system, and adding in a cut-off at high frequencies corresponding the maximum frequency of the discrete, finite system. Both of these models have a single free parameter, and give valuable insight into the thermodynamics of the solid state. However, neither of them can be applied to all problems of structural phase behaviour as they stand. The assumption of a trivial frequency distribution removes all notion of geometry from the problem, and when comparing structures of identical density, like *fcc* and *hcp*, we are fundamentally concerned with geometry as this is the only way in which the structures differ. Any applicable theory must therefore include these geometrical differences, and a popular example of such a theory is the harmonic approximation.

The harmonic approximation & higher-order schemes

The harmonic approximation operates by assuming that soft-potentials can be accurately represented by simple spring-like interactions. That is, the interaction potential is truncated to second-order, in terms of the displacements of the particles from their ground-state ('perfect-lattice') positions. The resulting set of equations is exactly soluble, and accurately represents the original system when the displacements of particles from their sites is small, i.e. in limit of low temperature [9, chapter 22].

The harmonic approximation was first proposed in 1912 by Born & von Kármán [11, 12], when the concept of a crystal lattice was still in the process of being confirmed as an experimental reality via X-ray diffraction. Since that time the harmonic approximation has been modified in many different ways, all with the aim of extending the range of physical conditions over which one may produce accurate results. For example, the quasi-harmonic approximation [9] attempts to model the (more experimentally realistic) constant pressure behaviour by including the effects of thermal expansion. Many extensions of the harmonic theory have also been proposed in order to extend the range of temperatures for which the theory holds true (see [9, chapter 25] & [13, 14]). These are largely peturbative methods, where the effect of the 3rd and 4th order terms of the displacement-expansion upon the free-energy is estimated under the assumption that the amount by which these terms alter the free-energy is small.

However, comparisons of these methods (for example [14]) have illustrated the fundamental problem with this class of approach; the uncontrolled nature of the approximation. One cannot know, *a priori*, whether a given approximate theory will be able to predict successfully the behaviour of the true system over the temperature range of interest. The harmonic approximation can be used to predict the behaviour of a system in the low temperature *limit* (a fact which will be used to check our Lennard-Jones results, see §6.4), but that is all one can be sure of. Also, the harmonic approximation is not well suited to dealing with the hard-sphere solid, as the interaction is strongly anharmonic (i.e. very badly approximated by a quadratic potential well) and so cannot be expected to fit within this framework.[2] To ensure accuracy and generality, one is required to work without uncontrolled assumptions, and that can only be achieved by attacking the problem numerically.

2.4.2 Simulation

Before examining the ways in which simulation can be applied to structural phase behaviour, some general concepts common to all condensed-matter simulation work should be outlined. Simulations of this kind can be described quite aptly by the phrase 'balls in a box', as that is essentially what one

[2] Although, it is interesting to note that there are striking similarities between the two systems [15], and in one dimension these two system can be shown to be equivalent [16].

creates in mind of the machine. We define our set of positions for the particles ($\vec{r}_i$), within some (usually Cartesian) coordinate system, apply some appropriate boundary conditions, and then allow the system to explore its configuration space so that we might determine the properties of interest.

One might imagine choosing the boundary conditions in order to represent the containers used for real experiments, for example one could simulate inside a hard-walled box. However, simulations that use this type of boundary condition tend to be dominated by surface effects. In general, the bulk behaviour can be simulated much more effectively using periodic boundary conditions (figure 2.4). This arrangement means that the simulation cell no longer has any true surfaces, and rather it sees copies of itself in all directions. As long as this does not allow particles to interact directly with themselves, one finds that the thermodynamic limit can be *better* represented by a *smaller* system than in the case of a surface-dominated simulation. Unfortunately, this does not mean that the effects of finite system size disappear entirely, and care must still be taken to ensure that the simulated system is 'big enough' to accurately represent the behaviour in the thermodynamic limit.

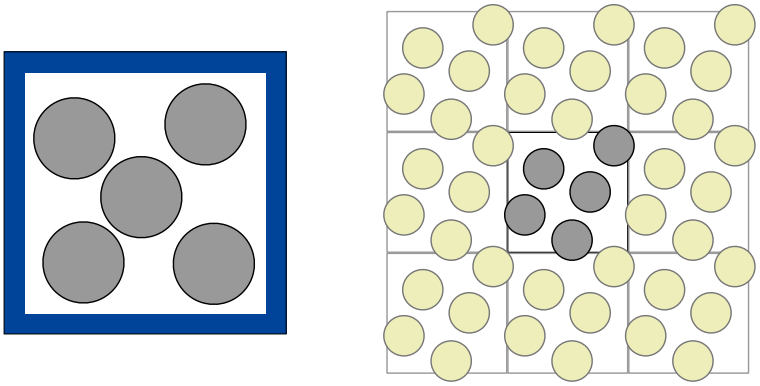


Figure 2.4: While a hard-walled container (left) can be used as the simulation box, periodic boundary conditions (right) give a more reliable representation of the bulk.

As well as the system size and boundary conditions, one must also choose the rules under which our simulation will evolve. There are two main options here, known as the Molecular Dynamics and Monte Carlo techniques.[3]

In classical Molecular Dynamics (MD) [17, chapter 4], the system is evolved deterministically by integrating Newton's equations of motion for the system. That is, one calculates the forces on every particle, and then updates the velocities and positions of those particles according to Newton's laws. As these laws conserve energy, this approach would seem to restrict the range of macroscopic environments one may consider. However, the MD scheme can be extended relatively easily: for example the simula-

[3]There are other approaches, such as hybrid Molecular Dynamics + Monte Carlo schemes like Brownian Dynamics [17], but these will not be considered here.

tion may be thermostated to allow the energy to fluctuate, while keeping the temperature fixed. Similar extensions have been carried out in order to use other ensembles (see [17, chapter 6] for more details).

The second approach, Monte Carlo (MC) simulation [17, chapter 3], is quite different. It is a general scheme for performing integrations, although much of the development of MC has focused on the type of 'balls in a box' simulation work considered here. New *trial* microscopic configurations are generated at random, and then accepted under certain conditions, in such a way as to explore the configuration space of the system of interest with the correct microstate probabilities. When these trial configurations are generated using so called 'physical moves' (which usually corresponds to jiggling single particles around over small distances) then the evolution of the system can be interpreted as a reasonable approximation of the true dynamics. However, the dynamical interpretation of a stochastic MC process must be performed carefully, and in cases where the dynamics are of interest (for example, when measuring diffusion coefficients) Molecular Dynamics may be the preferred approach.

Modelling the kinetics

The simplest approach one might take to determining the structural phase behaviour of a system is just to simulate its evolution directly, and then wait to observe which is the most likely phase. However, for this to work, it is necessary to define an order parameter which allows the simulation to identify which regions of the simulation box are currently in which phase (and so where the interfaces are), and this may not be straightforward when the differences between structures are small.

A more serious hurdle comes from the restrictions imposed by the distinctly finite amount of simulation time available.[4] This can mean that a simulation is unable to visit all the valid regions of configuration space in a reasonable time. To use the parlance of the field, the simulation can become non-ergodic. An extreme example of this problem is that of glassy materials, where systems can become non-ergodic on the experimental timescale [18]. In the context of structural phase transitions, the ergodic problem manifests itself as the hysteresis associated with first-order phase transitions. When moving across an *fcc* to *hcp* phase boundary (for example), one finds that the time required for the 'new' (*hcp*) phase to nucleate out of the 'old' (*fcc*) is extremely long, and one must push the system far into the *hcp* region in order to observe the transition on a reasonable (experimental) timescale.

For a simulation, this long timescale is simply inaccessible, and the system will be unable to freely explore the different structures available to it. Put another way, if we begin a simulation in the *hcp* phase (say), it tends to stay in that phase for the duration of the simulation. In this case it is usually true that

[4] The available computing power is, of course, constantly increasing. However, it is difficult to imagine when simulating macroscopically sized systems for experimentally long times will ever become feasible.

while the simulation is non-ergodic overall, it does possess 'in-phase ergodicity'. The lack of 'cross-phase ergodicity' has been overcome in some cases by using a 'flexible' periodic cell [19, 20], but this approach in not well suited to comparing structures that are not related by a simple deformation.

Along with the 'wait and see' approach, there is another related technique where a form of gentle encouragement is used. This is based around the idea of applying a strain to the simulation cell, in such a way as to *encourage* the structure to change. This technique was originally used to investigate the relative stabilities of the *fcc* and *bcc* structures of the Lennard-Jones solid [21]. However, like the 'flexible-cell' approach, this technique can only be applied to structures for which a strain-induced transition is possible. Also, the hysteresis associated with this transition makes it difficult to determine which of the phases is indeed thermodynamically stable for any given set of macroscopic constraints [22, 23].

Free-energy determination

Instead of performing long, physically realistic simulations, one can use simulation to determine the free-energies of the structures which one considers likely candidates for the equilibrium phase. The problem of calculating the free-energy is essentially that of determining $\mathcal{Z}$ (eqn. 2.12) which requires the solution of the $3N$-dimensional integral of all possible particle arrangements (eqn. 2.8). One might imagine calculating this integral by some multi-dimensional quadrature scheme but unfortunately, for even a small number of particles, this sum becomes impossibly large. Imagine a system of 100 particles in three dimensions, and then imagine considering just 10 possible values of each degree of freedom. Clearly, this approach would require the calculation and summation of some 10^{300} terms! This problem is exacerbated by the fact that in most cases, only a very small subset of the configuration space contributes strongly to the integral (i.e. $W(\{\vec{q}\})$ has a tendency to be sharply peaked). This means that even if one does apply the quadrature approach (as one can in the case of small N), the chances of including those configurations which contribute the most to $\mathcal{Z}$ is small.

The failure of the brute-force approach naturally leads to the idea of using a physical simulation to measure the free-energy. It has already been noted that a computer simulation can be used to explore configuration space with the appropriate microstate probabilities, thus circumnavigating the vast regions of unlikely configurations. The free-energy can *in theory* be directly determined as an observable within this type of simulation, but unfortunately the particle configurations which contribute most strongly to the free-energy tend to be suppressed, and the results are rarely statistically reliable [24].

However, despite the fact that the free-energies themselves are very difficult to determine numerically, simulations *can* be used to evaluate differences in free-energy, or differentials of the free-energy. In the former case, a single simulation *directly* calculates the free-energy difference of interest (§2.4.3). In

the latter case, a number of simulations are used to compute free-energy differences, using numerical integration to sum the derivatives of the free-energy along some inter-system pathway. The pathway usually lies between the system of interest and an ideal system for which the free-energy is known. These multi-stage integration methods will be considered in §2.4.4.

2.4.3 Direct differences

Any simulation which is free to visit two different phases can determine the free-energy difference between them from their relative probabilities (eqn. 2.13). This idea can be traced back to the seminal publication of Bennett [25], and underpins almost all of the work performed in this research. The algorithm presented in [25] is quite general, allowing the free-energy difference between two different systems to be estimated by measuring the energy cost of changing from one Hamiltonian to the other, *but without actually performing that 'move'*. The ratio of the partition functions for systems operating under two different Hamiltonians can be expressed as

$$\begin{aligned} \frac{\mathcal{Z}_2}{\mathcal{Z}_1} &= \frac{\sum \exp\left[-\beta H_2\right]}{\sum \exp\left[-\beta H_1\right]} && (2.17)\\ &\equiv \frac{\sum \exp\left[-\beta(H_2-H_1)\right]\exp\left[-\beta H_1\right]}{\sum \exp\left[-\beta H_1\right]} && (2.18)\\ &\equiv \overline{\left(\exp\left[-\beta(H_2-H_1)\right]\right)_1}\,, && (2.19) \end{aligned}$$

where the sum notation $\sum$ has been used as a shorthand for a sum over all microstates of the system, and where $\overline{\left(\exp\left[-\beta(H_2-H_1)\right]\right)_1}$ denotes the mean value of $\exp(-\beta(H_2-H_1))$, as determined while operating under Hamiltonian 1. Using this, the free-energy difference between the two systems can be determined as

$$\begin{aligned} \Delta F_{1,2} &= F_2 - F_1 \\ &= -kT\ln\left(\frac{\mathcal{Z}_2}{\mathcal{Z}_1}\right) \\ &= -kT\ln\left(\overline{\left(\exp\left[-\beta(H_2-H_1)\right]\right)_1}\right)\,. && (2.20) \end{aligned}$$

Unfortunately, as it stands, this technique is not practically useful in the majority of cases. This is because the particle configurations generated using Hamiltonian 1 must be very close to those which would be generated by Hamiltonian 2. If this is not the case, there will be large statistical fluctuations in the estimated value of $\frac{\mathcal{Z}_2}{\mathcal{Z}_1}$, and the timescale required for a sufficiently accurate measurement of $\Delta F_{1,2}$ will become prohibitively large. In general, the configurations associated with two distinct phases are so different that this approach cannot be used to compare them. For example, Moody *et al.* [26] attempted to use this style of calculation to determine the free-energy difference between the *fcc* and *hcp* phases of

a system of particles interacting via a Morse potential.[5] They found that a single-stage calculation could not be used, and consequently employed a sequence of seven separate direct-difference calculations, creating an inter-phase pathway from *fcc* to *hcp*. While the results of these calculation were good enough to argue that *fcc* was *probably* the stable structure, the errors in their calculation were of the order of 30% of the *fcc–hcp* free-energy difference itself.

2.4.4 Integration methods

In the absence of any more direct approach, most of the previous work in this field has concentrated upon building some kind of multi-stage inter-phase path between the two phases, like that mentioned above. For example, is it easy to show that the derivative of the Helmholtz free-energy (eqn. 2.12) with respect to the (inverse) temperature is

$$\left.\frac{\delta(F/T)}{\delta(1/T)}\right|_{NV} = E. \tag{2.21}$$

This tells us that the energy of a system at a given temperature and density can be used to estimate the gradient of the free-energy at this point. This idea is quite general; for any system with a partition function of the form

$$\mathcal{Z} = \sum \exp\left[-\lambda\mathcal{H}\right] \ , \tag{2.22}$$

the change in the free-energy due to a change in λ can be estimated as

$$\frac{\delta F}{\delta\lambda} = \overline{\mathcal{H}} \ . \tag{2.23}$$

One can use these differentials to determine the *total* free-energy by constructing an integration pathway, along which a suitable system parameter is changed so that the system of interest may be transformed into a system for which the total free-energy is *known*. For example, eqn. 2.21 could be used to measure how the free-energy changes as the temperature of a crystal is increased. This information could then be combined with the harmonic approximation estimate of the free-energy, which is exact in the low temperature limit, to provide an estimate of the total free-energy for that temperature.

To determine the free-energy difference between two candidate structures, one must first construct pathways from some ideal reference system to each of the structures of interest (figure 2.5). One then determines the nature of the observable to be measured in order to find the change in the free-energy along those pathways (via eqn. 2.23). The integration involves performing a series of simulations and then using some appropriate quadrature rule to determine the total free-energy of each phase. Providing

[5] A Morse potential is similar in form to the Lennard-Jones interaction, but is based on a combination of opposing exponential functions.

that the total free-energies have been determined to sufficient accuracy, the preferred phase will be revealed by their difference (eqn. 2.13). The reference system does not have to be the same for all phases under consideration, but good reference systems are hard to come by, and usually only one such system is employed in a given calculation.

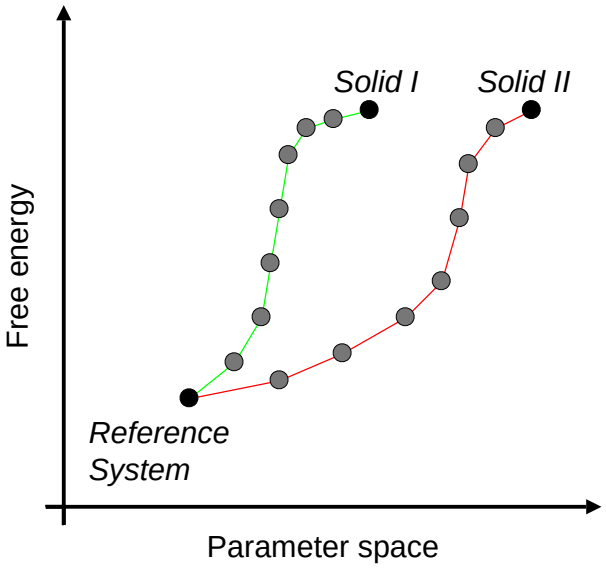


Figure 2.5: Schematic illustration of integration methods. The figure shows two paths in parameter space and the associated free-energy changes along the way. As shown here, the free-energy difference between the two target systems may be small on the scale of the free-energy difference between either of those systems and the reference system.

The reason why some reference systems are more useful than others can be seen by taking the example of integrating along a path of increasing density from an ideal gas to the solid state. In this case the integral one wishes to calculate would be (from equations 2.6 & 2.23)

$$F = F_{ig}(V_\infty) + \int_{V_\infty}^{V_{min}} P \, \mathrm{d}V, \tag{2.24}$$

where F_{ig} is the free-energy of the ideal gas, V_{min} is the (minimum) volume for which we wish to evaluate the free-energy, V_∞ is some volume sufficiently large that the interaction effects are negligible and the gas is effectively ideal.

Unfortunately, this approach will not work, because integration methods require that the observables (in this case, P) behave sensibly. This volume integration pathway goes through a region of strong hysteresis, i.e. a first-order phase transition from the fluid to the solid state. At this point the simulation ceases to sample configuration space efficiently, the pressure becomes ill-defined, and the free-energy evaluation fails. For this reason, whatever integration pathway one chooses, the integration is required

to be reversible (without hysteresis) and the simulation must retain both in-phase ergodicity and cross-phase non-ergodicity at all times.

The harmonic crystal

As indicated in §2.4.1, many systems are accurately represented by the harmonic approximation in the limit of low temperature. Therefore, as suggested above, expression 2.21 can be used to evaluate how the free-energy changes as the temperature is increased from some suitably low value. However this approach, first used by Hoover *et al.* [27], does have a number of limitations due to the restricted applicability of the harmonic approximation (see §2.4.1). Also, as will be seen in §6.4, calculating the harmonic free-energy can be somewhat troublesome, even for the simple systems considered here. Despite these complications, the harmonic crystal can act as an good reference crystal in many cases. It is therefore surprising that, during this work, very few publications have been found which use this approach.

The single-occupancy cell model

First introduced by Kirkwood [28], this method is based around an integration from low densities (where the system behaves as an ideal gas), through to the high density solid-state. This approach attempts to avoid the hysteresis associated with the first-order freezing transition by confining each particle to its Wigner-Seitz cell; every particle becomes associated with a lattice site, and the centre of each particle is not allowed to move outside the space which is nearer to its site that any other. In the low-density limit, we have a partitioned ideal gas, for which the free energy is known. In the opposite limit, the particles will no longer interact with the cell walls (because the particles will interact with each other first[6]) and so the single-occupancy cell (SOC) model becomes equivalent to the real system at sufficiently high densities (fig. 2.6).

While this approach has been used by many workers in the field (in the case of hard spheres, see for example Hoover & Ree [29], or more recently Woodcock [30]), its reliability has fallen into question [31]. At the very least, the cusp which appears in the P-V curve upon ‘freezing’ causes significant problems for the accurate evaluation of the free-energy, requiring a large number of integration points. At worst, this cusp may in fact be a weak first-order phase transition, and the integration pathway may not be reversible after all.

[6]Note that this is only true if the centre-of-mass of the system has been fixed with respect to the lattice, to avoid the entire system drifting into the walls. Unfortunately, it is often not made clear in the literature whether this constraint has been implemented or not.

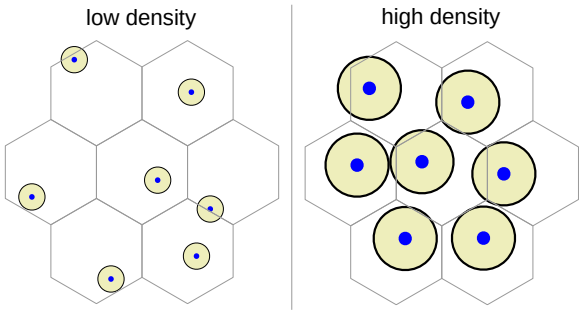


Figure 2.6: The single-occupancy cell model. At low densities (left) one has a relatively simple weakly-interacting system, and at high densities (right) the particles no longer 'see' the walls, and the true system is recovered.

The Einstein crystal

This approach, developed more recently than the previous two, is the most flexible and powerful integration technique to appear so far. For this reason, the method (as first described in [32, 33], and discussed in [17]) will be explored in some detail here, in the context of the hard-sphere solid.

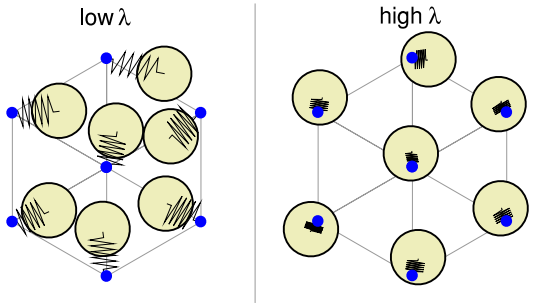


Figure 2.7: Einstein crystal integration. As the spring constant λ is increased, one moves from a system dominated by hard-sphere interactions (left) to an Einstein crystal (right).

The Einstein crystal is simply a system of particles that interact only with their lattice site, as if attached there by a spring. If one combines this interaction with the inter-particle hard-sphere repulsion, then one can see (fig. 2.7) that while at low spring-strength (small λ) the effect of the springs will be negligible, at high enough λ ($\lambda \rightarrow \lambda_{max}$, say) the particles will cease to interact with each other and we will recover the Einstein crystal, for which the free-energy $F_{einstein}$ is known analytically. The two systems are combined via a modified configurational energy,

$$E_{IM} = E + \lambda \sum_{i=1}^{N} (\vec{r}_i - \vec{R}_i)^2, \tag{2.25}$$

where the original expression has been augmented by a harmonic term based on the distance each particle

is from its lattice site ($\vec{R}_i$). In contrast to the temperature and density integrations considered earlier, our integration variable is no longer a thermodynamic variable but a model parameter. To determine the free-energy of the system we must evaluate the integral,

$$\Delta F = \int_{\lambda=0}^{\lambda_{max}} \overline{\left(\frac{dF(\lambda)}{d\lambda}\right)} d\lambda. \tag{2.26}$$

As shown in ref. [17, pp. 154,216] (see also eqn. 2.23) the value of $\overline{\left(\frac{dF}{d\lambda}\right)}$ can be estimated as the average of the square of the particle displacements, and eqn. 2.26 becomes,

$$\Delta F = \int_{\lambda=0}^{\lambda_{max}} \overline{\left((\vec{r}_i - \vec{R}_i)^2\right)_\lambda} d\lambda. \tag{2.27}$$

So by measuring the fluctuations in particle positions for a series of values of λ (from 0 up to some λ_{max}) one can determine the free-energy of the system.[7] The question remains, however, as to how large should λ_{max} be? Unfortunately, this question has no easy answer and leads to a series of complications to the technique (see fig. 2.8).

This complexity arises from the fact that one cannot actually take λ_{max} very close to ∞ in a numerical calculation. Thus, the 'ideal' system against which the hard-sphere solid is actually being compared is in fact a *weakly-interacting* Einstein crystal. In [33], the amount these weak interactions contribute to the free-energy is estimated from a 'virial-like' expansion ($F_{interaction}$) and an empirically determined correction to this expansion ($F_{correction}$). Example values of all of the contributions to the total free-energy are given in table 2.1.

Free-energy Contributions (per particle)	*fcc*$(N = 6^3)$	*hcp*$(N = 6^3)$
$F_{einstein}$	7.9568	7.9568
$F_{interaction}$	0.0183	0.0183
$F_{correction}$	-0.0024	-0.0008
ΔF	$-2.9754(10)$	$-2.9761(9)$
$F_{idealgas}$	0.9432	0.9432
$F_{C.o.M.}$	-0.0246	-0.0246
F_{total}	$5.9159(10)$	$5.9168(09)$
$F_{fcc} - F_{hcp}$		$-0.0009(14)$

Table 2.1: Example values of the components of the free-energy from the Einstein crystal calculation in ref. [33]. The errors associated with these estimates are given in parenthesis, indicating the uncertainty in the final digit. This convention will be used throughout this thesis.

So, having performed this long and complex calculation for both the *fcc* and *hcp* structures, we find that while the *total* free-energy has been very accurately determined (to within 0.02%), this is not enough to

[7] Note that as for SOC model, this is only true if the calculation is carried out in the centre-of-mass frame. In this case, this is because the unconstrained hard-sphere solid is free to move away from the lattice-sites used for the Einstein crystal calculation. This centre-of-mass diffusion means that the value of $\left\langle (\vec{r}_i - \vec{R}_i)^2 \right\rangle_0$ will diverge unless the centre-of-mass is fixed.

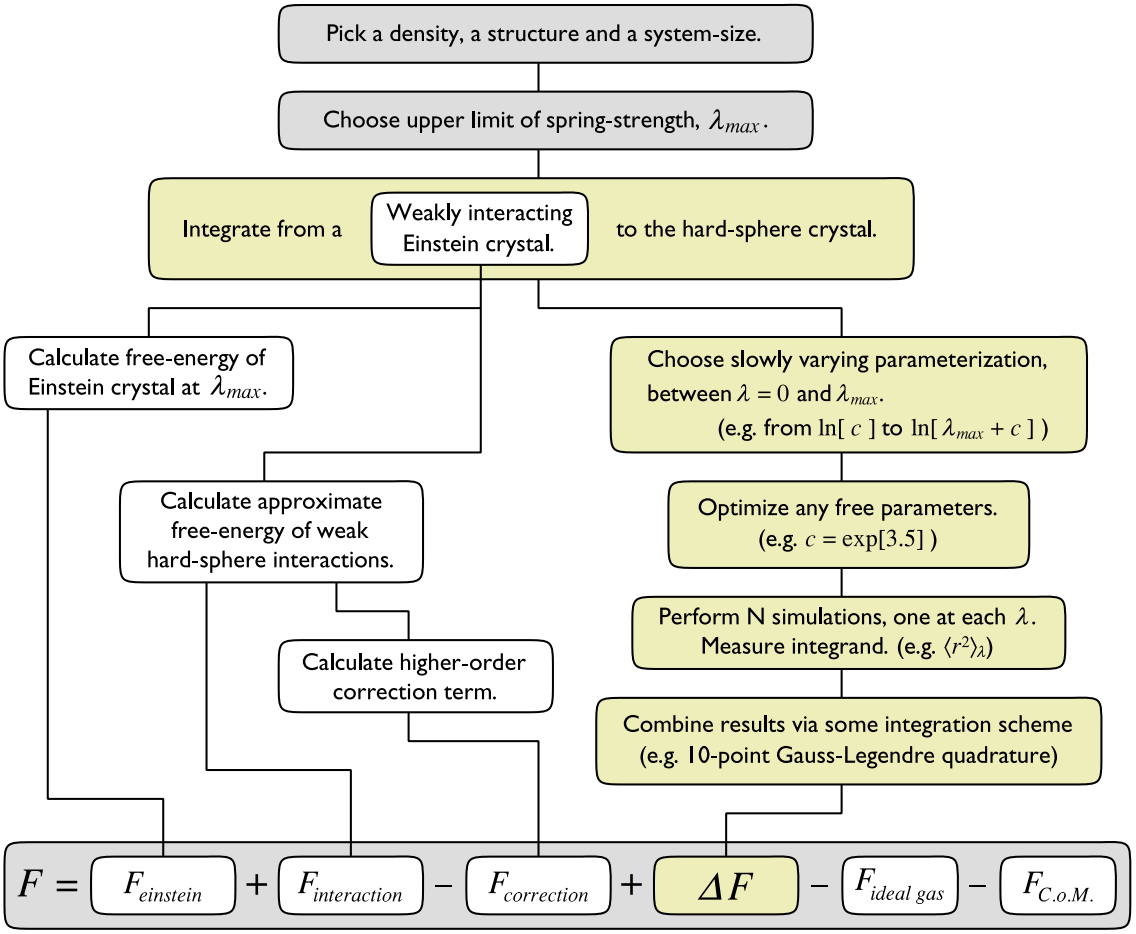


Figure 2.8: Flow diagram showing the major elements of the Einstein crystal integration method, as implemented in [33]. The final estimate of the free-energy is composed of six elements. Three are associated with the free-energy of the weakly-interacting Einstein crystal ($F_{einstein}$, $F_{interaction}$ & $F_{correction}$). ΔF is the difference between the 'ideal' and the hard-sphere systems. $F_{idealgas}$ is used to express the total free-energy relative to that of an ideal gas. $F_{C.o.M.}$ is a simple correction to compensate for the fact that the calculation is carried out in the centre-of-mass frame.

resolve the free-energy difference between *fcc* and *hcp*. In fact, the entropy difference determined using this method is significantly smaller than the difference in the empirically determined "correction to the correction" term, $F_{correction}$. To improve these results, one must begin the calculation again, increasing the value of λ_{max}, using more steps in the integration and using longer run-times at every λ.

The method outlined here is highly specific to the hard-sphere solid. For example, any long-range (soft) interaction, such as the Lennard-Jones potential, will never cease interacting no matter how strongly the particles are bound to their sites. A modified scheme [33] works by changing the configurational energy of the system such that

$$E' = \lambda_E E_{sys} + (1 - \lambda_E) \times \lambda_r (\vec{r}_i - \vec{R}_i)^2, \tag{2.28}$$

where E_{sys} is the configurational energy of the system of interest, λ_r is the spring constant of the reference Einstein crystal, and λ_E is the integration parameter. The integration now runs from $\lambda_E = 0$ (Einstein crystal 'switched on', E_{sys} 'switched off') through to $\lambda_E = 1$ (E_{sys} 'switched on', $E_{einstein}$ 'switched off'). This scheme avoids the calculation of the weak-interaction correction term, because the $\lambda_E = 0$ system is a true (non-interacting) Einstein crystal. However, the value of the reference-system spring-constant (λ_r) has a significant effect on the efficiency of the procedure and requires careful tuning.

2.5 Discussion

In general, one may reliably determine the total free-energies of two candidate structures using the Einstein crystal integration method. However, the procedure is complex, and after having painstakingly carried it out one may still find that the all-important *difference* in free-energy between structures has not been accurately determined. It would surely be preferable to find a way to focus in on this difference by designing a computational approach which is free to visit both structures, and thus determine the free-energy difference directly. This would allow a direct comparison of candidate structures in a single-stage process.

But how can one devise an algorithm capable of overcoming the ergodic block between structures? In the MD context at least, it is difficult to see how one may surmount this barrier. However, much of the power of the MC approach lies in the fact that as long as certain conditions are met, *any* move can be attempted, whether 'physical' or not.[8] As is it relatively easy to incorporate non-physical moves into a MC simulation, this freedom might allow one to find ways of overcoming the restrictions imposed by the distinctly finite simulation time available. The development of such a technique forms the core of this thesis, but before this development can begin we must explore the Monte Carlo method itself in

[8] Such moves *can* be implemented in the MD context, but the process is usually somewhat more straightforward in the case of an MC simulation.

some detail.

Chapter 3

Tools of the Trade: Part I

Monte Carlo simulation & extended sampling

In this chapter, the essential elements of Monte Carlo (MC) simulation are presented, along with some more recent ideas and developments.[1] This will provide all the basic tools required for the simulation of the two systems of interest. Also, the tools necessary for the *analysis* of the hard-sphere system (to be presented in chapter 4) will be outlined here. Those analysis techniques which are only relevant to the work done investigating the Lennard-Jones system will be described further on, in chapter 5.

3.1 Monte Carlo techniques

The Monte Carlo simulation technique was originally developed to study particle diffusion problems, for example to predict the yields of nuclear reactions in the complex geometries of reactors or storage facilities, and its name arises from a secret file on the Manhattan project [35, chapter 7]. Since that time, Monte Carlo simulation has grown to become one of the most widely used techniques for direct simulation in statistical mechanics. As an illustrative aid in the following exploration of Monte Carlo techniques, a simple toy model will be introduced and then referred to throughout.

[1] For a fuller introduction to the Monte Carlo method, the reader is directed to the textbooks by Allen & Tildesley [34] and Frenkel & Smit [17].

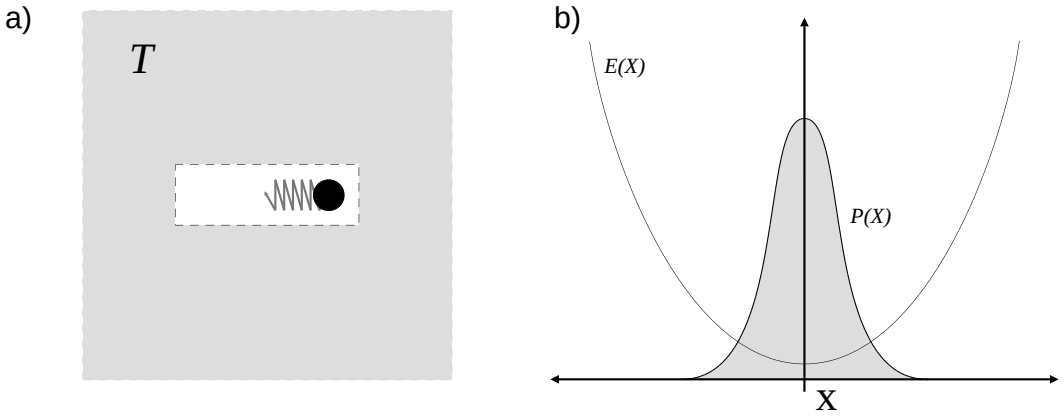


Figure 3.1: The toy model: (a) A one-dimensional harmonic oscillator in a heat bath of fixed temperature (T). (b) Graph showing both the harmonic potential $E(X)$ (thinner line), and the probability distribution of the particle's position (thicker line and filled grey), which is Gaussian for Boltzmann statistics.

3.1.1 A toy model

We consider a particle in a one-dimensional harmonic well, coupled to a heat bath (fig. 3.1(a)). That is, we have one degree of freedom (the particle's position x) and a simple configurational energy,

$$E = \lambda x^2, \tag{3.1}$$

where λ is the harmonic spring constant. The particle's exploration of x is chosen to obey Boltzmann statistics (eqn. 2.5), and so the probability distribution function (p.d.f.) of the particle's position is described by,

$$P(x)\mathrm{d}x = \frac{\exp\left[-\beta E\right]}{\mathcal{Z}}\mathrm{d}x = \frac{\exp\left[-\beta\lambda x^2\right]}{\mathcal{Z}}\mathrm{d}x\ , \tag{3.2}$$

where β is the inverse temperature ($1/kT$) of the heat bath. Clearly, the distribution $P(x)\mathrm{d}x$ is Gaussian, with a mean of $x = 0$ and standard deviation $\sigma = \sqrt{kT/2\lambda}$ (see fig. 3.1(b)). This system is sufficiently simple that we can evaluate the partition function exactly (using a standard Gaussian integral identity [36]),

$$\mathcal{Z} = \int_{-\infty}^{+\infty} \exp\left[-\beta\lambda x^2\right] \mathrm{d}x = \sqrt{\frac{\pi}{\beta\lambda}}\ . \tag{3.3}$$

Having determined $\mathcal{Z}$, all other properties of the system will follow via its derivatives. However, if we had been unable to find the partition function analytically, how could Monte Carlo techniques be used to solve the problem?

3.1.2 The simulation concept

Before describing the Monte Carlo technique in detail, it is necessary to introduce the language associated with computer simulation along with the most basic ideas involved. The central concept is that of

the discretization of time: the computer produces a sequence of microstates x_t, where t is the current simulation time-step, and the simulated system 'jumps' from one state to the next. Having accepted this, one can go on to produce the following (slightly abstract) formulation of a generic computer simulation, which is intended to introduce the essential form of input, process and output in the simulation context. Each step is also explained in terms of our toy model.

	In General	**Toy Model**
1	Describe the system.	Set up the position x as a variable and write a subroutine to calculate the energy.
2	Define some initial conditions.	e.g. $x_0 = 0$.
3	Generate a new (trial) state.	Pick a new x.
4	(Conditionally) accept this new state.	Decide whether to accept the new x or keep the old one (i.e. $x_{t+1} = x$ or $x_{t+1} = x_t$).
5	Repeat steps 3 & 4 *ad infinitum*.	Thus generating x_1, x_2, x_3...
6	Periodically measure ('sample') some property(ies) of interest.	e.g. the position, x_t, or the energy $E(x_t)$.

In this formulation, there remain two questions that must be answered: how do we generate new states from the old (the trial moves), and under what conditions will these moves be accepted? Having resolved these two issues, the simulation can be used to produce a series of microstates of the system. The particular choice of generation mechanism and acceptance rule will determine how we estimate properties of interest from the simulated microstates. For example, in a molecular dynamics simulation, the next microstate is generated by finite-difference integration of the equations of motion of the particles that make up the system. In this case, the new 'trial' state is always accepted, because it is exactly what should happen according to the laws of motion invoked.[2] Monte Carlo simulations tend to differ somewhat from this 'realistic' approach.

[2] Or at least, as 'exactly' as can be expected for a discrete-time computer simulation.

3.1.3 Random sampling

In this most basic form of Monte Carlo integration, one samples the configuration space of a system at random. That is, we generate the trial states by choosing random values for *all* the degrees of freedom of the system, and we always accept the move to this new state. The macroscopic control parameters (such as the temperature) do not affect the evolution of the system, and are simply used to calculate the system's properties from the simulation data. To concretize this idea, we shall examine the case of the toy model and how one can use random sampling to evaluate its partition function,

$$\mathcal{Z} = \int_{-\infty}^{+\infty} W(x)\mathrm{d}x \ , \tag{3.4}$$

where

$$W(x) = \exp\left[-\beta\lambda x^2\right] \ . \tag{3.5}$$

This 'integration over all microstates' can be approximated by an integral over a *finite* range of x,

$$\mathcal{Z} \approx \int_{-x_{max}}^{+x_{max}} W(x)\mathrm{d}x \ , \tag{3.6}$$

which becomes exact as $x_{max} \to \infty$. This approximate expression can be rewritten (via the mean value theorem [35]) as,

$$\mathcal{Z} \approx 2x_{max}\overline{W(x)} \ , \tag{3.7}$$

where $\overline{W(x)}$ is the mean value of the integrand over the range $-x_{max}$ to $+x_{max}$. This can be interpreted as a statement that the area under the $W(x)$-curve (i.e. the value of $\mathcal{Z}$) is equal to the area of a rectangle of width $2x_{max}$ and height $\overline{W(x)}$ (see fig. 3.2(a)).

Equation 3.7 tells us that if we can estimate the value of $\overline{W(x)}$ this will in turn produce an estimate of the value of $\mathcal{Z}$. This mean value can be calculated by randomly sampling $W(x)$ over the range $-x_{max} < x < +x_{max}$. That is, we generate $i = 1...N_s$ values of x which are randomly and uniformly distributed between $-x_{max}$ and x_{max}, then estimate the average value of the Boltzmann weight, which can be calculated using,

$$\langle W(x)\rangle_{x_{max}} = \frac{1}{N_s}\sum_{i=1}^{N_s} W(x). \tag{3.8}$$

The x_{max} subscript of the left-hand side is used to indicate that the average depends on the chosen value of x_{max}. The value of $\langle W(x)\rangle_{x_{max}}$ can be used to produce an estimate $\mathcal{Z}$ via a slightly modified version of eqn. 3.7,

$$\mathcal{Z} \overset{\text{EB}}{=} 2x_{max}\left\langle W(x)\right\rangle_{x_{max}} \ , \tag{3.9}$$

where we have used $\overline{W(x)} \overset{\text{EB}}{=} \langle W(x)\rangle_{x_{max}}$, meaning that the true mean value of $W(x)$ is *estimated by* (hence $\overset{\text{EB}}{=}$) the average value of $W(x)$, as measured in a Monte Carlo simulation. The same $\langle\mathcal{O}\rangle$ notation,

as defined in eqn. 3.8, will be used throughout this thesis to indicate a simulation-based estimate of the mean value of an observable ($\overline{\mathcal{O}}$).

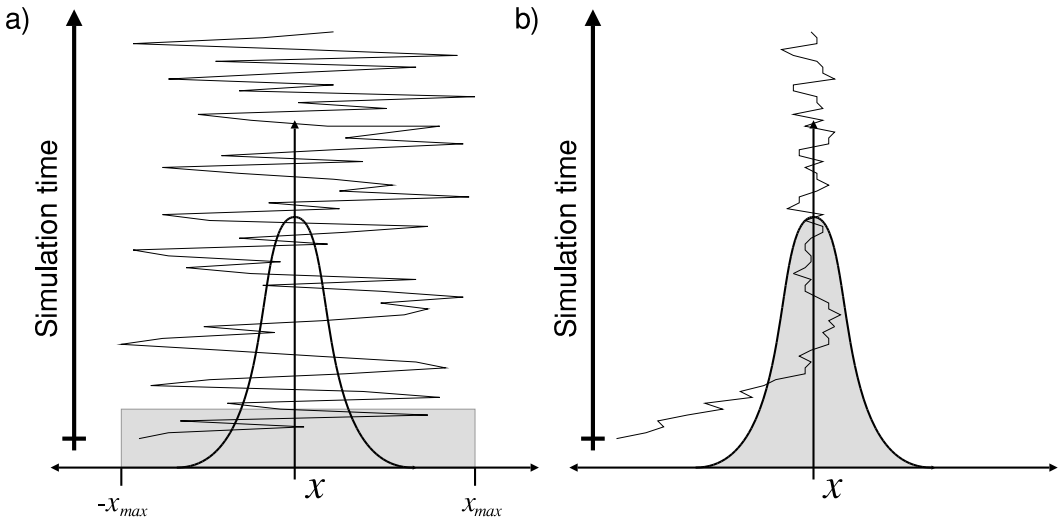


Figure 3.2: Comparison of the evolution of a simulation of the toy model for the two MC algorithms. (a) Random sampling: A broad range of x is sampled uniformly, and the Boltzmann factor does not affect the evolution of the simulation. (b) Importance sampling: By controlling the transition between states in a way that incorporates the Boltzmann weights, the microstates are automatically sampled in proportion to their importance.

While this process will work for this model, allowing the numerical value of $\mathcal{Z}$ to be estimated for any temperature, it is somewhat unsatisfying. It would be preferable if the integration process dedicated more time to the values of x for which the Boltzmann weight is high, as opposed to uniformly sampling some large and arbitrary range of x. Although the related problem of how to choose an appropriate value for x_{max} is something of an artifact of this particular toy model, for the kind of many-body systems we are really interested in it is generally true that only very small regions of configuration space are physically significant (i.e. of high statistical weight). For this reason, random sampling will invariably fail us when studying many-body systems, but it does provide a useful stepping stone on the way to the more powerful MC techniques at our disposal.

3.1.4 Importance sampling

The essence of this powerful technique lies in controlling the simulation's transitions between microstates in a way that ensures that those microstates are sampled according to their importance (see fig. 3.2(b)). In other words, the ways in which trial moves are generated and the rules under which a trial move is accepted or rejected are chosen such that each microstate of the system occurs with a frequency

proportional to its Boltzmann weight. If we consider measuring the average energy of our toy model,

$$\bar{E} = \int_{-\infty}^{+\infty} E(x)P(x)\mathrm{d}x \ , \tag{3.10}$$

where

$$P(x) = \frac{W(x)}{\mathcal{Z}} \tag{3.11}$$

then it is clear that if we can explore x with the correct (Boltzmann) probabilities $P(x)$, then the energy of the system can be estimated via a simple average of the observed energy of the simulation,

$$\bar{E} \overset{\text{EB}}{=} \langle E \rangle = \frac{1}{N_s} \sum_{i=1}^{N_s} E_i \ , \tag{3.12}$$

and indeed any other observable of the system can be treated in the same way.

The key feature of importance sampling is that the procedure does not rely on any prior knowledge of the actual value of $\mathcal{Z}$, but will nonetheless produce a sequence of microstates with the correct probabilities. To see how this is achieved, we must first re-express the evolution of the simulation in terms of a Markov chain.

3.1.5 The Markov chain

A simulation will form a Markov chain if the sequence of microstates it produces are generated such that the next microstate depends only on the current microstate, and not on any previous microstates in the sequence. We shall write the probability of observing any given microstate as $P(i)$, which is shorthand for the probability of the system occupying a particular point in configuration space, $P(\{\vec{q}\}_i) \prod_{\vec{q}} \mathrm{d}\vec{q}$ (for the toy model this is simply $P(x)\mathrm{d}x$). As the simulation's evolution is Markovian, the change in $P(i)$ (per unit time, i.e. per simulation step), $\Delta P(i)$, can be written as,

$$\Delta P(i) = \sum_{j \neq i} P(j)P(j \to i|j) - \sum_{j \neq i} P(i)P(i \to j|i), \tag{3.13}$$

where $P(i \to j|i)$ is the probability of a transition from i to j given that the system is in microstate i. The first sum on the right-hand side of equation 3.13 describes the flux of probability into microstate i from all the others, and the second sum describes the flux out of state i.[3] The concept of a probability flux is perhaps most easily grasped in this context by considering the ensemble corresponding to the toy model. Imagine we have a very large number (an ensemble) of toy-model simulations, each of which is exploring its configuration space independently from all the others. In this case, the probability flux from one microstate to another is essentially the *number* of systems which make this transition per unit time.

[3] Of course, in all the model systems considered here the configuration space is continuous, and so these sums in fact correspond to integrations over this space. However, for the sake of clarity the discrete sum notation will be used here.

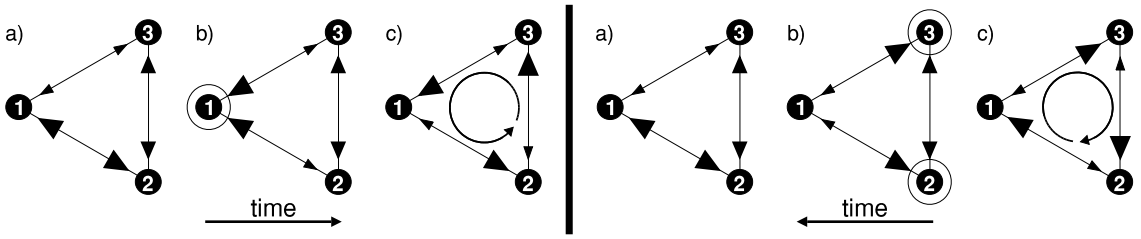


Figure 3.3: Time-reversal properties of a number of different transition schemes for a three-microstate system. The size of the arrows between two states i and j indicate the magnitude of $P(i)P(i \to j|i)$. (a) Detailed balance, which produces a steady distribution and is invariant under time reversal. (b) Example of an unphysical scheme which does not produce a steady distribution, and does not possess time-reversal symmetry. (c) A scheme that does produce a steady distribution, but for which there is a degree of circulation of probability, and so time-reversal symmetry is broken. Detecting the presence of this circulation has been suggested as a foolproof means of identifying whether a particular model satisfies detailed balance [37].

Our aim is to design a Markov chain which will converge onto the Boltzmann distribution. This will involve a period of equilibration, as the simulation approaches equilibrium from its initial conditions, and after this time the simulation should visit microstates according to the Boltzmann probabilities. To ensure convergence, it is necessary (but not sufficient, see below) that the transition probabilities should produce a steady distribution. In other words, the probability of any microstate should not change over time ($\Delta P(i) = 0$), and therefore in equilibrium,

$$\sum_{j \neq i} P(j)P(j \to i|j) - \sum_{j \neq i} P(i)P(i \to j|i) = 0. \tag{3.14}$$

There are any number of ways in which equation 3.14 may be satisfied,[4] but the usual choice is to ensure that the transitions conform to the condition of *detailed balance*. This means that the two sums are made to cancel term-by-term,

$$P(j)P(j \to i|j) = P(i)P(i \to j|i). \tag{3.15}$$

This choice of transition rule is very popular in the context of physical simulation. This is because the detailed balance condition embodies microscopic time-reversal symmetry, which is a property of all classical and quantum-mechanical equations of motion. The relationship between the detailed balance condition and time-reversal symmetry can be clarified by considering a simple system of three microstates, as shown in figure 3.3.

It should be noted that while detailed balance can easily be shown to produce a steady distribution (by satisfying eqn. 3.14), this does not constitute a proof that, given enough time, a Markov chain obeying detailed balance will *converge* on that steady distribution *from any and all other possible distributions*. However, mathematically rigorous (if somewhat dry) proofs have been around for some time, such as references [38] and [39] (which give proofs relating to systems of a finite and an infinite number of

[4] Indeed, for a system with continuous degrees of freedom (or a discrete system of infinite extend) there are an infinite number of possible solutions.

microstates, respectively). As well as these, a recent publication [40] has provided a more physicist-friendly proof that detailed balance does indeed ensure convergence to equilibrium.

The work in [40] also shows that *any* scheme which satisfies eqn. 3.14 will converge properly, and so as implied above detailed balance is indeed a stricter condition than is necessary. A similar argument was presented in reference [41], in order to justify the use of certain MC move generation schemes which had been shown to break detailed balance in its strictest form. The work in this thesis, however, will 'play safe' by ensuring that detailed balance is strictly observed at all times.

To make use of the detailed balance condition we substitute the Boltzmann probability of each microstate ($P(i)$, eqn. 2.7) into eqn. 3.15. We find that the troublesome partition functions cancel, leaving us with the requirement that

$$\frac{P(i \to j|i)}{P(j \to i|j)} = \frac{W(\{\vec{q}\}_j)}{W(\{\vec{q}\}_i)}. \tag{3.16}$$

From the arguments presented above, it should be clear that if transition probabilities are chosen such that equation 3.16 is satisfied, then the simulation should automatically converge onto the Boltzmann distribution.

3.1.6 The Metropolis algorithm

The original and most widely used algorithm in Monte Carlo simulation was introduced by Metropolis *et al.* [42]. To describe this algorithm, it is first necessary to split the transition probability into its two constituent parts: P_{gen}, the probability of generating any given move, and P_{acc}, the probability of accepting that move,

$$P(i \to j|i) = P_{gen}(i \to j|i)P_{acc}(i \to j|i). \tag{3.17}$$

The Metropolis algorithm requires that there is no bias in the way in which moves are generated, i.e. that

$$P_{gen}(i \to j|i) = P_{gen}(j \to i|j). \tag{3.18}$$

For the toy model, we might imagine generating new microstates by randomly changing x by a small amount. If this were done such that (for example) moves to the left were generated more frequently than moves to the right, then eqn. 3.18 would not be satisfied. The simplest way to ensure that 3.18 is satisfied in this case would be to generate moves to the left and to the right in a uniform and unbiased manner.[5] This property (eqn. 3.18) is referred to as the underlying symmetry of the Markov chain, and shifts the implementation of detailed balance entirely into P_{acc}. Metropolis *et al.* [42] showed that a

[5] An alternative to this condition is that any bias that *is* introduced into the generation of moves is *removed* by an additional acceptance condition so that eqn. 3.18 is recovered. Roughly speaking, if moves to the left are generated twice as often as moves to the right, then moves to the left should only be accepted half the time.

suitable acceptance condition is

$$P_{acc}(i \to j|i) = \min\left\{1.0, \frac{W(\{\vec{q}\}_j)}{W(\{\vec{q}\}_i)}\right\} , \tag{3.19}$$

which can be easily verified by substituting eqn. 3.19 into eqn. 3.16 for each of the three possible cases:

$$\begin{aligned} W(\{\vec{q}\}_j) &> W(\{\vec{q}\}_i) , \\ W(\{\vec{q}\}_j) &< W(\{\vec{q}\}_i) , \\ \text{and } W(\{\vec{q}\}_j) &= W(\{\vec{q}\}_i) . \end{aligned} \tag{3.20}$$

The only assumption required for this proof is a trivial one; that the values of $W(\{\vec{q}\}_i)$ are always positive, which must be true for the exponential weights given by eqns. 2.5 and 2.6.

3.2 Ensembles & algorithms

The precise form of the acceptance condition (eqn. 3.19) depends on $W(\{\vec{q}\})$, and so on the ensemble in which the simulation is to be performed. The work in this thesis will concentrate on two main physical ensembles (the canonical NVT and isothermal-isobaric NPT ensembles). The explicit forms of eqn. 3.19 for both of these cases will be presented here, along with the basic move-generation algorithms employed.

3.2.1 Canonical ensemble algorithms

For the canonical ensemble, the move-generation algorithm simply changes the positions of one or more particles. Using W_{NVT} from eqn. 2.5, we find the acceptance condition of any such move to be

$$P_{acc}(i \to j|i) = \min\left\{1.0, \exp\left[-\beta(E(\{\vec{r}\}_j) - E(\{\vec{r}\}_i))\right]\right\} . \tag{3.21}$$

One may interpret this rule as follows. When we attempt to change the position of any particle(s), we must evaluate the increase or decrease in energy associated with that change. If the energy of the system would be lowered by the move, then it is *always* accepted. However, if the energy of the system would be raised by that move, then the move is only accepted *some of the time*, with a probability based on the change in Boltzmann weight (see fig. 3.4). The role of temperature is therefore to control how likely a given increase in energy is, with an infinite temperature meaning that *all* moves will be accepted.

Trial moves

During the course of this work, a number of different particle move generation algorithms were implemented and compared. These belong to the class of moves known as ‘physical’ or ‘local’ moves, where

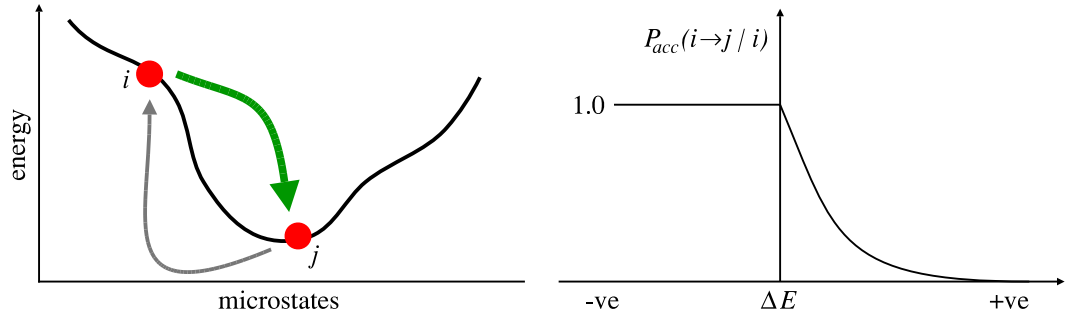


Figure 3.4: Schematic illustration of the Metropolis acceptance rule. For transitions which lower the system's energy (i to j, $\Delta E < 0$) the acceptance probability is always 1. For moves that cost the system energy (j to i, $\Delta E > 0$) the acceptance probability diminishes exponentially in the size of ΔE. The rate at which the acceptance probability diminishes is controlled by the temperature. The more energy that is 'available from the heat bath', the 'higher up' in configurational-energy-space the system can explore.

the next trial microstate is generated by changing the position of a single particle by a small amount, up to some maximum specified by a step-size parameter Δr. The units of time for a Monte Carlo simulation are usually defined in terms of these moves; simulation times are measured in Monte Carlo Sweeps (MCS), where one MCS is the amount of time required to attempt one 'physical' move of every particle in the system.

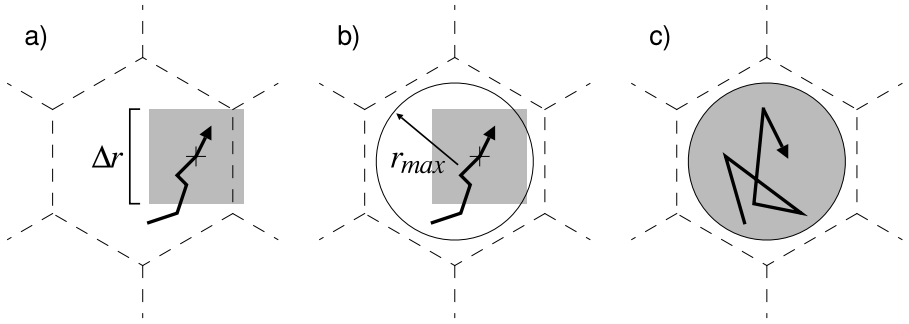


Figure 3.5: Illustration of the three particle-move types used in this work: (a) the random-walk, (b) the bounded random-walk and (c) the top-hat.

The random walk (RW) This is perhaps the simplest MC particle move, where the position of the chosen particle is changed by the addition of a small, randomly generated vector (fig. 3.5(a)). In this case, the parameter Δr refers to the side-length of the cube from which displacement vectors are generated. Although this type of algorithm is usually the best approach for fluid-phase simulation, it can present drawbacks when the system under consideration is crystalline. These drawbacks occur because at low densities, the particles will no longer be caged-in by their neighbours. Therefore, the RW algorithm makes it possible for particles to swap positions, and can even allow the entire system to spontaneously melt. This behaviour is essentially physical, but if the crystalline microstates are the only ones we wish

to consider, this kind of structural failure must be avoided.

The bounded random walk (BRW) In this case, each particle is associated with a lattice site, and is constrained to stay near that site (fig. 3.5(b)). The particle moves are implemented in much the same way as the RW algorithm, but as well as a maximum step-size Δr, there is a maximum possible displacement from this site r_{max}. Any moves that would take the particle further away from its site are rejected. Like the single-occupancy cell model (§2.4.4), the simulation should ensure that the lattice sites are kept in the centre-of-mass frame in order to stop the entire system from drifting into the cell walls.

The top-hat (TH) This algorithm is a simplified version of the BRW scheme, where a new trial position is generated completely at random within the bounded volume around each lattice site (fig. 3.5(c)). There is no random-walk parameter (Δr), and the dynamics are defined purely by the maximum displacement r_{max}. As in the case of the BRW, the simulation should be carried out in the centre-of-mass frame. Both the BRW and TH schemes will successfully avoid the melting problem, but the question of which is most efficient still remains. Furthermore, we must ensure that the cut-off required by the BRW and TH algorithms does not significantly affect the simulation results. The simulation must avoid the fluid phase without also refusing access to a significant proportion of the crystalline microstates.

Acceptance rates

All of these algorithms must be made to maximise their efficiency, i.e. to maximise the rate at which the simulation explores its configuration space. This is done by tuning the value of the maximum step-size Δr in order to minimise the autocorrelation time of the observable(s) of interest. Whether a given particle move is accepted or not is controlled by eqn. 3.21, which depends on the change in energy associated with that trial move. As we shall see later, very small step-sizes mean that trial moves are almost always accepted, whereas larger step-sizes mean that moves are accepted less frequently but that each move will correspond to a bigger leap through configuration space. The choice of step size is best expressed in terms of the 'acceptance rate', which is simply the number of accepted moves divided by the total number of attempted moves. The role of the acceptance rates in controlling the efficiency of our simulations will be explored in §4.10.1 and §6.6.1.

3.2.2 Isobaric-isothermal ensemble algorithms

In the case of the NPT ensemble, the simulation must sample the configurations associated with all the possible simulation-cell volumes, as well as those associated with the different possible particle configurations. To do this, it is possible to add simple volume dilations, which alter the volume of the simulation cell without modifying the particle positions (figure 3.6, left). The decision to accept this move would then be based upon the Boltzmann weight W_{NPT} given by equation 2.6. However, there is a more efficient method with which to implement volume moves (see [17, sec. 5.4] for details of this and other types of volume move). The idea is that instead of just changing the volume by some scaling factor, all particle positions within the periodic cell are also scaled (see figure 3.6, right).

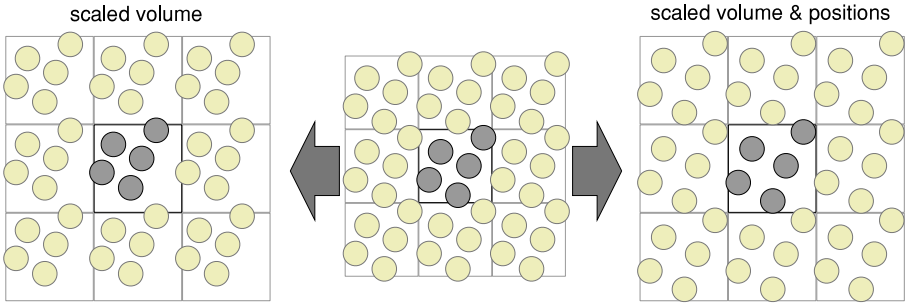


Figure 3.6: Volume moves. While just changing the volume of the periodic cell (left) leads to a simpler acceptance condition, the simulation takes some time to re-adjust to the new volume. If the particle positions are scaled with the volume, the resulting configuration is nearer the equilibrium configurations for that volume, and so the simulation is more efficient.

This all-scaling scheme requires the particle positions to be re-expressed as,

$$\vec{r}_i = \overline{\overline{L}}\vec{s}_i \tag{3.22}$$

where the deformation matrix $\overline{\overline{L}}$ has zero off-diagonal elements, and has parameters a, b and c along the diagonal. Under this constraint, the $\overline{\overline{L}}$ matrix describes a mapping from a cuboid simulation cell of dimensions $a \times b \times c$ onto a unit cube. Substituting this form into the equation for the partition function (2.16) we find,

$$\mathcal{Z} = \int \mathrm{d}V \left[\prod_i \int \mathrm{d}(\overline{\overline{L}}\vec{s}_i)\right] \exp\left[-\beta E(\{\vec{r}\}) - \beta PV\right], \tag{3.23}$$

$$= \int \mathrm{d}V \left[\prod_i \int \mathrm{d}\vec{s}_i\right] V^N \exp\left[-\beta E(\{\vec{r}\}) - \beta PV\right], \tag{3.24}$$

$$= \int \mathrm{d}V \left[\prod_i \int \mathrm{d}\vec{s}_i\right] \exp\left[-\beta E(\{\vec{r}\}) - \beta PV + N \ln V\right], \tag{3.25}$$

$$\tag{3.26}$$

where we have use the fact that the volume of the simulation cell, V, is equal to the product abc. Therefore, in this scaled system the Boltzmann weight is now,

$$W'_{NPT} = \exp\left[-\beta E(\{\vec{r}\}) - \beta PV + N \ln V\right], \tag{3.27}$$

and so the acceptance condition becomes,

$$P_{acc}(i \rightarrow j|i) = \min\left\{1.0, \exp\left[-\beta(E(\{\vec{r}\}_j) - E(\{\vec{r}\}_i) + (PV_j - PV_i)) + N \ln\left(\frac{V_j}{V_i}\right)\right]\right\}. \tag{3.28}$$

By comparing this to the NVT acceptance condition (eqn. 3.21), it is clear that the NPT acceptance condition has the same form as the NVT case, but with the addition of two extra terms. The PV term embodies the external pressure by encouraging moves to smaller volumes, whereas the $N \ln V$ term favours moves that increase the volume available to each particle (i.e. moves that increase the entropy).

Trial moves

Two different volume-move generation schemes were employed during the course of this research, with the only difference being that in one case the shape of the periodic simulation cell is fixed, and in the other the lengths of the sides of the simulation cell are allowed to fluctuate. The only constraint is that the trial moves must be generated in an unbiased fashion, in order that eqn. 3.18 is satisfied.

Fixed aspect-ratio volume moves (FVM) Under this scheme, new volumes are generated in much the same way as new particle positions, by altering the current volume randomly by some small amount uniformly distributed over a range $V \rightarrow V + [-\Delta V, +\Delta V]$. Having generated the new volume, the elements of the deformation matrix $\overline{\overline{L}}$ are determined under the condition that the aspect ratio of the cell remains constant (i.e. the ratio of any two side lengths, for example a/b and a/c, are unchanged). This is not the only way to perform fixed aspect ratio volume-scaling moves, for example a biased move-generation scheme that performs a random walk in $\ln V$ instead of V has been used in reference [43]. However, this latter scheme was not implemented during this work.

Unconstrained aspect-ratio volume moves (UVM) Here, the side-length parameters (a, b and c) are not constrained to yield a fixed aspect ratio. The volume-move procedure is split into three separate steps, where each of the three parameters (a, b and c) are individually modified by a small amount, $[-\Delta l, +\Delta l]$. After each of the three parameters has been modified, the new configurational energy is calculated and the same acceptance condition (eqn. 3.28) is used to decide whether the move should be accepted.

Acceptance rates and attempt frequencies

As in the case of particle moves, we have the freedom to tune the size of the volume moves (and so alter the associated acceptance rate) in order to maximise the efficiency of the simulation. This issue is explored in §4.10.1 & §6.6.1. We also have the freedom to choose how often to attempt a volume move, but for this we choose to use the standard prescription [17, §3.4] [34, §4.5], where volume moves are attempted at random intervals, with an average frequency of once per Monte Carlo sweep.

3.3 The random factor

For any of the algorithms outlined here to work, we need a source of random numbers with which to carry out the Monte Carlo moves. However, computers are strictly deterministic machines (excepting hardware faults, which are extremely rare in modern computers) and so there can be no such thing as a perfect random number generator (RNG). The sequence of numbers generated by a computational algorithm will be well defined, and will always (eventually) repeat. Many years of research have been dedicated to the problem of *pseudo-random* number generators (see ref. [34, §6] for an introduction), but the best we can hope for is that, for the volumes of random numbers we require, the sequence will fail all known tests designed to detect ordered behaviour.

Unfortunately, some simulation algorithms can be extremely sensitive to any correlations in a random number generator's output, and the only way to ensure that all is well is to check that we recover statistically indistinguishable results whatever RNG we choose. So, while most of this work was performed using one random number generator (RAN250/521), occasionally a simpler algorithm (RAN250) was employed in order to check the reliability of our results. Briefly, the RAN250 algorithm is a straightforward linear congruential algorithm, whereas the RAN250/521 generator combines the results of two independent (congruential) algorithms to produce an overall sequence of significantly greater statistical fidelity than the RAN250 alone. For more information on these algorithms, see refs. [44, 45, 46] and references therein. No RNG dependence was detected during this research.

3.4 Error analysis

Given that we have a successful importance sampling simulation, the question remains as to how we might estimate any properties of interest from the simulation data. We have already seen that the energy can be estimated by periodically measuring the simulated values and finding the average, and this is true

for any observable

$$\overline{\mathcal{O}} \overset{\text{EB}}{=} \langle \mathcal{O} \rangle = \frac{1}{N_s} \sum_{i=1}^{N_s} \mathcal{O}_i \ , \tag{3.29}$$

where N_s is the total number of times the value of $\mathcal{O}$ was sampled. However, we need to be able to estimate our statistical uncertainty in this average, that is we need to compute the standard error associated with the mean value of $\mathcal{O}$.[6] It is well known (see e.g. [48]) that if the N_s observations of $\mathcal{O}$ are uncorrelated, then the standard error of the mean can be calculated as

$$\sigma(\overline{\mathcal{O}}) = \frac{1}{\sqrt{N_s}} \left[\langle \mathcal{O}^2 \rangle - \langle \mathcal{O} \rangle^2 \right] \ . \tag{3.30}$$

Of course, as a Markov chain is generated by making a long series of (usually small) changes to the system, the sequence of microstates will be correlated to some degree. But how can we ensure that these correlations are not skewing the averages, and that the statistical errors are reliable? One option is to calculate the autocorrelation function of the data, and then estimate the autocorrelation time from that function. One can then ensure that the data is only sampled on a timescale significantly greater than the autocorrelation time. However, this process is rather time consuming and slightly arbitrary (as it depends on the particular form chosen to fit to the autocorrelation function data). A more straightforward approach is known as block averaging (see [17, Appendix D.3]), which avoids the explicit calculation of the autocorrelation time. The individual N_s measurements are gathered into N_b blocks of length L_b, so we have $B = 1 ... N_b$ individual block averages,

$$\langle \mathcal{O} \rangle_B = \frac{1}{L_b} \sum_{i=1+(B-1)L_b}^{L_b B} \mathcal{O}_i . \tag{3.31}$$

From this set of averages, we can compute the overall average,

$$\langle \mathcal{O} \rangle = \frac{1}{N_b} \sum_{B=1}^{N_b} \langle \mathcal{O}_b \rangle_B \ , \tag{3.32}$$

and the associated variance,

$$\sigma_b^2(\mathcal{O}) = \frac{1}{N_b} \sum_{B=1}^{N_b} \langle \mathcal{O} \rangle_B^2 - \langle \mathcal{O} \rangle^2 \ . \tag{3.33}$$

It can be shown that the following error estimate (based on the assumption that the blocks are uncorrelated),

$$\sigma(\overline{\mathcal{O}}) = \frac{\sigma_b(\mathcal{O})}{\sqrt{N_b}} \tag{3.34}$$

will be a constant for all block sizes which are sufficiently large that the individual block averages are uncorrelated. Therefore, we can check that the block size is sufficiently large by calculating $\sigma(\overline{\mathcal{O}})$ for a number of different block sizes (a range of values of N_b) and then adopt the value of $\sigma(\overline{\mathcal{O}})$ from the range of N_b for which it is constant.

[6] For an excellent introduction to error-handling techniques in general, see Taylor [47].

3.5 Successes & failures

Importance sampling Monte Carlo, as outlined so far, is a widely used and powerful tool for investigating many-body systems. However, there are a number of important limitations. For example, as mentioned in §2.2, continuous phase transitions present some serious difficulties for the computational physicist. A Monte Carlo simulation performed near a critical point tends to suffer from 'critical slowing down', meaning that the progress of the simulation through configuration space becomes extremely slow. Also, such a simulation will suffer from serious finite size effects, because at a critical point the spatial correlation length (i.e. the size of the fluctuations) diverges, and so only an infinite system can give a truly accurate representation of the system's behaviour. However, much progress has been made by using 'smart' MC moves to speed up the simulation's progress (e.g. cluster moves [17, Chapter 12]), and by using finite-size scaling techniques to determine what a series of finite (differently) sized simulations can tell us about the infinite system [49].

However, for the work under consideration here, the core limitation of an importance sampling MC simulation is that it cannot determine the absolute free-energy of a system. However, an MC simulation can estimate the *difference* in free-energy between two phases *if* it is free to visit them both (§2.4.2). Unfortunately, for the cold, solid-state systems under consideration here, the configuration space consists of a number of distinct fragments, corresponding to different phases, each of which is not only (formally) metastable, but on any reasonable *simulation* timescale is effectively stable.

An analogous problem can be invented by making a small alteration to our toy model. Instead of a single harmonic well, imagine that there are two minima in the configurational energy (see fig. 3.7). An importance sampling simulation (using 'physical' moves) will tend to linger in one well or the other, only crossing the barrier occasionally (see fig. 3.7(a)). If we make the barrier higher, the system spends less and less time moving between the wells, i.e. the probability of crossing from one well to the other gets smaller. With a high enough barrier, the simulation will become stuck (on the computational timescale) in one of the two wells, and never 'see' the other one(see fig. 3.7(b). Note that lowering the temperature of the simulation shown in figure 3.7(a) would have the same effect as increasing the barrier height.

3.6 Extended sampling

The essential idea behind extended sampling is to determine the behaviour of the system for some microstate weighting scheme other than the Boltzmann distribution, i.e. the simulation is biased. This biased simulation can overcome the probability barrier between two regions of configuration space, and so the simulation can visit both of them. The trick, however, is to do this in such a way as to allow us

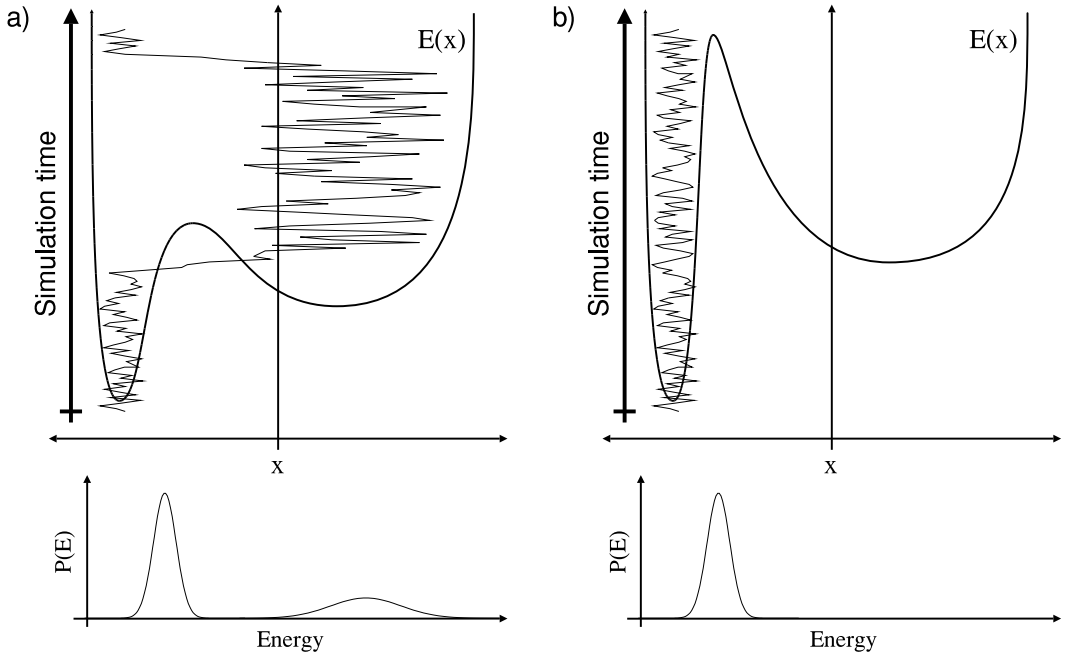


Figure 3.7: The toy model with a double-well potential: (a) If the barrier is relatively small, the simulation will move between the two on a reasonable timescale. (b) As the barrier is made taller, the simulation tends to get stuck in whichever basin it started off in.

to calculate what would have happened *if* the Boltzmann distribution *had* been used, so that we may remove the bias after the simulation has run its course.

This idea has been around for some time, and there are a number of different biased sampling techniques available for a number of different contexts (see [17, 24] and references therein). However, this work will focus on one particular technique, generally referred to as multicanonical Monte Carlo (MCMC). The genesis of this method can be traced back to the "umbrella sampling" of Torrie & Valleau [50], which was later generalised by Berg & Neuhaus [51] to create the multicanonical sampling algorithm. The name, incidentally, comes from the idea of a single simulation that visits many different temperatures, i.e. it samples many different canonical ensembles. The implementation of MCMC used in this work follows the form given in the Smith & Bruce review article [24].

3.6.1 Multicanonical Monte Carlo

The essence of the kind of problem which multicanonical Monte Carlo is well suited to dealing with is presented in figure 3.8(a). We have some macroscopic order parameter $\mathcal{M}$ for which there are two distinct sets of values, corresponding to two separate regions of configuration space in which the Boltzmann weight is high. In the case of our double-well toy model (fig. 3.7), this order parameter could be the energy. The order parameter space is assumed to be discretized, such that for each microstate i, its

associated macrostate value $\mathcal{M}$ belongs to a discrete macrostate bin $\mathcal{M}_m$ where there are $m = 1...N_M$ macrostates altogether. The probability of visiting the macrostates of $\mathcal{M}$ in between the two phases becomes (exponentially) small as one moves away from the (approximately Gaussian) peaks, and so the probability of crossing this region in an importance-sampling simulation is very small.

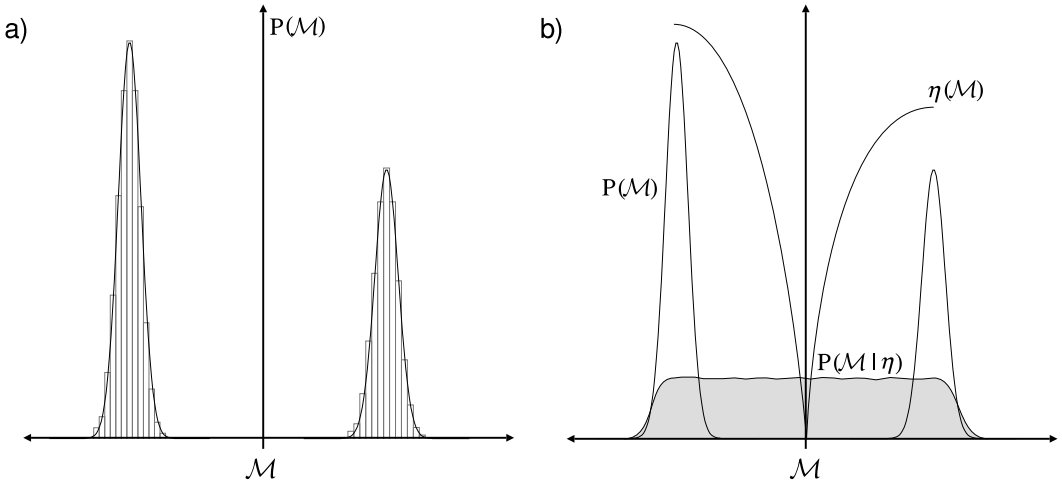


Figure 3.8: Illustration of the essential ideas of multicanonical Monte Carlo. (a) We have some double-peaked distribution, which we wish to explore in full. The discrete $\mathcal{M}$ states are shown as a grey histogram. (b) A multicanonical simulation can visit all relevant $\mathcal{M}$-values uniformly (grey-filled plot), using an appropriate weight function η (thick line). The true distribution can be determined by removing the bias introduced by the weights. The curves have been drawn as continuous functions for reasons of clarity.

In multicanonical Monte Carlo, the original Boltzmann weight is augmented by a 'weight function' which depends on $\mathcal{M}$ (via the index m),

$$\mathcal{W}(\{\vec{q}\}) = W(\{\vec{q}\}) \exp\left[\eta_m\right]. \tag{3.35}$$

This can be used to modulate the probability of visiting each macrostate of $\mathcal{M}$, and so potentially allows the full range of $\mathcal{M}$ to be explored under any distribution which seems appropriate. The usual choice [24] is for the macrostate space to be explored approximately uniformly,[7] as illustrated in figure 3.8(b).

The bias introduced by the weight function can be removed from the simulation data as follows. Any measurements of an observable taken in the m'th macrostate have had their weighting augmented by a factor $\exp\left[\eta_m\right]$, and so by reweighting each observation by a factor of $\exp\left[-\eta_m\right]$, the bias will be removed. Our simple averages for observables (eqn. 3.29) become,

$$\langle \mathcal{O} \rangle = \frac{\sum_{i=1}^{N_s} \mathcal{O}_i \exp\left[-\eta_m\right]}{\sum_{i=1}^{N_s} \exp\left[-\eta_m\right]}, \tag{3.36}$$

[7] The question of what the optimal augmented distribution should be is by no means trivial to answer, but the work presented in reference [52] suggests that the gains over using a uniform distribution are small, because "...multi-canonical sampling is never bad in the way that Boltzmann sampling can be bad".

where the i'th observation belongs to the m'th macrostate. This process is sometimes referred to as 'unfolding the weight function'. Attempts to determine the distribution of $\mathcal{M}$ are carried out in a similar vein, by using a set of equations based on eqn. 3.36 (one for each macrostate). As the simulation proceeds, we record the amount of time it spends in each macrostate, thus accumulating a histogram of the states visited. We denote this histogram (obtained with some weight function η_m) as $H(m|\eta_m)$. Clearly, the *biased* distribution is estimated by,

$$P(\mathcal{M}|\eta_m) \stackrel{\text{EB}}{=} \frac{H(m|\eta_m)}{\sum_{m=1}^{N_M} H(m|\eta_m)}, \tag{3.37}$$

and the unbiased distribution can be found as,

$$P(\mathcal{M}) \stackrel{\text{EB}}{=} \frac{H(m|\eta_m)\exp\left[-\eta_m\right]}{\sum_{m=1}^{N_M} H(m|\eta_m)\exp\left[-\eta_m\right]}. \tag{3.38}$$

The explicit forms of the algorithms in both the NVT and NPT ensembles are both trivially modified by this new MCMC term; only the acceptance probabilities are changed. In the NVT ensemble, we find

$$\tilde{P}_{acc}(i \to j|i) = \min\left\{1.0, \exp\left[-\beta(E(\{\vec{r}\}_j) - E(\{\vec{r}\}_i)) + \eta_l - \eta_m)\right]\right\}. \tag{3.39}$$

and in the NPT ensemble,

$$\begin{aligned}\tilde{P}_{acc}(i \to j|i) \quad = \quad & \min\Bigg\{1.0, \exp\Bigg[-\beta(E(\{\vec{r}\}_j) - E(\{\vec{r}\}_i) + (PV_j - PV_i)) \\ & +N\ln\left(\frac{V_j}{V_i}\right) + \eta_l - \eta_m\Bigg]\Bigg\},\end{aligned} \tag{3.40}$$

where m denote the macrostate to which microstate i belongs, and l denotes the macrostate to which j belongs. The symbol $\tilde{P}_{acc}$ has been used to distinguish these acceptance rules from the unbiased forms (P_{acc}). From equations 3.39 and 3.40, it is clear that the changes to the core of the Monte Carlo algorithm are only slight and thus easy to implement.

Although *using* the weight function is simple, the same cannot be said of the task of generating a suitable weight function in the first place. There are a range of different procedures by which we may determine a suitable weight function [24], but it is difficult to known which will be most efficient beforehand. For this reason, determining which is the most effective technique (or combination of techniques) for a given problem is something of a black art. However, it is important to realise that it does not really matter how we generate the weight function. We are free to use any means, fair or foul, to put together a weight function that works. The mechanism for utilising and unfolding the weight function is the only part of the process which can affect the accuracy of the results.

3.6.2 Generating the weight function

The simplest approach, which is suitable for many cases, is the visited states (VS) method. This is an iterative method by which a revised weight function ($\eta^{(n+1)}$) is derived from the previous weight function ($\eta^{(n)}$) combined with a histogram of $\mathcal{M}$ obtained by using those weights. The update scheme is written as,

$$\eta_m^{(n+1)} = \eta_m^{(n)} - \ln\left[H^{(n)}(m|\eta_m) + 1\right] + k, \tag{3.41}$$

where the arbitrary constant k is fixed by the convention that the lowest point on the function will be placed at zero ($\min_m(\eta_m) = 0$). This arbitrary constant arises because the dynamics of the system are determined only by the *change* in the weight function as one moves between microstates, and so the absolute values of η are unimportant.

This iterative process may be initialised by assuming a flat weight function, thus reducing the initial simulation to normal importance sampling. The histogram thus obtained can then be used to modify the weight function (through eqn. 3.41) so that the probabilities of macrostates which were infrequently visited will be enhanced, and the more popular macrostates will find their popularity diminished. Eventually, this iterative process will converge; the measured histogram will become uniform, and the weight function will be unchanged from one iteration to the next. At this point, the weight function precisely describes the distribution $P(\mathcal{M})$, as the bias now exactly opposes this distribution. However, this method tends to be very slow to converge, and a large number of iterations may be required when the twin peaks of $P(\mathcal{M})$ are narrow in comparison with the distance in $\mathcal{M}$ between them. This poor performance can be improved by block-wise evaluation, where a set of separate simulations are used, each of which is constrained to explore different regions of $\mathcal{M}$. The results of these Blocked Visited States (BVS) simulations can be used to estimate different sections of the weight function (via eqn. 3.41), which can then be spliced together to provide an estimate of the overall weight function. Once a reasonable weight function has been determined, the standard VS algorithm can be used to refine it further.

A second blocking technique is based on using a slowly moving barrier, which allows the system to explore only the $\mathcal{M} = 0$ macrostate at first, then the $\mathcal{M} = -1...+1$ states, and then the $\mathcal{M} = -2...+2$ states, and so on. The barrier is moved outwards at regular intervals of the order of the equilibration time. This allows each side of the double-peaked distribution to be evaluated separately, after which the two estimates can be combined to produce an estimate of the overall distribution. This will be referred to as the Mobile Barrier Visited State (MBVS) technique.

The macrostate transition probability matrix

While the VS methods have the advantage of being relatively simple, a significantly more efficient approach has been invented, based on monitoring the rate of transitions between macrostates [24]. During this work, many different algorithms have been employed for the purposes of collecting and utilising the transition data. However, before describing them, we must first outline the formal mathematical structure upon which all of these methods are based.

In a Monte Carlo simulation, the sequence of *microstates* forms a Markov chain (eqn. 3.13). Consequently, the sequence of *macrostates* generated by a Monte Carlo simulation can be also described in those terms [24],

$$\Delta P_m = \sum_{l \neq m} \mathcal{P}_{ml} P_m - \sum_{l \neq m} \mathcal{P}_{lm} P_l, \tag{3.42}$$

where (to distinguish this equation from its microscopic equivalent) $\mathcal{P}_{ml}$ is used as a shorthand for $P(m \to l|m)$. The matrix $\mathcal{P}_{ml}$ is the macrostate transition probability matrix (MTPM). The macrostate Markov chain can be shown to obey a detailed balance condition,

$$\mathcal{P}_{ml} P_m = \mathcal{P}_{lm} P_l. \tag{3.43}$$

It should be noted, however, that this is true only under the assumption of fast local equilibration. As the simulation moves between microstates, it takes some time for the history of the simulation to be forgotten (the autocorrelation time). If this time is significant on the scale of the time it takes to move from macrostate to macrostate then the measured transition rates will depend on the recent history of the system, not just the current state, and the macrostate chain becomes non-Markovian. Fortunately, each macrostate will usually correspond to a large number of microstates, and so the rate of transitions between macrostates tends to be sufficiently slow. Even if the macrostate transitions are found to occur too rapidly, this can be circumvented by 'locking' the simulation into each macrostate for a small period of time.

The actual transition rates ($\mathcal{P}_{ml}$) themselves represent the total probability of all possible transitions between two macrostates. Therefore, to measure the transition matrix, we should count the number of accepted transitions from one macrostate to another, remembering to include the rejected moves as transitions from the macrostate in question unto itself. These quantities are recorded in a matrix, $\mathcal{T}_{ml}$, thus allowing the transition rates to be estimated as,

$$\mathcal{P}_{ml} \overset{\text{EB}}{=} \frac{\mathcal{T}_{ml}}{\sum_k \mathcal{T}_{mk}}. \tag{3.44}$$

In other words, the probability of a transition from m to l ($P(m \to l \mid m)$) is estimated using the number of observed transitions from m to l divided by the total number of transitions observed for state m.

In fact, this estimator can be improved further by recording all possible outcomes of each transition, weighted by the probability that they would have occurred [53]. In this case, every time a microstate move is attempted (with an associated macrostate transition $m \to l$), $\mathcal{T}_{ml}$ is incremented by the (Boltzmann) microstate transition probability (P_{acc}), and $\mathcal{T}_{mm}$ is *also* incremented by the probability of rejecting the move ($1 - P_{acc}$). Note that the accumulated data does *not* use the MCMC acceptance probability ($\mathcal{P}_{acc}$), and so the recorded probability is that which *would* have been observed *if* Boltzmann importance sampling was being used. This scheme makes maximum use of all the available information, and so the transition probability (TP) work in this thesis will usually employ this approach.

Making use of the MTPM

Once the MTPM has been collected, the p.d.f. of the macrostate order-parameter can be determined in a number of ways. The simplest is to use the macrostate detailed balance condition (eqn. 3.43) in an iterative form. One chooses some arbitrary value for P_m, and then generates P_{m+1} using eqn. 3.43 in the form,

$$P_{m+1} = P_m \frac{\mathcal{P}_{m(m+1)}}{\mathcal{P}_{(m+1)m}}. \tag{3.45}$$

This iterative process can be started at any value of $\mathcal{M}$, and used to trace the shape of the distribution outwards from that point. The constraint of normalisation ($\sum_m P_m = 1$) is then applied, which removes the arbitrariness injected by the initial choice of P_m. This 'shooting method' can suffer by accumulating errors along its trajectory, but the results can be checked by using different forms of this approach. For example, the same iterative procedure can be used to run from high to low m, from low to high, and from the mid-value of m outwards (in both directions). Furthermore, it should be noted that eqn. 3.45 only utilises the data from nearest-neighbour transitions, and there is a less wasteful approach which takes all the recorded transitions into account. The distribution we want, P_m, is in fact the steady-state eigenvector of $\mathcal{P}_{ml}$, and standard matrix solution routines can be used to determine this eigenvector, thus making use of all the transition data. Some initial work was done to compare these different schemes, using the three shooting methods mentioned above and the eigenvector method to determine the probability distribution of the magnetisation of a two-dimensional Ising model. The advantages of using the more complex and computationally expensive eigenvalue approach were found to be slight, as were the differences between the three shooting methods, which all produced reasonably accurate results.

Having determined P_m, the weight function can be calculated as,

$$\eta_m^{TP} = \ln\left[P_m + 1\right] + k, \tag{3.46}$$

in much the same way as VS (eqn. 3.41). If necessary, this weight function can then be refined either via the visited states method, or by collecting more transition information. Fortunately, when using the

MTPM approach the data from different runs is easy to combine: the Boltzmann macrostate transition probabilities do not depend on the probability of visiting each state in the simulation (i.e. one records the Boltzmann probability of a transition from i *given* that we are in state i). This means that the information we collect is not dependent on the algorithm used to explore the macrostate space (under the condition of fast local equilibration). Reference [53] includes a proof of this fact in the context of MCMC simulation. It shows that the data from many different runs can be combined by a simple summation of the individual elements, even if different weight functions were used. This allows the MTPM to be steadily built up, while the MCMC weight function is slowly improved.

Artificial dynamics

The TP weight-generation algorithms considered so far have all been based in the 'natural' dynamics of an importance or MCMC sampling simulation. To speed up the evaluation of the transition probabilities, a number of 'artificial' dynamics can be used. For example, we have used the Multiple Initialisation (MITP) method (as used in [24]), where the data from many simulations, each of which has been initialised somewhere in between the two equilibrium peaks, are combined to produce an estimate of the weight function. Another example of this concerns block-wise evolution (BTP), similar to the blocked VS algorithm mentioned above. Under the BTP scheme, a number of simulations can be used, each locked into separate ranges of macrostate space, and the transition matrix elements from each of these can be combined to produce the overall transition matrix. The mobile-barrier MBVS scheme mentioned previously can also be used to control the evolution of the simulation while the *transition rates* are measured (MBTP). As well as these techniques, which have been taken from the literature [24, 49] a related technique, referred to here as "strong sampling", was invented during the course of this work.

Strong sampling

This new (unpublished) TP method has been designed to sample the whole range of $\mathcal{M}$ *automatically* and so remove the 'black art' from weight function determination. The algorithm is based around a four step process intended to *force* the simulation to visit the entire range of macrostates of $\mathcal{M}$ in a near uniform manner. It is not dissimilar to the BTP technique mentioned above, but the 'block' into which the simulation is locked is only one macrostate wide. This is implemented by using two 'barriers', each of which forces any microstate transitions which would cause the simulation to cross the barrier to be rejected. For example, if the order parameter was chosen to be the energy, and if the macrostates were naturally discrete, then this would simply amount to implementing a microcanonical (fixed energy) simulation in the MC context. Although the system is locked in a single macrostate, the simulation is

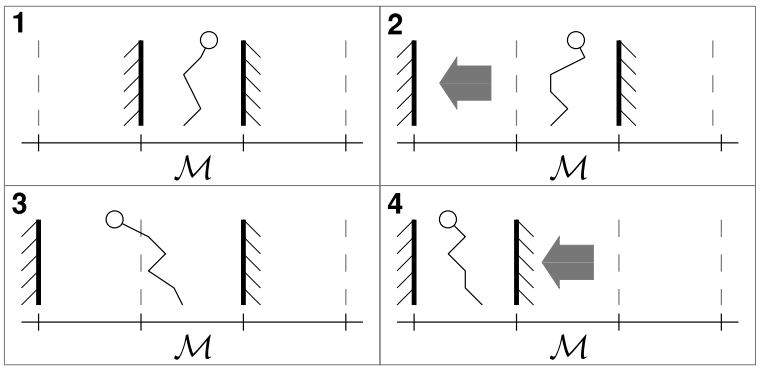


Figure 3.9: The strong-sampling dynamics can be broken down into 4 steps. See main text for details.

still free to evolve by exploring the set of accessible microstates which belong to that macrostate. The simulation proceeds as follows (see fig. 3.9 for an illustration of each step):

1. Lock the simulation into the current microstate, for a period of n_{lock} attempted moves. Two barriers are used to enforce a block which is a single macrostate wide.

2. After n_{lock} moves have been attempted, pick one of the two barriers at random, and widen the block in that direction by 1 macrostate. The block is now two macrostates wide.

3. Wait until the system happens to move from the original macrostate to the new macrostate made accessible by step 2.

4. Once the system has moved to the new macrostate, move the other barrier so that the system is now locked in the new macrostate. Return to step 1 and repeat.

While this process is controlling the simulation's evolution, *all* attempted transitions are recorded in the MTPM with their Boltzmann probabilities. As long as the simulation is able to equilibrate locally (i.e. n_{lock} is sufficiently large) this process should build up an accurate estimate of the entire MTPM. One can view this idea as turning the MCMC process on its head, forcing (or at least strongly encouraging) the system to visit the macrostates uniformly and then *inferring* the weight function, instead of *teaching* it to do so by slowly *evolving* the weight function.

However, it should be noted that this 'strong-sampling' (SS) algorithm does not strictly adhere to detailed balance, as a direct consequence of the presence of the moving barriers. For this reason, the sampling algorithm will only be used as a way of generating an estimate of the weight function, before using normal MCMC to collect the data from which any physical results are to be taken. The SS algorithm will be tested against the other TP methods and the VS approach during the course of this work (see below).

In support of the strong-sampling algorithm, it should be noted that the approach bears more than a passing resemblance to a number of published techniques. In particular, the relationship of this scheme to the broad histogram method (BHM) of Oliveira *et al.* [54, 55] is worth exploring. The BHM employs a series of independent microcanonical simulations to evaluate the energy-macrostate TP matrix, each exploring a single macrostate while monitoring and refusing any transitions to other energies. This parallel exploration of the energy macrostates is essentially the only difference between the BHM and the SS algorithm, which explores the $\mathcal{M}$-space diffusively. There are many variants on the original form of BHM [56, 57, 58, 59], all of which bear some resemblance to the strong-sampling algorithm but with different macrostate dynamics. It should be noted that while the reliability of some of these algorithms is perhaps uncertain [60, 61], there is an ongoing development of ideas (e.g. [62]) which could eventually produce a BHM-style method that is as reliable as (but more efficient than) MCMC.

Methodology of this work

Although the direct transition-matrix evaluation techniques may eventually supersede the multicanonical approach, the case for these techniques is not yet strong enough to allow the present research to rely solely upon them. Therefore, while a number of these techniques have been applied during this work, they were only used to form *initial estimates* of the weight function required as input to the MCMC scheme. This reliable and trustworthy scheme was used for all of our free-energy estimates. The different weight-generation schemes applied during this work (VS, MBVS, TP, MITP, MBTP & SS) are compared in §4.10.5 & §6.5.1.

3.6.3 Limitations

The MCMC approach has been successfully applied to a number of phase-coexistence problems, for example in lattice-gauge theory [63], ferromagnets [51], the Lennard-Jones fluid [64], protein folding [65] and isostructural phase transitions [66]. The fruitful application of this approach to the Lennard-Jones liquid-gas phase-boundary problem is of particular interest here, as this involved the study of a first-order phase transition for one of the systems we are concerned with. The simulation work in [64] was based on choosing the order parameter to be the density, and using MCMC to ensure that the simulation visited both the high-density (liquid) and the low-density (gas) phases. The simulation progressed between these two extremes via a series of two-phase configurations consisting of coexisting liquid and gas regions separated by an interface. Unfortunately this approach has some severe limitations when applied to solid-liquid and structural phase transitions, essentially as a consequence of the choice of order parameter.

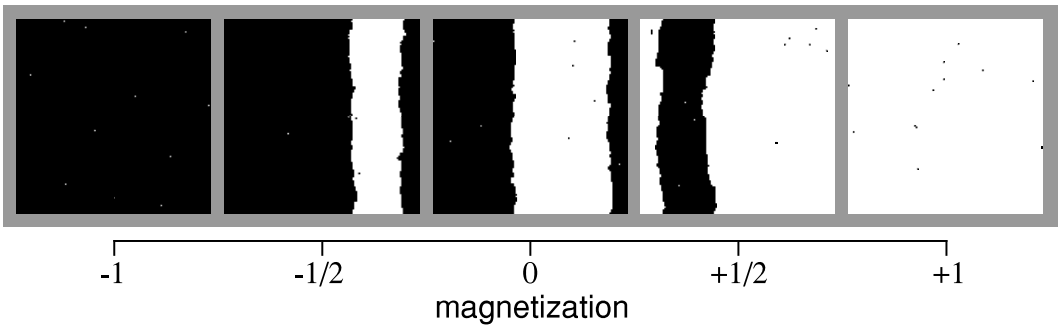


Figure 3.10: Multicanonical Monte Carlo simulation of a 128^2 Ising model at low temperature (coupling constant $J/k_BT = 0.65$). A normal importance sampling simulation at this temperature would become locked in one of the extreme (+1 or -1) microstates. However, multicanonical weighting allows both extremes of magnetisation to be explore, via a two-phase coexistence pathway.

The conventional strategy is to weight multicanonically along the path indexed by the 'conventional order parameter' (e.g. density for the liquid-gas problem, magnetisation for the Ising model). In this case, the macrostates of intermediate m consist of heterogeneous two-phase states with one or more interfaces separating the different phases (see fig. 3.11(a)). For some systems, this scheme can work very well (see fig. 3.10 for an example), but for problems involving crystalline phases this approach tends to suffer from a number of problems. These problems arise because the simulation attempts to traverse a pathway of two-phase coexistence. In the liquid-gas case, the cost of forming and moving the interface is relatively small, but in denser systems the interfaces become very expensive to create. In the case of most solid-solid phase transitions,[8] this approach requires substantial (physically slow) restructuring, and the simulation has a strong tendency to become stuck in some midway non-crystalline state. The central aim of this work is therefore to try and identify some mechanism (and some associated order parameter) that would allow a simulation to move between two distinct solid-state structures *without* passing through any kind of interface-laden mixed-phase states. In other words, the simulation should remain crystalline at all times (see fig. 3.11(b)).

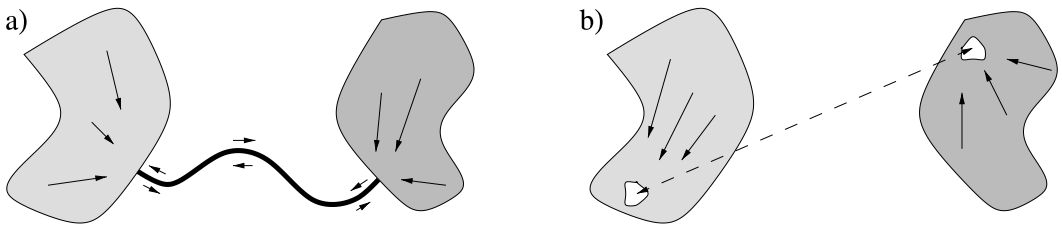


Figure 3.11: Schematic illustration of two different dual-phase sampling paradigms. (a) Different phases can be explored via a two-phase pathway along which one phase is slowly transformed into another. (b) Alternatively, if a suitable transformation can be found, one can switch directly between the configuration space associated with each of two phases. The nature of the inter-phase transformation will determine the size of 'gateway' region (shown in white).

[8] But not all, see for example the *fcc-fcc* isostructural phase transition studied in [66].

3.7 Discussion

The greatest freedom intrinsic to the Monte Carlo approach is the freedom to invent ‘special moves’. Instead of employing only ‘physical’ moves, such as moving single particles around by small amounts, one can invent composite moves, where many degrees of freedom are updated simultaneously. This is the idea behind the cluster moves used to efficiently simulate the Ising model, where the speed of the simulation’s progress through configuration space is boosted by flipping many spins at once [17, Chapter 12].

The aim of the work in this thesis is to design a special move that will map a configuration of one crystalline structure onto some configuration of another. Large moves in configuration space are generally found to be somewhat unlikely to occur, and so we will probably have to apply extended sampling techniques in order to encourage this phase-switch to occur. The work presented here explores the development of just such a technique, making use of a global coordinate transformation to perform the switch, and defining a suitable order parameter so that this move may be multicanonically encouraged. This interface-avoiding dual-phase simulation algorithm was first developed for the hard-sphere solid, and this forms the subject matter of the next chapter.

Chapter 4

The Hard-Sphere Solid

The structural phase behaviour of the hard-sphere solid is a long-standing problem in statistical physics which has only recently been dealt with satisfactorily (see §4.1 below). This fact, together with the model's simplicity and broad range of interest, make the hard-sphere solid an excellent testing ground for our new simulation technique. In this chapter, the ideas presented in §3.7 will be developed into an algorithm capable of reliably determining the structural phase behaviour of the hard-sphere system. This will show that while the original *implementation* of the lattice-switch method [1] was slightly flawed (leading to erroneous results) the *ideas* behind that research were sound. The revised simulation technique (as published in [2]) will be used to explore a range of issues concerning the crystalline phase behaviour of hard-spheres.

4.1 Background

As mentioned in §2.1.1, the equilibrium phase behaviour of a system of hard-spheres is controlled by the density. We express the density of a system of N hard-spheres (in a volume V) in terms of the reduced number density,

$$\tilde{\rho} \equiv \frac{1}{\tilde{v}} \equiv \frac{\rho}{\rho_{cp}} \equiv \frac{N/V}{\sqrt{2}/\sigma^3} \ , \tag{4.1}$$

where ρ_{cp} is the number density at close-packing, σ is the diameter of the spheres, and $\tilde{v}$ is the volume per particle. For the NPT ensemble, the pressure is defined in reduced units such that,

$$\tilde{\mathrm{p}} \equiv \frac{\mathrm{p}\sigma^3}{kT} \ . \tag{4.2}$$

A factor of kT has been folded into the pressure and so the temperature only affects the system indirectly, in that the value of the reduced pressure is dependent upon it.

The different phase regimes are shown in figure 4.1, both schematically and through the form of the pressure-density curve of the system. At low densities ($\tilde{\rho} \lesssim 0.663$), the hard-sphere fluid is the stable phase. As the density is increased above $\tilde{\rho} \approx 0.663$, the equilibrium state consists of coexisting fluid and crystal regions. Beyond $\tilde{\rho} \approx 0.733$, the equilibrium structure is a pure crystalline solid. The dashed region of the solid-phase curve shows the behaviour of the metastable crystal, terminating in a spinodal point below which the solid will always spontaneously melt. Also, if the system is quickly 'quenched' by rapidly increasing the density (or the pressure), then the metastable (glassy) fluid phase may be observed. This metastable branch is believed to terminate at $\tilde{\rho} \approx 0.86$, at which point the pressure will diverge as the structure becomes randomly close-packed (*rcp*) [67, 68]. For the crystalline phase, the pressure increases with the density and finally diverges in the crystalline close-packing limit (by definition, at $\tilde{\rho} = 1.0$).

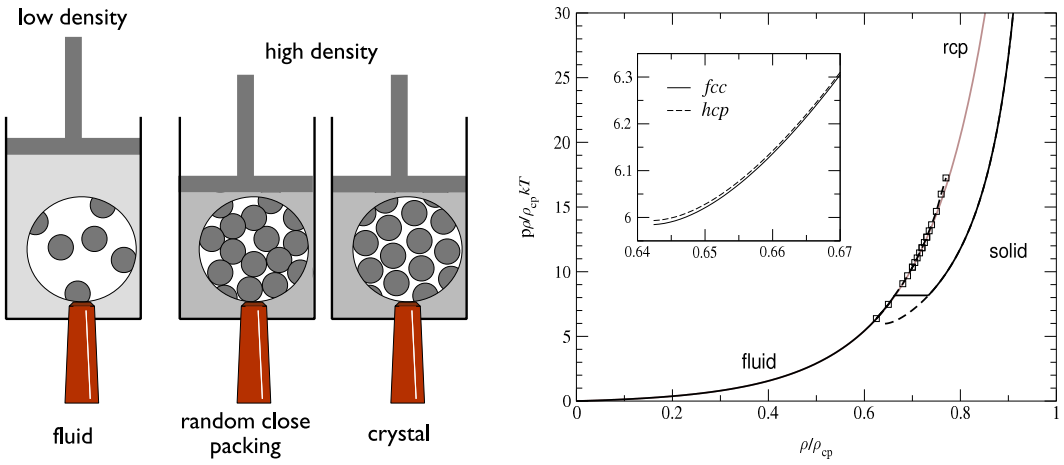


Figure 4.1: Illustrations of phase behaviour and the phase diagram of the hard-sphere system. At low densities, a fluid is observed for which the semi-empirical Carnahan-Starling equation [69] provides an accurate equation of state. At higher densities, the fluid branch (leading to random close-packing) is metastable with respect to the coexisting fluid-crystal or the ordered crystal phase. The fluid-crystal tie-line and the random close-packed data/equations are taken from [67]. The crystalline solids' equations of state were taken from [70]. Inset shows a close-up of the spinodal region of the crystalline phase, where the difference between the *fcc* and *hcp* equations of state (taken from [70]) can be clearly seen.

As explained in §2.1.1, the candidates for the equilibrium structure of the hard-sphere solid are the close-packed crystals and of these structures the face-centred cubic and hexagonal close-packed are the most important (the other possible close-packed structures will be introduced in §4.2). In the early (1960's) work on this problem, the only measurable difference between the *fcc* and *hcp* structures was the pressure difference [71, 7] (as shown by the inset graph on the right-hand side of fig. 4.1). However, the pressure difference on its own cannot be used to determine which is the thermodynamically favoured phase; this can only be achieved by explicitly evaluating the total free-energy difference. Of the publications during this period, only reference [71] contains a calculation of the entropy difference: they estimated that

$s_{fcc} - s_{hcp} \equiv \Delta s = 0.002Nk$ (i.e. favouring *fcc*) at close-packing ($\tilde{\rho} = 1$) via a simple integration-method calculation (under the assumption that $\Delta s = 0$ at melting and that the pressure difference was constant over all densities). However, the authors could provide no estimate of the errors associated with their value of Δs.

Since that time, there have been a number of attempts to clarify this issue, both analytically (e.g. [72, 73]) and numerically (e.g. [74, 75]). However, these publications were isolated, inconsistent and uncorroborated. Even in the mid-1980's, almost 20 years after reference [71] appeared, the careful Einstein-crystal integration of Frenkel and Ladd [33] was still inconclusive, yielding bounds upon the entropy difference such that $-0.002Nk < \Delta s < 0.001Nk$ (near melting, at $\tilde{\rho} = 0.736$). This issue remained unsolved and largely ignored until the 1997 publication by Woodcock [30]. However, this publication produced a value of Δs of the order of $0.005Nk$ for all densities, which is incompatible with results of Frenkel and Ladd [33]. This disagreement initiated a flurry of publications (which will be covered in more detail in §4.10.7), including two publications involving the work presented here [1, 2].

4.2 Meet the structures

Consider the creation of a close-packed structure, by taking a number of close-packed (hexagonal) layers and placing them one on top of another. There are three possible layer positions, which (using the standard notation) are referred to as A, B and C. Face-centred cubic corresponds to stacking patterns with a repeating unit of three, written as ABC, and hexagonal close-packed has a repeating unit of 2, denoted as AB. These two stacking patterns are shown in figure 4.2. The local (first nearest neighbour) environment that each particle 'sees' is shown in figure 4.3.

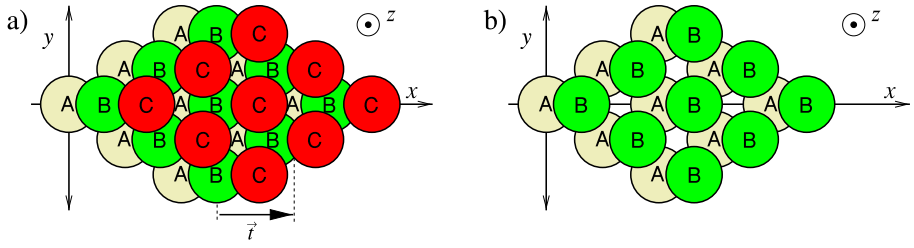


Figure 4.2: Schematic illustration of the *fcc* (ABCABC...) and *hcp* (ABAB...) stacking patterns. The hexagonal planes are parallel to the x-y plane, and are stacked in the z-direction. The layer can be identified as being an A, B or C layer by its displacement in the x-direction; for A the displacement is zero, for B it is $+{}^1/_2\vec{t}$ and for C it is $+\vec{t}$. The coordinate-system chosen for our calculations will be explained in more detail in §4.8.1.

For *hcp*, the spacing between the stacking planes (c) and the separation between particles within each plane (a) can change. For measurement purposes, we define the c/a ratio for both *hcp* and *fcc* structures

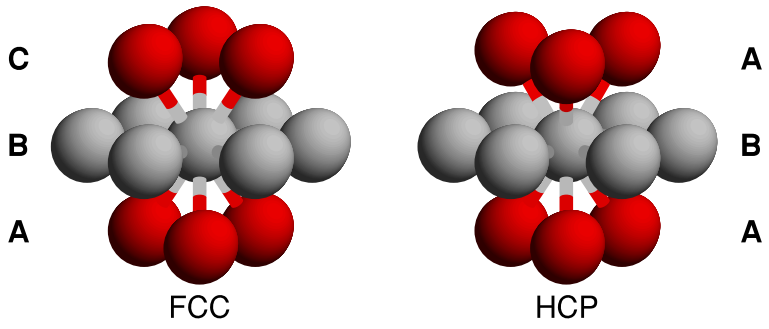


Figure 4.3: Three-dimensional plot of the positions of the first-nearest neighbours around every lattice site for the *fcc* and *hcp* structures. The particle environments are very similar, the only difference being the orientation of the triplets above and below each particle (in the z-direction). For *fcc*, the triplets point in opposite directions, and for *hcp* they both point in the same direction.

by $c/a = 2d_0/\overline{S}$, where d_0 is the mean separation between close-packed planes and $\overline{S}$ is the mean separation between neighbouring sphere-centres within the close packed planes. In the *hcp* structure $c/a = \sqrt{\frac{8}{3}} = 1.63299$ if the packing is *ideal* [9]. The cubic symmetry of the *fcc* structure, combined with the radial symmetry of the hard-sphere interaction, mean that there can be no systematic deformation of the *fcc* structure along some preferred direction. The symmetries of the system simply deny the possibility that any direction may be preferred, and c/a *must* be ideal. This fact can therefore be used as a check on the implementation of our constant-pressure calculations (§4.10.3).

As well as *fcc* and *hcp*, there are any number of stacking patterns which all have the same density, and so all of these possibilities should be considered. Examples of this type of structure are twinned *fcc*, where *fcc* regions are separated by a single *hcp*-like layer (e.g. ABCABCBACBA), and random hexagonal close-packing (*rhcp*), where the stacking pattern is stochastic. However, these structures can always be broken down into a sequence of *fcc* and *hcp* domains, as illustrated in figure 4.4. Our primary objective, therefore, should be to compare the *fcc* and *hcp* structures themselves. The role of the mixed structures will be considered further in §4.10.9.

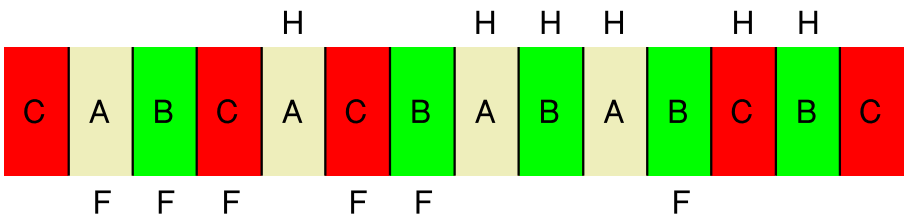


Figure 4.4: Illustration of a random hexagonal close-packed structure, where the stacking planes (lain vertically, from left to right) are arranged randomly. The resulting structure can be thought of as a mixture of *hcp* and *fcc* domains (denoted as H and F respectively) by identifying those planes for which the neighbouring planes are different (e.g. ABC) or the same (e.g. ABA).

4.3 Lattice-switch Monte Carlo for hard spheres

Before we can consider how to switch *between* two crystalline structures, we must be able to *identify* a configuration as 'belonging to' a given crystalline phase. This can be done quite naturally, and in the traditions of lattice dynamics (see §2.4.1 and §6.4), by decomposing the particle position coordinates into a sum of 'lattice' and 'displacement' vectors (see fig. 4.5),

$$\vec{r}_i = \vec{R}_i^{\alpha} + \vec{u}_i \ . \tag{4.3}$$

Here, $\{\vec{R}\}_\alpha \equiv \{\vec{R}_i^{\alpha}, i = 1...N\}$ is a set of fixed (configuration-independent) vectors describing the crystalline structure (labelled α). The set of positions of all N sites will be referred to as the 'lattice', although this is not strictly the usual crystallographic meaning of the word.[1] The $\{\vec{u}\}$ vectors represent displacements with respect to those lattice sites.

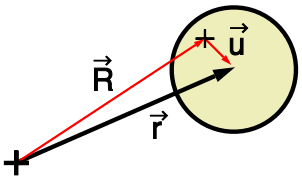


Figure 4.5: The decomposition of particle positions $\vec{r}$ into a lattice-site position $\vec{R}$ (marked with a cross) and a displacement $\vec{u}$ from that site.

This framework provides a number of ways of identifying a configuration as being associated with a particular structure. One might adopt the criterion that all particle displacements, with respect to the associated lattice sites, should lie within some *spatial* cut-off, chosen to be sufficiently large that the results are independent of its specific value. The bounded random-walk and top-hat move generation schemes (§3.2.1) are implementations of this constraint. Alternatively, we might identify the relevant configurations as the set that are accessible from *some* member of the set (e.g. the perfect crystalline state) within some given *temporal* cut-off. The actual length of the temporal cut-off should not be important as long as it is 'long enough', meaning that it allows the configurations associated with a particular structure to be sampled efficiently. This has the merit of being what, in practice, computer simulations actually do: if a computer simulation is begun in *fcc*, say, then it will tend to stay in that structure over any and all reasonable simulation timescales. This style of cut-off is invoked implicitly by choosing to implement the (unbounded) random-walk move generation mechanism (§3.2.1). Of course, if a simulation is able to change structure spontaneously, then the lattice-switch approach will

[1] A crystalline structure is usually described in terms of the *crystallographic lattice* (defining the periodic unit) and the *basis* (which defines the positions of the lattice sites within the repeating cell). By convolving the lattice with the basis, the positions of all the sites in the system are determined, and it is this that we refer to as the 'lattice'.

fail because the algorithm will not be 'aware' that this transition has occurred, and will mis-classify the explored microstates. If necessary, a spatial cut-off can be used to restrict the sampling to a particular structure.

Having associated a set of lattice sites (labelled α) with a set of microstate configurations, the nature of the 'lattice-switch' (LS) move can be made clear. The structure label is now treated as another stochastic variable, so that as well as updating, for example, the particle positions ($\vec{u}_i \rightarrow \vec{u}_i'$), we are also free to attempt a structure-update ($\alpha \rightarrow \alpha' \equiv \{\vec{R}\}_\alpha \rightarrow \{\vec{R}\}_{\alpha'}$). However, there are still decisions to be made concerning the precise nature of the LS transformation.

4.3.1 The choice of mapping

In principle, there are many transformations which will map the displacements from one set of lattice sites onto the other. The simplest type of mapping swaps the lattice sites while the displacement vectors ($\{\vec{u}\}$) remain unchanged, and most of the transformations considered here will be of this type. More generally, there are $N!$ possible site-site mappings, corresponding to all the possible ways in which every site of phase α can be mapped onto each of the N sites of the conjugate phase α'. In fact, the only constraint is that the mapping preserves the volume of information held by the displacement vectors, thus ensuring that the move is reversible and in turn that detailed balance is satisfied. We are free to write the position decomposition as

$$\vec{r} = \vec{R}^\alpha + \overline{\overline{T}}^\alpha . \vec{u} \, , \tag{4.4}$$

where $\vec{r}$, $\vec{R}^\alpha$, and $\vec{u}$ are now column vectors with 3N elements and $\overline{\overline{T}}^\alpha$ represents two $3N \times 3N$ non-singular matrices, one for each structure (c.f. eqn. 4.3). The use of such a transformation matrix allows many different styles of mapping, which will be classified as follows. Firstly, there are the direct site-site mappings (SSM) mentioned above. Secondly, a simple site-site mapping can be modified such that the displacement vectors associated with each site are also transformed (TSSM). For example, given that the local neighbour environments of *fcc* and *hcp* differ only in the orientation of the trio of particles above or below each site (fig. 4.3), we might imagine designing a mapping where the displacement vectors are reflected or rotated to take this into account.

Also, it is possible to create 'many-to-many' mappings, where the displacement of a particle in one phase corresponds to the displacements of many particles in the other. This will be referred to as a non-local mapping (NLM). Note, however, that this means using a non-unitary $\overline{\overline{T}}$-matrix in eqn. 4.4, and this will destroy the underlying symmetry of the Markov chain. To restore this symmetry, we require the form of the configurational energy to be modified such that

$$E'(\{\vec{r}\}) = E(\{\vec{r}\}) - \ln \left[\det \overline{\overline{T}}^\alpha \right] \, . \tag{4.5}$$

This representation is significantly more complex than the simple local mappings, and introduces a number of computational overheads. Also, the problem of identifying the most efficient NLM is very difficult to tackle (it is, effectively, a $3N$-dimensional minimisation problem), and the work presented in §4.10.4 suggests that (for hard-spheres at least) there is little to gain here by this kind of tuning. For these reasons, all the mappings used in this thesis will be one-to-one (SSM) mappings. The possible advantages of using a complex mapping will be considered in §7.2.

While a number of different site-site mapping have been tested (see §4.10.4), one particular mapping was used during almost all of this work. As illustrated schematically in fig. 4.6, this mapping exploits the fact that *fcc* and *hcp* differ only by their stacking pattern (ABC versus AB), by simply *translating* the appropriate close-packed planes. By 'translate' we mean, more precisely, 'relocate at a position defined by an appropriate translation vector'; the planes should not be thought of as 'sliding through' the intermediate space.

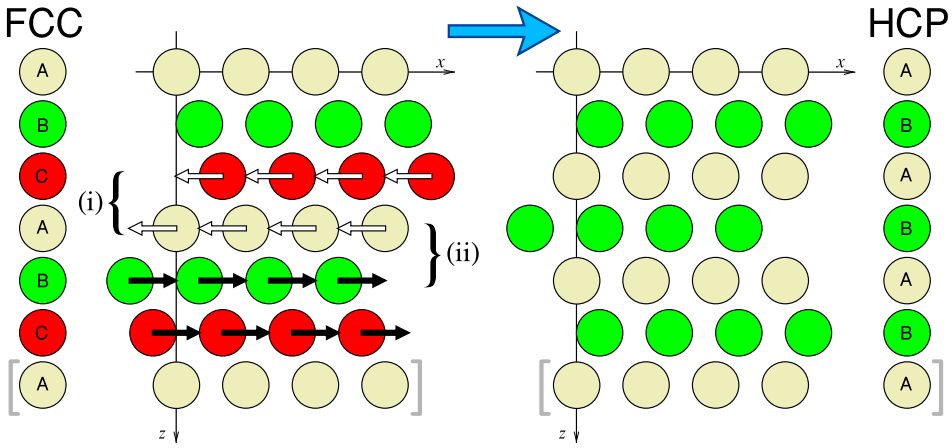


Figure 4.6: The main LS transformation, shown for the perfect-crystal configuration. Six close-packed (x-y) layers are shown, with an additional [bracketed] layer at the bottom showing the position of the periodic image of the top ($z = 0$) layer. The circles show the boundaries of hard spheres located at the sites of the two close-packed structures. In this realization of the *fcc*→*hcp* lattice switch, the top pair of planes are left unaltered, while the other pairs of planes are relocated by translations, specified by the vectors $-\vec{t}$ (white arrows) and $\vec{t}$ (black arrows). The vector $\vec{t}$ is identified in fig. 4.2.

Fig. 4.6 shows the lattice-switch as applied to the *perfect*-crystal configurations, where all particles are on their lattice-sites ($\vec{u}_i = \vec{0} \ \forall \ i$). In this case, the two structures are clearly 'energy matched' because the lattice switch cannot cause any particles to overlap, and thus will be accepted. However, for a 'typical' equilibrium microstate (see fig. 4.7), the two configurations related by the LS operation will not be energy matched. While the planes which are displaced together (type (i) in fig. 4.6) cannot produce any overlaps between themselves during the switch, the planes which are translated differently (type (ii) of fig. 4.6) may, and indeed with overwhelming probability *will*, map a realizable (no overlaps) configuration of one structure onto an unrealizable (overlapping) configuration of the other. A MC

lattice-switch move will, therefore, be rejected for *most* configurations, but *not* for *all*. A number of configurations for which the lattice-switch move may be accepted (i.e. 'gateway' states) *must* exist. For example, from the definition of our lattice-switch operation, configurations for which all of the particles are 'close enough' to the perfect-lattice positions must fall into this category. One could choose these 'small-displacement' configurations to act as the gateway states, and design a multicanonical procedure accordingly. However, we can avoid having to make this explicit choice and, instead, let the system *find* gateway configurations itself. To do this, we first define a measure of the mismatch between the energies of the configurations linked by the lattice-switch.

4.3.2 The overlap order parameter

For a system of hard-spheres, the configurational energy cost of the lattice switch is always either zero or infinity, and using this to measure the 'energy mismatch' will not produce an effective MCMC weighting procedure. The algorithm must be able to tell if the configurations are 'getting closer' to the gateway states, instead of simply whether the LS produces any overlaps or not. For this reason, the mismatch is best quantified by the number of pairs of overlapping spheres created by the lattice switch. To this end, we let $M(\{\vec{u}\},\alpha)$ denote the number of overlapping pairs associated with the displacements $\{\vec{u}\}$ within the structure α. From this, we define the *overlap order parameter*,

$$\mathcal{M}(\{\vec{u}\}) = M(\{\vec{u}\},hcp) - M(\{\vec{u}\},fcc)\ . \tag{4.6}$$

As $M(\{\vec{u}\},\alpha)$ is necessarily non-negative, and must be zero for any realizable set of displacements for the structure α, then $\mathcal{M}$ will be ≥ 0 (≤ 0) for realizable configurations of the *fcc* (*hcp*) structure. Figure 4.7 provides a concrete example. The gateway states may then be identified (*without* prejudging their microscopic character) as the set of configurations for which $\mathcal{M} = 0$, as these configurations are realizable in *both* structures.

4.3.3 Multicanonical weighting

As mentioned above, the lattice switch will usually cause overlaps for a 'typical' configuration of a given phase. However, we can use multicanonical sampling (§3.6.1) to enhance the probability of visiting the unlikely $\mathcal{M} = 0$ gateway states. The MCMC algorithm supplied in §3.6.1 can be applied to this problem by simply identifying the multicanonical order parameter (the $\mathcal{M}$ of §3.6.1) as the overlap order parameter ($\mathcal{M}$ of eqn. 4.6). This order parameter is naturally discretized (the number of pairs of overlapping spheres can only be an integer), and the MCMC tools presented in §3.6.1 can be used 'straight out of the box'. The combination of the LS transformation, the overlap order parameter and

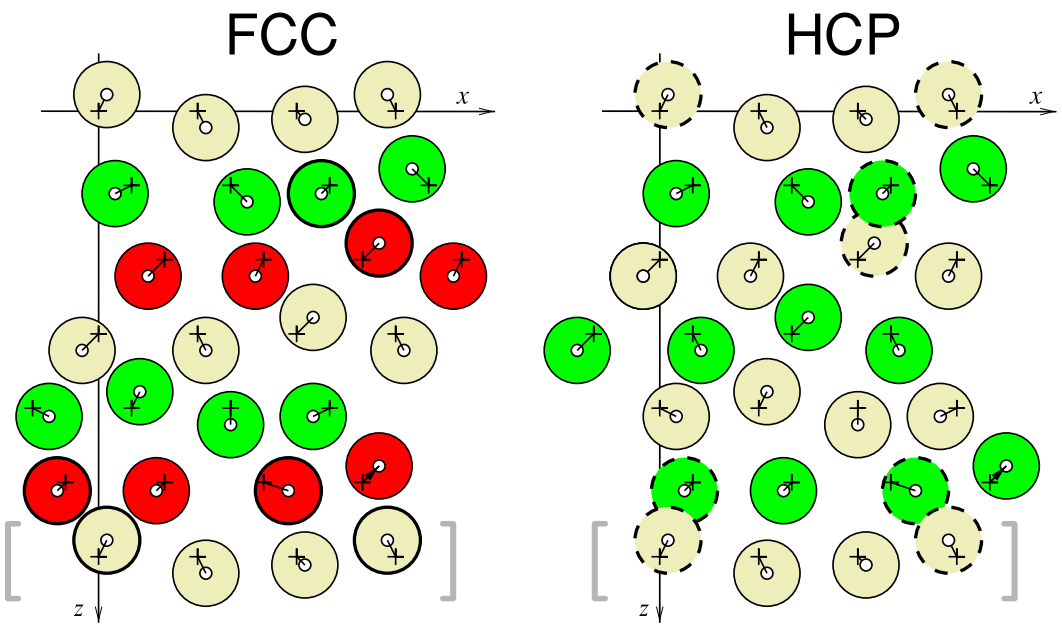


Figure 4.7: Schematic representation of the LS transformation applied to a 'typical' configuration of hard spheres. The crosses identify the lattice sites, while the small circles locate the sphere centres for this configuration of displacements $\{\vec{u}\}$. This configuration is realizable (gives no overlaps) in the *fcc* structure, but under the LS transformation it is mapped onto an (unrealizable) *hcp* configuration with three overlapping pairs of hard spheres (shown with broad, dashed boundaries). Thus, for this configuration, the overlap order parameter $\mathcal{M}(\{\vec{u}\}) = 3$ (eqn. 4.6).

the MCMC algorithm can now be used to evaluate the distribution $P(\mathcal{M}|\mathcal{C})$ for any set of conditions $\mathcal{C}$ (referring to, for example, the constant volume or constant pressure ensembles), via eqn. 3.38. By breaking the distribution into its two (*fcc* and *hcp*) parts, the relative probabilities of the two phases can be estimated (c.f. eqns. 2.11 & 3.36) using,

$$\frac{P(fcc|\mathcal{C})}{P(hcp|\mathcal{C})} \overset{\text{EB}}{=} \frac{\sum_{\mathcal{M}>0} H(\mathcal{M}|\mathcal{C},\{\eta\}) \exp\left[-\eta_m\right]}{\sum_{\mathcal{M}<0} H(\mathcal{M}|\mathcal{C},\{\eta\}) \exp\left[-\eta_m\right]} \ . \tag{4.7}$$

4.4 The constant-volume ensemble

The weight of any configuration in the NVT ensemble is simply $\exp\left[-\beta E\right]$ (eqn. 2.5). The energy E is zero for any realizable configuration of hard-sphere, and so each of these configurations has unit weight. Configurations with overlapping spheres are prevented by the infinite energy barrier, i.e. the configurational weight of this energy is zero. As the weight can only be zero or one, the value of the temperature does not affect the partition function. We have chosen to set $T = 1$, implying that the Helmholtz free-energy is equal in magnitude (although opposite in sign) to the entropy,

$$f(N,V,\beta,\alpha) \equiv -s(N,V,\alpha) = -\frac{1}{N} \ln \mathcal{Z}(N,V,\beta,\alpha) \ , \tag{4.8}$$

where both f and s are expressed in units of the Boltzmann constant (k). The entropy difference can be expressed as,

$$\Delta s_{fcc,hcp} \equiv s(N,V,fcc) - s(N,V,hcp) = \frac{1}{N}\ln \mathcal{R}_{fcc,hcp}(N,V) \ , \tag{4.9}$$

where

$$\mathcal{R}_{fcc,hcp}(N,V) \equiv \frac{\mathcal{Z}(N,V,fcc)}{\mathcal{Z}(N,V,hcp)} = \frac{P(fcc|N,V)}{P(hcp|N,V)} \ . \tag{4.10}$$

As a consistency check, the method of Eppenga and Frenkel [43] will be used to measure the pressure in this ensemble. This should agree with the pressures and densities observed in the constant pressure ensemble (at least in the thermodynamic limit), and can be checked against the results of ref. [70]. The pressure calculation in [43] proceeds as follows: At any time, there is some pair of values of i and j for which $d \equiv (|\vec{u}_j - \vec{u}_i| - \sigma)$ is a minimum (the 'closest-approach', d_s). The pressure can then be determined from,

$$\frac{\tilde{\mathrm{p}}}{\tilde{\rho}} = 1 + \frac{\sigma}{3N\langle d_s \rangle} \ , \tag{4.11}$$

where the average value of the closest-approach ($\langle d_s \rangle$) is determined by block analysis of a sequence of recorded values of d_s.

4.5 The constant-pressure ensemble

In this case, we require the ratio $\mathcal{R}_{fcc,hcp}(N,\tilde{\mathrm{p}})$ of the partition functions associated with each phase

$$\mathcal{Z}(N,\tilde{\mathrm{p}},\alpha) = \int \mathrm{d}\tilde{v}\mathcal{Z}(N,V,\beta,\alpha)\exp\left[-\tilde{\mathrm{p}}N\tilde{v}\right] \ , \tag{4.12}$$

where $\tilde{\mathrm{p}}$ is the dimensionless pressure, $\tilde{\mathrm{p}} = \mathrm{p}\sigma^3/kT$, and $\tilde{v}$ is the dimensionless intensive volume $\tilde{v} = 1/\tilde{\rho}$. We wish to estimate the Gibbs free-energy,

$$g(N,\tilde{\mathrm{p}},\alpha) \equiv -\frac{1}{N}\ln Z(N,\tilde{\mathrm{p}},\alpha) \ , \tag{4.13}$$

such that, in analogy with eqn. 4.9,

$$\Delta g_{fcc,hcp} \equiv g(N,\tilde{\mathrm{p}},fcc) - g(N,\tilde{\mathrm{p}},hcp) = -\frac{1}{N}\ln \mathcal{R}_{fcc,hcp}(N,\tilde{\mathrm{p}}) \ , \tag{4.14}$$

where

$$\mathcal{R}_{fcc,hcp}(N,\tilde{\mathrm{p}}) \equiv \frac{\mathcal{Z}(N,\tilde{\mathrm{p}},fcc)}{\mathcal{Z}(N,\tilde{\mathrm{p}},hcp)} = \frac{P(fcc|N,\tilde{\mathrm{p}})}{P(hcp|N,\tilde{\mathrm{p}})} \ . \tag{4.15}$$

As a further consistency check, we can use the following (thermodynamic) relationship between the Gibbs free-energy difference and the entropy difference [76]:

$$\Delta g(\tilde{\mathrm{p}}(\tilde{\rho})) = \Delta f(\tilde{\rho}) + \mathcal{O}(\Delta\tilde{\mathrm{p}}^2) \ . \tag{4.16}$$

In other words, we expect the Gibbs and Helmholtz free-energies to differ by terms that are *second order* in the pressure difference between the two phases (at least in the thermodynamic limit). That pressure difference is extremely small, as is the difference between the measured densities of the two structures (see fig. 4.1 and ref. [70]). Therefore, we expect the magnitude of Δg to be close to that of the entropy density Δs.

4.6 The close-packed limit

As well as simulating at constant density ($\tilde{\rho} < 1.0$) or at finite pressure, it is also possible to simulate the close-packed limit ($\tilde{\rho} \to 1.0$, $\tilde{\mathrm{p}} \to \infty$) directly. A full treatment of this limit can be found in reference [72], and only an outline of the crucial elements is supplied here. In this limit, the configurational integral (eqn. 2.8) may be rewritten as the product of two terms:

$$\mathcal{Z}(N,V,\alpha) = \mathcal{Z}_0(N,V)\mathcal{Z}_{cp}(N,\alpha)\ . \tag{4.17}$$

The first term is defined by

$$\mathcal{Z}_0(N,V) = \left[\frac{\sigma\epsilon}{1-\epsilon}\right]^{3N}\ , \tag{4.18}$$

with

$$\epsilon = 1 - \sqrt[3]{\tilde{\rho}}\ . \tag{4.19}$$

The contribution to the entropy due to this term diverges in the close-packed limit, but is independent of the phase. In this limit, the sphere separation becomes infinitely small (embodied in ϵ), the sphere surfaces are essentially flat (fig. 4.8(a)), and the system can be visualised as a set of hard dodecahedra[2] (fig. 4.8(b)).

Clearly, then, in this limit only the parallel separation between neighbouring particles can be of importance, because any finite perpendicular displacement will be insignificant on the scale of the particle size. The *difference* between the displacements of every pair of particles from their associated lattice sites ($\vec{u}_{ij} = \vec{u}_j - \vec{u}_i$) is re-expressed in terms of the parallel and perpendicular components of that vector with respect to the direction of each nearest-neighbour site (fig. 4.8(a)),

$$\vec{u}_{ij} = u_{ij}^{\parallel}\hat{n}_{ij}^{\alpha} + \vec{u}_{ij}^{\perp}\ . \tag{4.20}$$

The unit vector $\hat{n}_{ij}^{\alpha}$ describes the direction from a lattice-site i to a nearest-neighbour site j, and therefore the system is dependent on the phase (α) via the geometry of the nearest-neighbour environment. The

[2] Which should be thought of as large in comparison with their separation.

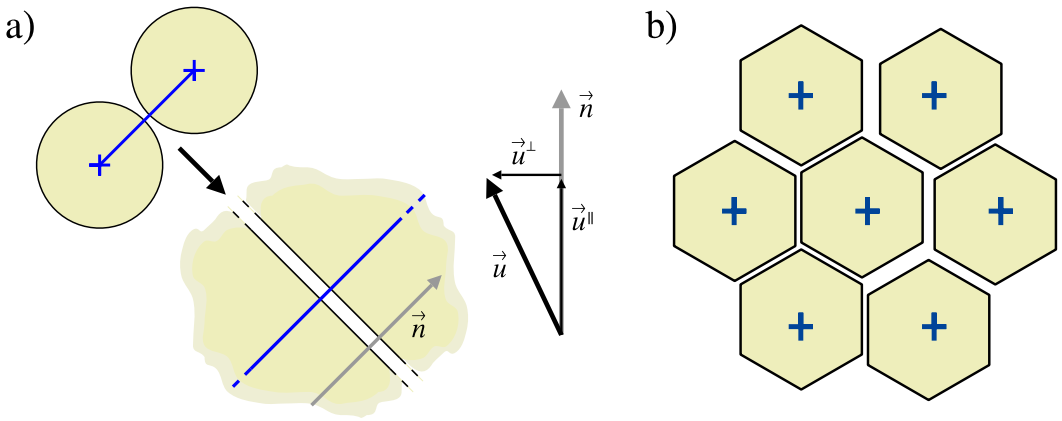


Figure 4.8: The close-packed limit. (a) As the mean separation of the spheres tends to zero, their curvature becomes effectively negligible and only the separation parallel to the nearest-neighbour vectors is important. (b) This limit can be thought of as a system of hard dodecahedra, or in two dimensions, a system of hard hexagons.

displacement coordinates are then rescaled by an ϵ-dependent factor, producing the following expression for the second contribution to the configurational entropy (eqn. 4.17),

$$\mathcal{Z}_{cp}(N,\alpha) = \prod_i \left[\int \mathrm{d}\vec{u}_i\right] \prod_{\langle i,j\rangle}^{nn} \Phi_{cp}(u_{ij}^{\parallel}) \left[1 + \mathcal{O}(\epsilon)\right] \ , \tag{4.21}$$

where the particle interaction is now described by

$$\Phi_{cp}(u_{ij}^{\parallel}) = 0 \qquad \text{if} \quad u_{ij}^{\parallel} < 1, \tag{4.22}$$

$$= 1 \qquad \text{otherwise.} \tag{4.23}$$

Note that all length-scales have now been removed from the problem (the diameter and the density no longer feature in the calculation). However, the actual simulation algorithm remains essentially the same as in the hard-sphere case, but the particle interactions are now made to obey eqn. 4.22; an 'overlapping' pair of particles is identified as those for which $u_{ij}^{\parallel} < 1$.

4.7 Polydispersity

Simulations were also performed for a simple model of a polydisperse hard-sphere system, where each particle now has a unique diameter. The diameter of each sphere, σ_i, is drawn randomly from a uniform distribution over the range $\sigma(1 - \frac{\Delta\sigma}{2}) < \sigma_i < \sigma(1 + \frac{\Delta\sigma}{2})$. The algorithm is altered so as to implement the modified interaction potential:

$$\phi(r_{ij}) = \infty \qquad \text{if} \quad r_{ij} < \left(\frac{\sigma_i + \sigma_j}{2}\right), \tag{4.24}$$

$$= 0 \qquad \text{otherwise.} \tag{4.25}$$

The degree of polydispersity can be set to any value up to the maximum allowed at the prescribed density,

$$\Delta\sigma_{max} = \frac{1-\sigma}{\sigma}. \tag{4.26}$$

This maximum is imposed because if the value of $\Delta\sigma$ were set any higher, the set of diameters could violate the hard sphere constraint when the spheres are placed on the perfect-lattice sites. Note that this is an artifact of the simplistic way in which this simulation was initialised.

A number of polydisperse simulations were performed, each with a separate set of diameters, $\{\sigma\}$, and the free-energy difference was determined as an average over this series. However, it should be made clear that this is only an approximate model of a polydisperse system, as it does not implement a full integration over the ensemble of all possible radii for every particle. It is likely that the true equilibrium structure of a polydisperse hard-sphere solid should show correlations between the radii of neighbouring particles. The approximate calculation described above is not capable of examining this 'size ordering', and so can only be expected to be accurate in the limit of small polydispersity. The reason for including this model was simply to show that it is possible to extend lattice-switch Monte Carlo to deal with such systems.

4.8 Implementation details

4.8.1 Lattice construction and the simulation cell

The *hcp* lattice was constructed as a set of close-packed planes stacked in the z-direction. The separation of nearest-neighbour lattice sites (r_{nn}) was set to unity. This requires that within the hexagonal x-y planes, neighbouring particles are separated by $s_y = 1.0$ in the y-direction and by $s_x = \sqrt{3}/2$ in the x-direction (fig. 4.9(a)). The separation of the stacked planes was $s_z = \sqrt{2/3}$ in the z-direction (see fig. 4.9(b)). The total size of the simulation cell was determined by the number of particles in each direction, such that $L_x = n_x s_x$, $L_y = n_y s_x$ and $L_z = n_z s_z$. The 'stacking identity' of a hexagonal layer (i.e. A or B) was defined by its position in the x-direction (0 and $+t$ respectively). For this arrangement, $\vec{t}$ has a magnitude of $1/\sqrt{3}$.

The *fcc* lattice was derived from the *hcp* lattice via the LS transformation, as in fig. 4.6. For every three AB pairs of *hcp* planes, the first was left unchanged (AB→AB), the second was shifted by $-\vec{t}$ (AB→CA), and the third was shifted by $+\vec{t}$ (AB→BC), thus producing the required ABCABC stacking pattern.

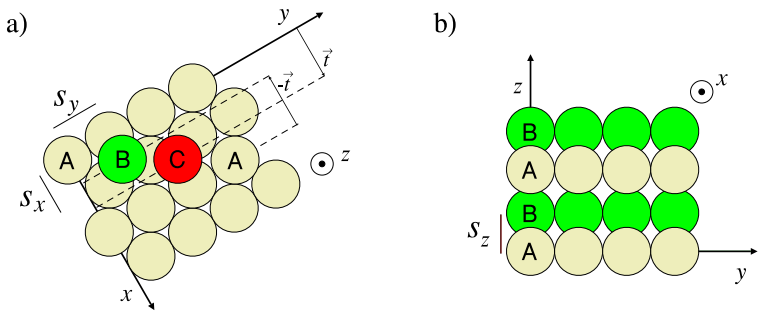


Figure 4.9: Geometry of the lattice construction, showing the close-packed layers lying in the x-y–plane, and stacked in the z–direction. Spheres are shown with unit radius, corresponding to the close-packed ($\tilde{\rho} = 1$) microstate.

Periodic boundary conditions

The two lattices described above, when used with normal periodic boundary conditions, clearly impose restrictions on the sizes of system we can simulate. The *fcc* lattice has a repeating unit of three stacking planes, whereas *hcp* repeats in two, and so the smallest possible system must consist of at least 6 stacking planes. It is possible to avoid this restriction by using more elaborate boundary conditions [77], but we chose to avoid this complication and simulate systems of 6^3, 12^3 and 18^3 spheres.

Another more subtle issue concerning the boundary conditions should also be addressed here as a potential pitfall when performing these simulations. In a usual MC simulation using 'standard' periodic boundaries, the particles (or rather the particle centres) are *always* kept inside the simulation cell. In this context, where the displacements have been decomposed into a lattice site and a displacement, this would require the simulation to change the position of the lattice site of any particle which moves across the simulation cell's 'walls' (see fig. 4.10(a)).

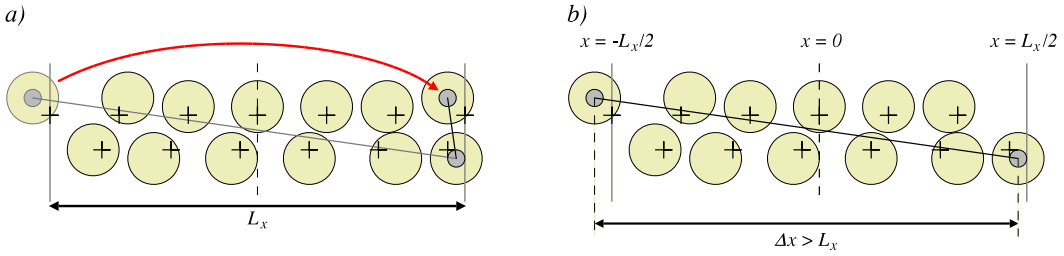


Figure 4.10: Implementations of periodic boundary conditions, for the x-direction. Dashed grey lines indicate the centre ($x = 0$) of the simulation cell, solid grey lines indicate the cell edges. (a) A full implementation of 'standard' periodic boundaries requires that any particle which moves outside the cell is wrapped around to the other side. For LSMC, this requires moving the particle's lattice site while leaving the displacement fixed. (b) If such particles are not wrapped around, then the separation between 'neighbouring' particles (in a given direction) can become greater than the cell size (in the same direction). This has consequences for the correct implementation of the boundary conditions (see main text).

However, it is not necessary to apply the periodic boundary conditions to the particle positions themselves if the calculation of the particle-particle interactions take them fully into account. That is, when determining the particle-particle separation ($\vec{r}_{ij}$) required by the pair-potential calculation, the boundary conditions are applied to the components of r_{ij} (Δx, Δy and Δz) so that the nearest image of the 'neighbouring' particle is always used. Originally [1], this was done using the 'standard' formula [34]

$$\Delta a \rightarrow \Delta a - L_a \text{int}\left(2\frac{\Delta a}{L_a}\right) \ , \tag{4.27}$$

where a refers to each of the x, y and z directions, and L_a refers the side-length of the simulation cell in that direction. Unfortunately, this scheme fails when particles are not constrained to stay within the same simulation cell-image [34, p. 326], because the separation between two particles may become greater than the cell-size (see fig. 4.10(b)). If this happens, then eqn. 4.27 will calculate the position of the *second* nearest image instead of the first. This can mean that neighbouring spheres do not 'see' each other, and so pass through one another. When using an (unbounded) random-walk algorithm, this mechanism lets the particles move far from their lattice sites, and causes the simulation to break down in a manner that is almost indistinguishable from the crystal actually melting.[3] However, this flaw can be remedied by using the following sequence of coordinate transformations

$$\begin{aligned} \Delta a &\rightarrow \Delta a - L_a \text{int}\left(\frac{\Delta a}{L_a}\right) \\ \Delta a &\rightarrow \Delta a - L_a \text{int}\left(2\frac{\Delta a}{L_a}\right) \ , \end{aligned} \tag{4.28}$$

which will always calculate the position of the first-nearest image, whichever simulation-cell image(s) the particles lie in.

4.8.2 Truncated interactions

To implement the calculation of the hard-sphere configurational energy (eqn. 2.2), the simulation *should* calculate the inter-particle potential energy for *all* distinct pairs of particles. However, this order N^2 calculation is extremely computationally expensive and indeed, unnecessary. In a short-ranged and strongly-caged system such as a hard-sphere crystal, each particle can usually be associated with the same lattice-site for the entire duration of even a long simulation.[4] Thus, particles which are separated by more than a few lattice-spacings will never interact, and so this interaction need not be included in the configurational-energy summation.

[3] This unphysical 'melting' was occasionally observed during the research published in [1], and a failure to trace this to the boundary conditions forced the original authors to use the top-hat move-generation scheme. This kind of algorithm forces particles to stay near their associated sites, and so prevents the 'melting' process.

[4] As long as the centre-of-mass motion is taken into account (see §4.8.4).

For a crystalline system, it is easiest to phrase the truncation of the interaction in terms of the number of nearest-neighbour shells for which the interaction is calculated. For the *fcc* and *hcp* lattices there are 12 nearest-neighbours (at a distance of $1r_{nn}$, see fig. 4.3), and 6 second-nearest neighbours (at $\sqrt{2}r_{nn}$). As these hard particles only interact upon touching, it should be sufficient to include just these 18 neighbour particles in the summation; more distant particles are extremely unlikely to touch one another. Nevertheless, such approximations should not be made carelessly, and so the consistency of this approximation was tested using the direct-difference method of Bennett [25] (see §2.4.3) to evaluate the free-energy difference between two different ranges of truncation. Although the direct-difference approach is not generally statistically reliable, we can reasonably expect the particle configurations associated with two different levels of truncation to be very similar and so it should be possible to apply this technique more or less as it stands. Having said that, the binary nature of the hard-sphere potential mean that it makes more sense to re-express the Bennett algorithm in terms of the number of configurations associated with a given set of constraints.

We use $\Omega(N, V, \alpha, C)$ to denote the number of configurations of a structure α, composed of N hard-spheres in a volume V, that satisfy a constraint C. We choose to consider a 'liberal' constraint that allows 2nd nearest neighbours to overlap (C_l), and a tighter constraint that disallows this (C_t). The entropy difference between the two systems operating under constraints C_t and C_l can be written as

$$\begin{aligned} \Delta s_c(N, V, \alpha) &= \frac{1}{N}\left[S(N, V, \alpha, C_t) - S(N, V, \alpha, C_l)\right] \\ &= \frac{1}{N}\left[\ln \Omega(N, V, \alpha, C_t) - \ln \Omega(N, V, \alpha, C_l)\right] \\ &= \frac{1}{N} \ln\left[\frac{\Omega(N, V, \alpha, C_t)}{\Omega(N, V, \alpha, C_l)}\right] . \end{aligned} \tag{4.29}$$

Now, under the assumption that the set of configurations that satisfy the tight constraint C_t form a consistent subset of the configurations under the looser constraint C_l, we may write

$$P(C_t|N, V, C_l) = \frac{\Omega(N, V, \alpha, C_t)}{\Omega(N, V, \alpha, C_l)} = \frac{\Omega(C_t|N, V, \alpha, C_l)}{\Omega(N, V, \alpha, C_l)} . \tag{4.30}$$

By consistent we mean that the subset of the configurations visited under the looser constraint which satisfy the tighter constraint must be identical to the configurations which would be explored using only the tighter constraint. This assumption is illustrated schematically in fig. 4.11, and in mathematical terms corresponds to the statement that

$$\Omega(C_t|N, V, \alpha, C_l) \equiv \Omega(N, V, \alpha, C_t) . \tag{4.31}$$

As long as this assumption holds, then the probability of satisfying the tighter constraint (eqn. 4.30) can be estimated as,

$$P(C_t|N, V, C_l) \stackrel{\text{EB}}{=} \frac{\mathcal{N}_t}{\mathcal{N}_l} , \tag{4.32}$$

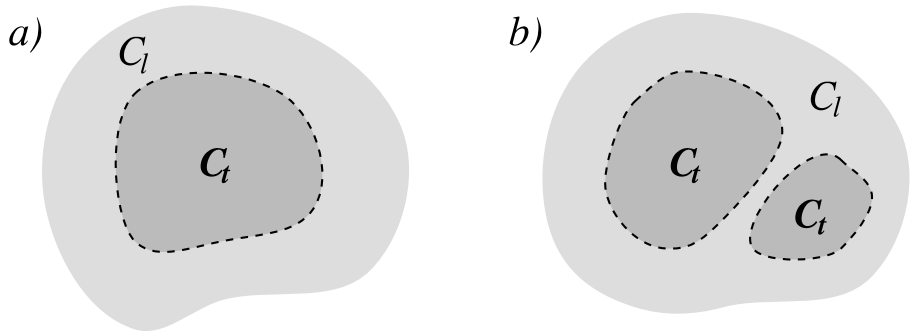


Figure 4.11: Configuration-space overlaps of the tight and liberal constraint microstates. (a) In our calculation, we assume that the tighter constraint forms a consistent (unique) subset of the more loosely constrained system. (b) If isolated fragments of configuration space associated the tight-constraint microstates can be explored as the loosely constrained system evolves, the calculation will over-estimate the size of the C_t configuration space, and so underestimate the free-energy difference between the two systems.

where $\mathcal{N}_t$ and $\mathcal{N}_l$ are the number of configurations satisfying constraints t and l which have been generated during a simulation which demands only that the looser constraint C_l is obeyed. To determine $\mathcal{N}_t$ and $\mathcal{N}_l$, we simply count the number of microstates satisfying each constraint generated by a simulation using 'ghost' 2nd nearest-neighbour interactions (see fig. 4.12). The entropy difference,

$$\Delta s_c(N, V, \alpha) = \frac{1}{N} \ln \left[\frac{\mathcal{N}_t}{\mathcal{N}_l} \right] , \tag{4.33}$$

can be calculated for both structures ($\alpha = fcc$ and $\alpha = hcp$), and this difference between $\Delta s_c(N, V, fcc)$ and $\Delta s_c(N, V, hcp)$ can be used to estimate free-energy cost associated with the truncation of the interaction with respect to the overall free-energy difference.

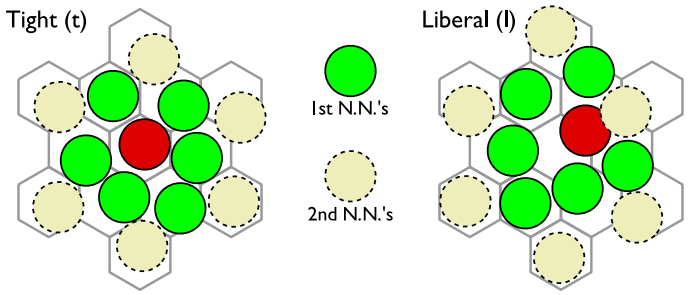


Figure 4.12: Simple illustrations of the two types of configuration to be counted in the estimation of the truncation error. When all particles satisfy the (tighter) constraint that no 2nd neighbours should overlap (left), it is counted as a type t (tight-constraint microstate. However, if *any* particle in the system is overlapping one of its 'ghost' second nearest-neighbours (right), then the configuration is counted as a type l (liberal-constraint) microstate.

4.8.3 Monte Carlo moves

The most basic MC move simply concerns changing the individual sphere displacements $\vec{u}_i$. All of the different move generation schemes described in §3.2.1 were tested, and the results of this comparison are presented in §4.10.1. However, before moving on to the other classes of MC move, there is one issue concerning the implementation of all of the move-generation schemes which should be mentioned here. In order to carry out the multicanonical weighting, it is necessary to know the current and trial values of the weight function, and the algorithm must therefore 'know' the value of the order parameter $\mathcal{M}$ at all times. To minimise the time spent calculating how a change in the displacement of one of the spheres affects the value of $\mathcal{M}$, we used a 'local' order parameter array. This array stores information on which neighbours currently overlap with each sphere in the conjugate phase. The changes to this array due to each change in the sphere displacements are calculated and stored, so that the associated change in $\mathcal{M}$ can be determined. Note also that for the random-walk algorithm, checks were put in place to test for the breakdown of the 'temporal cut-off' (described in §4.3). At regular intervals, the distance from each particle to its neighbouring sites was calculated. If any particle was found to be nearer another particle's lattice site that its own, the simulation run would be automatically halted.

In the canonical ensemble the only other type of Monte Carlo move is the lattice switch. After every sphere move, the value of $\mathcal{M}$ is checked. If we are in the gateway macrostate ($\mathcal{M} = 0$), then the lattice switch move is accepted.

In the isothermal-isobaric ensemble, the simulation implements volume dilations as well as particle and lattice-switch moves. The unconstrained aspect-ratio algorithm (UVM, see §3.2.2) was used for all of the work presented in this chapter. Note that the value of $\mathcal{M}$ depends on the size of our simulation cell as well as the local sphere displacements, and so the global nature of a volume move requires that both the global overlap order parameter $\mathcal{M}$ and the local overlap array be re-calculated. Therefore, the volume moves are computationally expensive, and so attempts to change the volume of the simulation cell were performed only once per sweep (on average). The volume-move parameter (Δl) was chosen such that the autocorrelation time of $\mathcal{M}$ was at a minimum (§4.10.1).

4.8.4 The centre of mass

When implementing the spatial cut-off as a means of associating a configuration with a lattice, it is important to understand the role of the centre-of-mass diffusion. Over time, the entire simulation cell will diffuse through space, and so the system will 'drift into the walls' of any spatial cut-off that is associated with the lattice. To investigate the importance of this effect the simulation code was designed

to have the option of proceeding in the centre-of-mass frame. Following the prescription given in ref. [17], the centre of mass was constrained as follows: after every successful particle move ($\Delta\vec{u} = \vec{u}'_i - \vec{u}_i$), the positions of *all* the particles were updated such that

$$\vec{u}_j = \vec{u}_j - \frac{1}{N}\Delta\vec{u} \qquad \forall\ j\ . \tag{4.34}$$

4.8.5 The close-packed limit

Here, instead of the hard-sphere potential, we implement the hard-dodecahedra condition (eqn. 4.22), where the parallel separation $u_{ij}^{\parallel}$ is calculated using

$$u_{ij}^{\parallel} = \vec{u}_{ij}.\hat{n}_{ij}^{\alpha}\ . \tag{4.35}$$

Again, an overlap array is used to keep an eye on the order parameter as the simulation proceeds. To speed up the calculation further, the set of nearest-neighbour vector unit vectors ($\hat{n}_{ij}^{\alpha}$) were stored in a look-up table.

To test that the implementation of the close-packed limit was correct, the following checks were used. It can be shown [76] that in the close-packed limit, the value of the distribution $P(u^{\parallel})$ at contact is

$$P(u^{\parallel} = 1) = \frac{1}{2}\ . \tag{4.36}$$

This result provides a way of testing the algorithm, by measuring $P(u^{\parallel} = 1)$ to check that the simulated value is indeed $^1/_2$. Furthermore, ref. [70] contains an estimate of the entropy in the high-density limit, of which one term is identified as the mean value of the square of the perpendicular component of the displacements at contact, i.e. $\langle\left[u^{\perp}\right]^2\rangle$ at $u^{\parallel} = 1$. This property can be measured during our MC simulations, and compared against the results given in [70]. The outcome of these tests is presented in §4.10.3.

4.9 Polydispersity

This calculation proceeded in an almost identical manner to that for the monodisperse case, as the only change is the implementation of the modified interaction potential (eqn. 4.24). As for hard-spheres, the interactions of each particle with its 1st & 2nd neighbours were included. However, the entropy cost associated with the truncation of the interactions was not estimated in this case.

4.10 Results

4.10.1 Move generation mechanisms

All three particle-move algorithms (§3.2.1) were compared; the random walk (RW), the top-hat (TH), and the bounded random walk (BRW). The relative efficiency of the different algorithms will be examined here, and in §4.10.8 their effect on the measured value of the free-energy difference will be examined.

When using the TH or BRW algorithms, the spatial cut-off must be made sufficiently large that it has no effect on the results. Unfortunately, as the spatial cut-off is increased, the efficiency of the TH algorithm drops rapidly, because large displacements are unlikely to be accepted and a wide top-hat will frequently generate unlikely trial positions. The BRW and RW algorithms should be (almost) identical in their efficiency, because as long as the spatial cut-off is large enough the two algorithms are essentially the same. Note that this is only true if the spatial and temporal cut-offs have the same effect, i.e. that no significant rearrangement of the crystal structure is observed under RW dynamics. This was indeed found to be the case for all of our simulations. Of course, this equivalence will fail should the system melt, as indeed it would at lower densities.

Figure 4.13(a) shows the acceptance rate of RW moves as a function of Δr, which appears to be exponential in form. The relative *efficiencies* of the RW and TH algorithms are compared in figure 4.13(b), using the autocorrelation time of the overlap order parameter in a single-phase (*fcc*) simulation as our criterion. Clearly, the RW algorithm has the two-fold advantage of yielding significantly shorter autocorrelation times while also avoiding the extra complications involved with implementing a spatial cut-off.

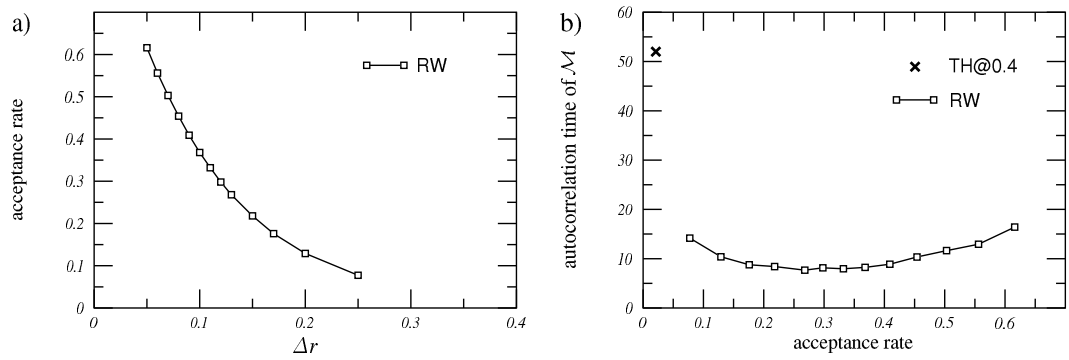


Figure 4.13: (a) The acceptance rate of RW moves as a function of the RW step parameter Δr. (b) Plot of the autocorrelation time of $\mathcal{M}$ as a function of the acceptance rate. Notice the TH result, showing the low acceptance rate and high autocorrelation time for a top-hat of width $r_{max} = 0.4$.

The fine-tuning of the RW algorithm is performed by varying the maximum step-size Δr, and to max-

imise the efficiency of the simulation, we minimise the autocorrelation time of $\mathcal{M}$ with respect to this parameter. As can be seen from fig. 4.13, the minimum in the autocorrelation time of the order parameter occurs at a RW acceptance rate of about 30%, corresponding to a maximum step size of $\Delta r = 0.13\sigma$ at our chosen density ($\tilde{\rho} = 0.7778$). As a convenient scale on which to consider 'small' lengths such as Δr, we define the minimum sphere separation,

$$\delta = \sigma(\tilde{\rho}^{-1/3} - 1) \, , \tag{4.37}$$

which is the shortest distance between two sphere surfaces in the perfect-crystal configuration. For the chosen density, the value of δ is 0.09σ, and so the best RW move parameter is ≈ 1.4 times greater than the minimum sphere separation.

In the NPT ensemble, a range of pressures were explored, and the density for each pressure was determined from the simulation data. A simple linear interpolation was then used to estimate the pressure for which the corresponding density is $\tilde{\rho} \approx 0.7778$. This produced a value for the pressure of $\tilde{\mathrm{p}} = 18.74$, for which the optimum value of Δl was found to be 0.005, corresponding to an acceptance ratio of approximately 50%.

For the simulation of the close-packed limit, the optimum value of Δr was found to be ≈ 0.5, and as in the finite-pressure case, the optimal acceptance rate was approximately 30%. For the polydisperse system, exactly the same parameters were used as for the monodisperse NVT simulations.

4.10.2 Neighbour-interaction entropy difference

The entropy difference between a simulation of $N = 6^3$ hard-spheres with only first nearest-neighbour interactions and the same simulation using both 1st & 2nd neighbours (Δs_c, eqn. 4.33) is shown in figure 4.14. While only a small range of densities have been investigated, it is clear that as one moves away from the melting density (~ 0.733), the entropy difference diminishes very rapidly. In fact the value of Δs_c *must* tend to zero in the close-packed limit, as only 1st nearest-neighbours can possibly interact when the particle separation becomes small.

However, the results shown in fig. 4.14 are not in agreement with the results presented in ref. [77]. For a system of $N = 8^3$ hard-spheres, ref. [77] estimates the value of Δs_c to be $8(2) \times 10^{-5}$ for both the *fcc* and *hcp* structures when $\rho/\rho_{cp} = 0.739$. In this work, the estimated entropy differences at that density are an order of magnitude smaller; $\Delta s_c = 5.0(2) \times 10^{-6}$ for *hcp* and $6.0(2) \times 10^{-6}$ for *fcc*.

The reason for this disagreement is unclear, and insufficient information about the details of their calculation is supplied by the authors of ref. [77] to allow any firm conclusions to be drawn. If the estimates presented here are incorrect, this must imply that the calculation's central assumption (eqn. 4.31) has

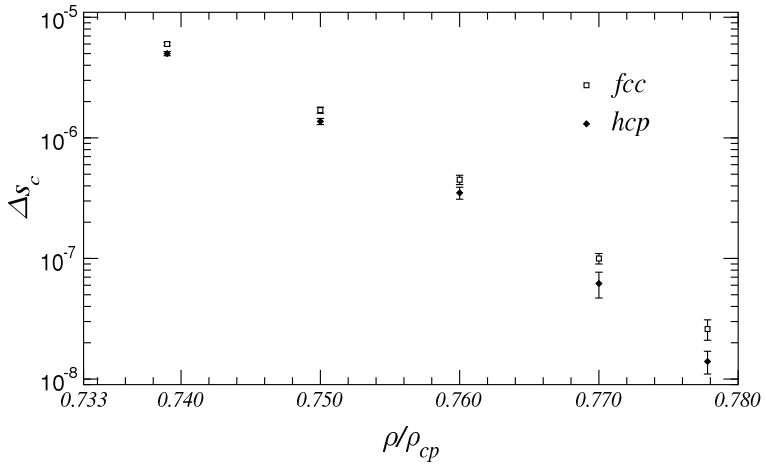


Figure 4.14: The entropic 'cost' of changing from 1st & 2nd to 1st nearest-neighbour interactions ($N = 6^3$). Δs_c (from eqn. 4.33) is plotted over a small range of densities close to melting. Note that the vertical (Δs_c) scale is logarithmic.

broken down, and that the sub-space of configurations that satisfy the tighter (1st & 2nd neighbour) constraint is in some way fragmented.

Despite this uncertainty, the *consequences* of these calculations (in terms of the effect on the *fcc–hcp* entropy difference) should not be severe. Using either set of results, the difference between the values of Δs_c for the two structures is small (on the scale of the errors in the *fcc–hcp* entropy difference, see §4.10.7). Moreover, all hard-sphere lattice-switch simulations included both 1st & 2nd nearest neighbour interactions (except, of course, those performed in the close-packed limit). Therefore the value of Δs_c is in fact a rather liberal estimate of the errors, and should really be thought of as an extreme upper bound: the errors associated with not including 3rd, 4th... neighbour interactions will surely be much smaller than this.

4.10.3 Consistency checks

In the canonical ensemble, the pressure of a $N = 6^3$ *hcp* crystal was measured (via eqn. 4.11) at the chosen density of $\tilde{\rho} = 0.7778$. This required measuring the 'closest approach' (d_s) distribution, and determining the mean value of d_s. The distribution appeared to be exponential in form; an observation supported by the fact that the mean value $\langle d_s \rangle = 0.1152(5) \times 10^{-3}$ is equal to the standard deviation $\sigma_{d_s} = 0.1149(6) \times 10^{-3}$ to within the statistical accuracy of the calculation. These measurements produce a pressure of $\tilde{\mathrm{p}} = 18.83(8)$, consistent with the result of $\tilde{\mathrm{p}} = 18.76$ quoted in ref. [70].

A more detailed analysis was performed in the constant-pressure ensemble. Table 4.1 shows the measured densities for the two smallest system sizes, for both structures, and compares the results to those

N	Phase	Simulation Time [MCS]	$\tilde{\rho}$ of [70]	$\tilde{\rho}$	c/a
6^3	*fcc*	3.5×10^6	0.77757	0.7775(1)	1.6332(5)
6^3	*hcp*	3.5×10^6	0.77753	0.7776(1)	1.6323(7)
12^3	*fcc*	$9 \times 5 \times 10^6$	0.77757	0.7773(1)	1.6333(3)
12^3	*hcp*	$9 \times 5 \times 10^6$	0.77753	0.7770(3)	1.6332(3)

Table 4.1: Densities and c/a-ratios measured in the NPT ensemble, at $\tilde{\mathrm{p}} = 18.74$, for the *fcc* and *hcp* crystals using two different system sizes, along with the density predicted by the equations of state from ref. [70]. The $N = 12^3$ single phase information was determined from a LS double-phase simulation by reweighting each volume measurement so as to unfold the bias introduced by the MCMC weight function.

in ref [70]. Our results agree with the predictions from the equations of state given in ref. [70], but the densities have not been determined accurately enough to resolve the *difference* between the *fcc* and *hcp* densities. Table 4.1 also shows the value of the c/a ratio measured during these simulations. None of these measurements is significantly different from the ideal value ($c/a = 1.63299$). Evidently, whatever difference there is between the true values of these quantities in each of the phases is not easily resolved using these methods.

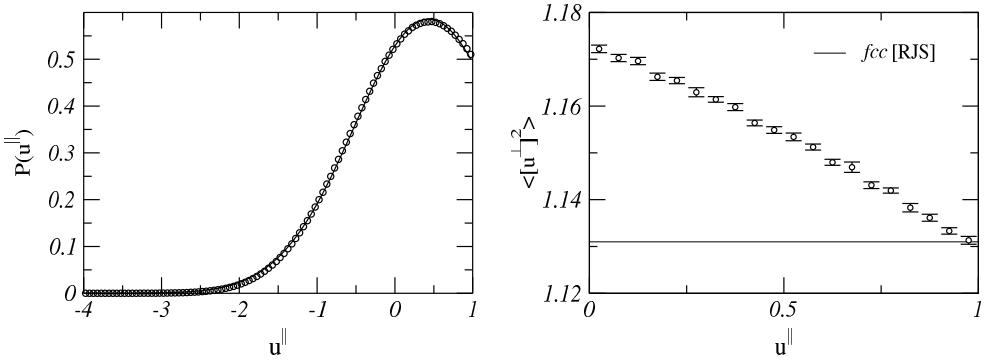


Figure 4.15: Measurements of $u^{\parallel}$ and $\langle [u^{\perp}]^2 \rangle$ in the close-packed limit, for $N = 6^3$ hard dodecahedra arranged on an *fcc* lattice. (a) The distribution $P(u^{\parallel})$, with a Gaussian fitted curve (solid line). The errors are significantly smaller that the symbol-size. (b) Plot of the average value of $[u^{\perp}]^2$ as a function of $u^{\parallel}$. The value of this property at $u^{\parallel} = 1$ is in agreement with the value given by ref. [70] (solid line, denoted as [RJS]).

The close-packed limit simulations were tested in two ways, based on the distribution of $u^{\parallel}$ and on the value of $\langle [u^{\perp}]^2 \rangle$ as a function of $u^{\parallel}$, and the results are shown in figure 4.15. As asserted in §4.6, the value of $P(u^{\parallel})$ when $u^{\parallel} = 1$ should be 0.5, and this is found to be the case ($P(u^{\parallel}) = 0.494(5)$). Also, the average value of $[u^{\perp}]^2$ at contact ($u^{\parallel} = 1$) was compared against the results provided in ref. [70]. Our measurements indicate that $\langle [u^{\perp}(u^{\parallel} = 1)]^2 \rangle$ is $1.130(1)$, in agreement with the value of 1.131 given in [70].

4.10.4 Variations of the lattice switch

The fundamental barrier to accepting the lattice-switch move is the number of overlapping spheres that would be created by it. The average number of overlapping pairs of particles that would be created by the switch in a normal importance sampling simulation is termed the 'equilibrium overlap count', and it makes sense to use any freedom we have concerning the definition of the lattice switch to minimise this number. As mentioned in §4.3.1, there are any number of different possible choices of lattice-switch mapping. Here, a range of different site-site mappings have been compared, using the equilibrium overlap count as an (inverse) measure of their efficiency. Table 4.2 shows the results for a variety of mappings, chosen to expose the different factors that control the mapping efficiency. Mapping number 1 is the one described in fig. 4.6: the notation $(0, -\vec{t}, +\vec{t})$ signifies that the three pairs of planes (counting from the top of fig. 4.6) are translated by 0, $-\vec{t}$ and $+\vec{t}$ respectively. A similar convention is used to label mappings 2 and 3. In mapping 4 ('random-plane') a *hcp* configuration is generated by taking an *fcc* configuration and re-stacking its close-packed planes in a random order, but in a *hcp* pattern. In mapping 5 ('random-site') a *hcp* configuration is generated by mapping the particle displacements in an *fcc* configuration randomly onto the sites of an *hcp* lattice.

mapping	description	effect	$m = \langle \mathcal{M} \rangle / N$
1	$(0, -\vec{t}, +\vec{t})$	*fcc* → *hcp*	0.150(1)
2	$(0, 2\vec{t}, -2\vec{t})$	*fcc* → *hcp*	0.183(1)
3	$(0, 3\vec{t}, -3\vec{t})$	*fcc* → *fcc*	0.194(1)
4	random-plane	*fcc* → *hcp*	0.373(2)
5	random-site	*fcc* → *hcp*	0.820(3)

Table 4.2: The efficiency of different mappings (for $N = 12^3$ and $\tilde{\rho} = 0.7778$), as measured by the average number of overlaps (per sphere) that they generate. See main text for details.

The random-site mapping (#5) shows the largest overlap count, which is perhaps unsurprising given that this transformation preserves the least information about how *neighbouring* sphere were arranged in the *fcc* phase. Using the random-plane mapping (#4) cuts the overlap count by a factor of (a little more than) 2 with respect to the random-site transformation. This efficiency gain simply reflects the fact that of the $6N$ possible overlaps between nearest neighbours, only the $3N$ associated with neighbours in different (but adjacent) planes can now contribute. Mapping 3 simply generates one *fcc* configuration from another (it is useful only because it is informative): its overlap count is cut by a further factor of 2. This reflects the fact that this mapping (as with mappings 1 and 2) moves the close-packed planes in

pairs, thus guaranteeing that no overlaps occur between the two members of each pair.

Mappings 2 and 1 show further – smaller but still practically useful – cuts in the overlap count. The origin of these gains is more interesting. It is clear that they must reflect the size of the translation vector used: mappings 1, 2 and 3 differ only in this respect. This vector controls the extent of the shear which the mapping introduces between successive pairs of planes, and so the following interpretation seems reasonable. The displacements of particles in adjacent planes will be correlated to some extent, with undulations in one surface (the z components of the displacements within each plane) matched to undulations in its neighbour. The smaller the shear, the more closely these undulations will *remain* matched to one another (in the conjugate configuration), and the smaller the overlap count. With increasing shear, this advantage is lost and the behaviour should (and indeed does) approach the limit (one quarter of the overlap count for mapping 5) one would expect in the absence of such correlations. The fact that this 'approach' is already apparent in the performance of mapping 2 is consistent with the fact that the measured correlation length of the surface undulations at the density concerned was found to be close to the magnitude of the translation vector $\vec{t}$.

These results help to clarify the factors which control the overlap count of the mapping (#1) we have actually used for the free-energy difference calculation. It is tempting to attribute the overlaps to the fact that the LS (*fcc→hcp*, say) maps each particle from an environment in which adjacent close-packed planes have different stacking labels (A and C, say) to one in which they have the same label (C, say). The results for mappings 1-3 show that it would be misleading to think this way. The overlaps simply reflect the numbers of particles that 'see' a new adjacent close packed plane (irrespective of its label), and the extent to which it is 'new'. Thus, any simple tuning of the displacement representation (the choice of TSSM transformation) is likely to be of no advantage here.

4.10.5 Evolving the weight function

For the smallest system ($N = 6^3$), the simple VS technique proved to be adequate, but significantly less efficient than the TP method (see figure 4.16 for an example). Each VS iteration was 5×10^6MCS in length, whereas the TP only required $\sim 10^5$MCS.

For the middle ($N = 12^3$) system, the VS method is no longer a viable approach. However, the TP method worked well here, generating a useful weight function in $\sim 10^6$ MCS. This 'first guess' was then refined using a 5×10^6MCS VS iteration. Finally, for the largest system ($N = 18^3$), it was necessary to use the mobile-barrier TP method, for which $\sim 10^6$ MCS were found to produce a satisfactory estimate of the weight function. A single 5×10^6MCS VS iteration was performed as a final refinement step.

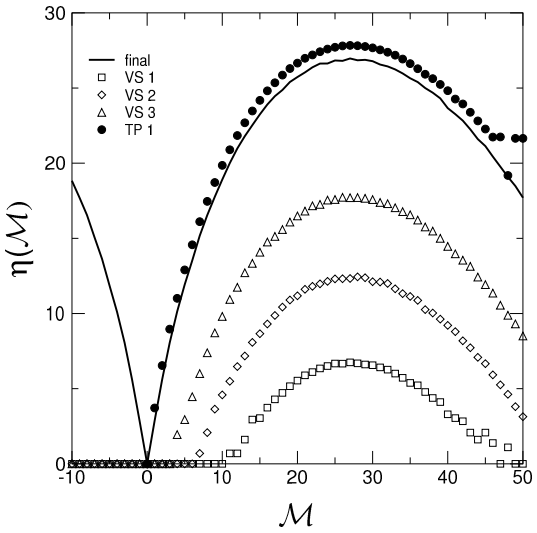


Figure 4.16: Weight evolution for $N = 6^3$ hard spheres, comparing the visited states (VS) and transition probability (TP) methods. The points marked VS are the results of the first 3 iterations of the visited-states algorithm, initiated from an *fcc* equilibrium state. The points marked TP emerge from one application of the transition probability method. The solid line shows a refined (usable) set of weights.

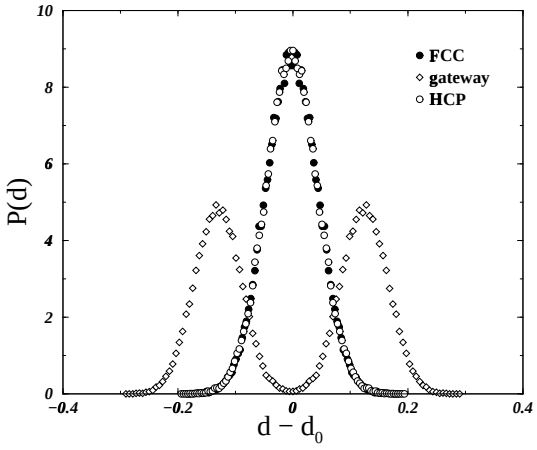


Figure 4.17: Distribution of the separation d between adjacent close-packed planes in a system of $N = 6^3$ spheres at $\tilde{\rho} = 0.7778$, in the equilibrium *hcp* and *fcc* macrostates, and in the gateway ($\mathcal{M} = 0$) macrostate. The separation is measured with respect to the equilibrium separation d_0 and is expressed in units of the equilibrium sphere separation δ.

4.10.6 Microscopic mechanics of the multicanonical biasing

The lattice switch can only be accepted when launched from a small subset of the configurations actually visited: these, by definition, are the 'gateway configurations'. As noted earlier, one could identify *a priori* those configurations (e.g. as characterised by 'small enough' displacements) which fall into this set. But we have elected, rather, to let the system (the algorithm) identify them on the basis of their defining characteristic — that they have zero overlap order parameter $\mathcal{M}$. It is therefore interesting to investigate the *microscopic* characteristics of the configurations picked out by this constraint.

Figure 4.17 shows the distribution of the separation, d, between adjacent close-packed $(x-y)$ planes for $\mathcal{M}$-macrostates corresponding to the equilibrium *fcc*, the equilibrium *hcp*, and the gateway $(\mathcal{M}=0)$ regions. The separation, d, is defined as the difference (in the z-direction) between the centre-of-mass positions of any pair of neighbouring planes.

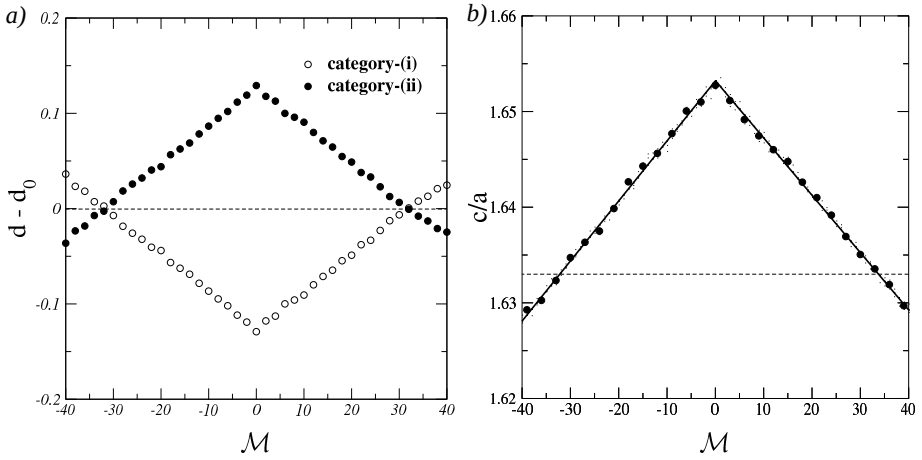


Figure 4.18: (a) The mean value of the separation d between adjacent close-packed planes in a system of $N=6^3$ spheres at $\tilde{\rho}=0.7778$, as a function of $\mathcal{M}$. The separation is measured with respect to the equilibrium separation d_0 in units of δ. Category-(i) planes (see fig. 4.6) are translated together by the LS transformation, while category-(ii) planes are translated through different amounts. (b) The evolution with $\mathcal{M}$ of the c/a-ratio in a constant-*pressure* ensemble ($N=6^3$, $\tilde{\mathrm{p}}=18.74$). The horizontal line marks the ideal close-packed value, and the equilibrium values of $\mathcal{M}$ appear to correspond to the undeformed states of both *fcc* and *hcp*.

The macrostates corresponding to the equilibrium crystal structures have similar, near-Gaussian, d-distributions. In contrast, for the gateway macrostate the distribution is *bi-modal*: in this macrostate, some planes are systematically moved closer to one another, while (in equal measure) others are shifted apart. On closer examination one finds that it is the planes which are *translated together* by the LS (e.g. the pair of planes marked (i) in fig. 4.6) that fall into the first category, while the planes that are *translated differently* by the LS (e.g. the pair of planes marked (ii) in fig. 4.6) fall into the second. The evolution, with $\mathcal{M}$, of the mean plane separation (for both categories) is shown in fig. 4.18(a). The

behaviour thus unearthed is entirely reasonable. The LS operation can *only* create overlaps between neighbouring planes which are translated by different amounts (sheared with respect to one another). The algorithm resolves the task set by the bias towards $\mathcal{M} = 0$ by moving these pairs of planes (the ones vulnerable to overlaps) further apart, at the expense of a compression of the others. The gateway configurations have further distinctive features (with respect to their equilibrium counterparts): the root-mean-square particle displacement, in the z-direction, is reduced; and the in-plane correlation length of these displacements (the undulations of the close-packed planes) is enhanced along the $\vec{t}$-direction.

In simulations conducted at constant pressure, the layer-compression (which is still present) is supported by a second effect. Fig. 4.18(b) shows that the algorithm now exploits the additional degrees of freedom (the *shape* of the simulation cell) to locate gateway states with values of the c/a ratio enhanced above the ideal close-packed value. This increases the distance between the stacking-planes, which will obviously enhance the probability of accepting the switch.

It is tempting to say that the sampling is 'intelligent'. In any event it is clear that the algorithm locates and utilises configurations which it would be difficult to exploit explicitly in the design of the switch operation.

4.10.7 Free-energy differences

The essential output of a LS-simulation is the normalised probability distribution of the overlap order parameter $P(\mathcal{M})$, reweighted to remove the bias in the multicanonical distribution actually measured during the simulation. Figure 4.19 shows the results for this distribution (at $\tilde{\rho} = 0.7778$) for three different N values.

As one would expect the distributions each comprise two peaks (one associated with each phase) each of which is approximately Gaussian and sharpens with increasing N. Note the close correspondence between the equilibrium overlap counts for the two structures. This result is not *required* by definition, or any obvious symmetry. Rather it should be seen as a further manifestation of the similarity of the local particle environments in the two structures. The relative weights of the two peaks is a direct measure of the difference between the entropies of the two structures (eqns. 4.9, 4.10, 4.7). Since the entropies are extensive the ratio of the peak weights grows exponentially with N. The fact that, at least for our smaller systems, the two peaks can even be displayed on the same scale is a reflection of the exceptionally delicate balance between the two entropy densities.

The fact that peak-weight difference diverges exponentially fast with N presents no serious computational problem: the multicanonical procedure is designed to cope with (and quantify) differences of this

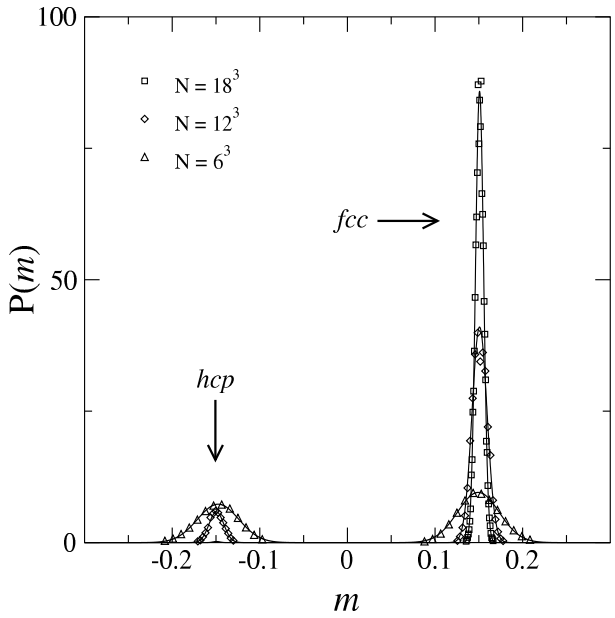


Figure 4.19: The probability distribution of the *intensive* overlap order parameter per particle, $m \equiv \mathcal{M}/N$, for three different system sizes, at $\tilde{\rho} = 0.7778$. The lines provide Gaussian guides to the eye. The statistical uncertainties on each of the data points are smaller than the symbol size. The entropy difference is identified from the logarithm of the ratio of the integrated weights of the two peaks. The *hcp* peak for the largest system is not visible on this scale.

scale. However, as the system size is increased the 'depth' of the probability trough that the MCMC sampling must negotiate also diverges, and this makes determining a suitable weight function more difficult.

Figure 4.19 allows one to *see* that *fcc* is the thermodynamically preferred structure. This conclusion is expressed quantitatively in the results gathered in table 4.3. The results of the present work are in full accord with the results (both from LSMC and via integration methods) reported by Pronk and Frenkel [81]. The close correspondence between the results for $N = 12^3$ and $N = 18^3$ confirms that the former system is already representative of the thermodynamic limit. While our results seem to be at variance with the original integration-method result of Woodcock [30], his revised estimate [78] appears to be consistent with our predictions (given the large uncertainty attached to that result). Table 4.3 also shows the results of our studies in the close-packed limit, using the hard-dodecahedron representation (eqn. 4.21). They are close to those reported by Mau and Huse [77], but the differences (for the smaller systems, particularly) appear to be statistically significant. Figure 4.20 gives an alternative view of these results. It utilises the parameterisation of the measured pressure difference between the two phases provided by Speedy [70] to determine the entropy difference as a function of density, *given* the entropy difference at a chosen reference density; we have used the results of the present work at $\tilde{\rho} = 0.7778$.

Table 4.4 shows the results of our studies in the constant pressure ensemble. The quantity of interest

ρ/ρ_{cp}	N	Δs	$(10^{-5} \times k)$	Method	Ref.
0.731	512	85	(10)	SM	[77]
0.736	12000	500	(100)	IM	[30]
0.736	12000	230	(100)	IM	[78]
0.736	12096	87	(20)	IM	[79]
0.736	13824	99	(6)	IM	[80]
0.739	512	90	(4)	LS	[77]
0.7778	216	132	(4)	LS	[81]
0.7778	1728	112	(4)	LS	[81]
0.7778	1728	113	(4)	IM	[81]
0.7778	216	133	(3)	LS	PW
0.7778	1728	113	(3)	LS	PW
0.7778	5832	110	(3)	LS	PW
1.00	12000	260	(100)	IM	[78]
1.0	512	110	(20)	SM	[77]
1.0	64	91	(5)	LS	[77]
1.0	216	107	(4)	LS	[77]
1.0	512	119	(3)	LS	[77]
1.0	1000	113	(4)	LS	[77]
1.0	216	131	(3)	LS	PW
1.0	1728	125	(3)	LS	PW

Table 4.3: The difference in the entropy densities of the *fcc* and *hcp* structures, $\Delta s \equiv \Delta s_{fcc,hcp}$ (eqn. 4.9); the associated uncertainties are in parenthesis. The results of reference [78] supersede those of reference [30]. IM stands for integration method; SM is the lattice shear method of [77]. To obtain our results (PW), the small system required a total of $\sim 25 \times 10^7$ MCS, the middle system took $\sim 12 \times 10^7$ MCS and the large system required $\sim 4 \times 10^7$ MCS. However, in terms of actual CPU runtime, the small system required a total of ~ 200 hrs, the medium system took ~ 190 hrs and the large-system runs lasted ~ 500 hrs.

here is the difference between the Gibbs free-energy densities at the chosen pressure, which follows from the relevant distribution with the aid of eqn. 4.14. As explained in §4.5, we would expect the magnitudes of Δg and Δs to be almost identical in the thermodynamic limit. This is indeed the case (within the uncertainties quoted, c.f. table 4.3), with the main difference between the ensembles being that the finite-size effects appear to be significantly smaller in the NPT case.

$\tilde{\mathrm{p}}$	$\tilde{\rho}_{hcp}$	$\tilde{\rho}_{fcc}$	N	Δg	$(10^{-5} \times kT)$
14.58	0.7776(1)	0.7775(1)	216	-113	(4)
14.58	0.7770(3)	0.7774(2)	1728	-112	(3)

Table 4.4: The difference between the Gibbs free-energy densities of the *fcc* and *hcp* structures $\Delta g \equiv \Delta g_{fcc,hcp}$ (eqn. 4.14); the associated uncertainties are in parenthesis.

Finally, we consider the effects of a small degree of polydispersity on the *fcc–hcp* entropy difference (§4.7). For our selected (NVT) density, $z = 0.7778$, we used the maximum polydispersity (eqn. 4.26)

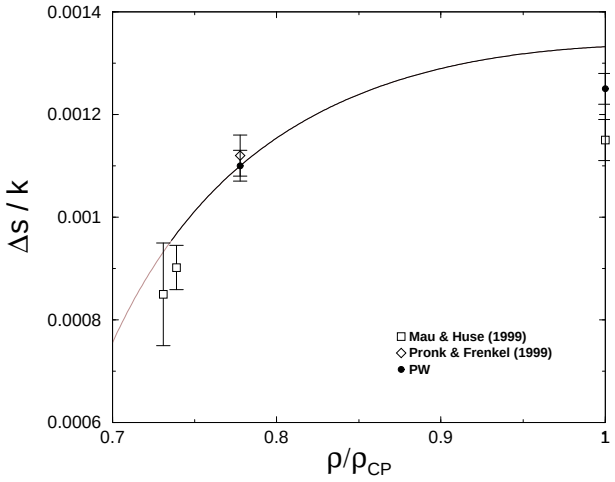


Figure 4.20: The difference in the entropy densities of the *fcc* and *hcp* structures, $\Delta s \equiv \Delta s_{fcc,hcp}$ as a function of reduced density. The data points are as given in table 4.3. The solid line is the result of an integration of the pressures of the phases [70]. Note that this line passes through our result at $\tilde{\rho} = 0.7778$ by *construction*.

of $\sigma = 0.087370\delta$. The simulation used 12^3 hard spheres, and predicted an entropy difference of $\Delta s_{fcc-hcp} = 128(5) \times 10^{-5} Nk$. This value is slightly larger than in the monodisperse case, but this result may be misleading as no finite-size effect analysis has been performed.

4.10.8 Getting it right

Having compared the *efficiency* of the various move-generation algorithms in §4.10.1, the *quality* of the results produced using these algorithms will now be explored. A number of different implementations of the algorithms were used to evaluate the free-energy difference between the *fcc*-*hcp* structures for the $N = 6^3$ hard-sphere system at a density of $\tilde{\rho} = 0.7778$, the results of which are presented in figure 4.21. Both the BRW and TH algorithms are in agreement (within statistical uncertainty) when the cut-off is small ($0.35/0.45 r_{nn}$), and these results are consistent with those in the original LS publication [1]. However, a different result is obtained when the cut-off is large (e.g. $BRW@2.0 r_{nn}$), or when the centre-of-mass of the system is fixed. In this latter case, the BRW algorithm give results in agreement with the RW algorithm (which is not centre-of-mass frame dependent).

Therefore, it appears that when using a bounded algorithm (such as BRW or TH) *without* fixing the centre-of-mass, the results are inconsistent with those found using the RW algorithm. The RW results have been corroborated independently by the authors of [81], so the questions remains as to why the small spatial cut-off should give incorrect results. As described in §4.8.4, the failure to conduct a simulation in the centre-of-mass frame when using a bounded algorithm may allow the particles to ‘drift

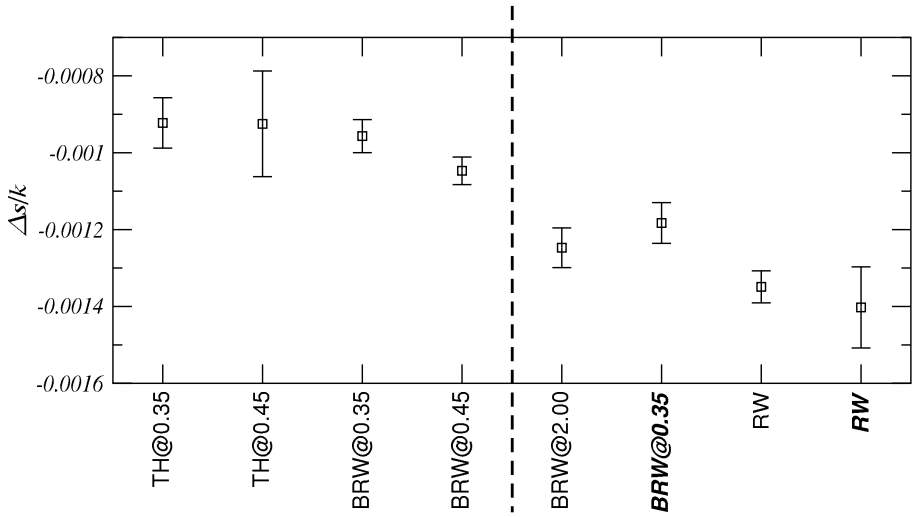


Figure 4.21: The entropy difference between *fcc* and *hcp*, as determined by a range of different algorithms. Simulations which employed a spatial cut-off are indicated with $@D_{max}$, where $D_{max} = 2r_{max}$ (see §3.2.1). The two evaluations performed in the centre-of-mass frame are indicated using bold type.

into the walls', but the consequences of this drift are difficult to estimate. However, as the number of particles increases, it is clear that the rate of diffusion of the centre of mass must tend to zero. Therefore the failure to take into account the centre-of-mass motion should act as a finite-size effect.

Figure 4.22 compares the entropy differences published in [1] (which used a TH@0.35 algorithm) to the results taken during this work (using a RW scheme); the disagreement between the two algorithms is clear. We have argued that this disagreement should be a finite-size effect, but these results show that if this is the case, very large system sizes would be required to overcome this problem. For this reason, we believe any algorithm which uses a spatial cut-off should *always* be implemented in the centre-of-mass frame.

4.10.9 Mixed-crystal stacking patterns

As mentioned in §4.2, there are any number of possible stacking patters, apart *fcc* and *hcp*, which may play a role in the equilibrium phase behaviour of the hard-sphere solid. No explicit evaluation of the entropies of these mixed structures has been performed here, but such calculations have been performed by other authors. Mau & Huse [77] used lattice-switch MC and a related shear technique to evaluate the relative entropies of a number of stacking patterns, and inferred that *fcc* is the equilibrium structure over all other stacking arrangements. This work is consistent with that of Pronk & Frenkel [81], who used LSMC to evaluate the cost of an *fcc–hcp* interface (see fig. 4.23(a)), and found the entropy of an

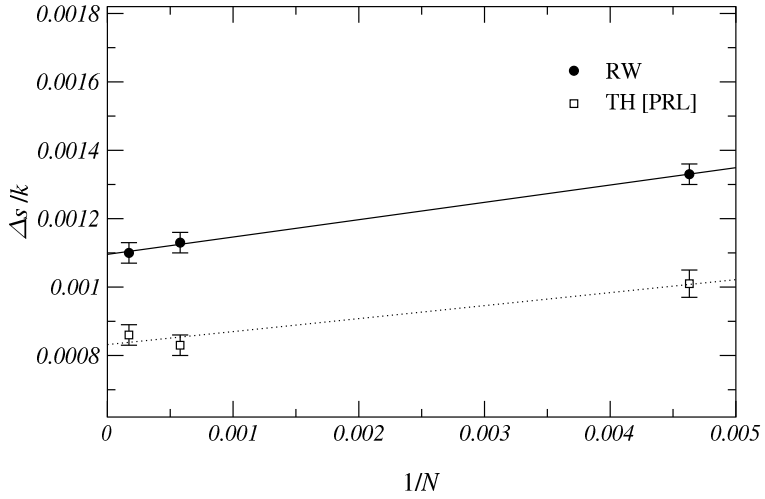


Figure 4.22: Plot showing the entropy-density difference calculated in the present work (using the random-walk algorithm, RW), and the results for exactly the same quantity from [1] (denoted as TH [PRL]). The entropies are plotted against the inverse of the system size, with linear fits for each set of data.

fcc crystal would be *reduced* by the introduction of *fcc–hcp* stacking faults.[5]

While there is little doubt that the thermodynamically stable phase (i.e. the preferred phase as $N \to \infty$) is indeed *fcc*, this is not necessarily true for a finite system. All of the work described so far is based on finite-system simulation, which have been used to *deduce* properties of the thermodynamic limit. But this is *not* enough to guarantee that the behaviour displayed will actually *be* that of the thermodynamic limit. To see this, we examine the stability of the perfect *fcc* crystal with respect to *hcp*-type stacking faults, and in particular the dependence of this stability upon the system size.

Displacement entropy versus stacking entropy

Consider a system of N hard spheres arranged in $N_{\parallel}$ close-packed layers of $N_{\perp}$ particles. Following reference [77] one may conveniently index each of the close-packed layers with a pseudo-spin (Ising-like) variable σ, where $\sigma_i = +1$ signifies that layer i has an *fcc* environment while $\sigma_i = -1$ implies an *hcp* environment (see fig. 4.23). The probability of a particular stacking sequence $\{\sigma\}$ then satisfies

$$\ln P(\{\sigma\}|N,V) = S(N,V,\{\sigma\}) + constant\ , \tag{4.38}$$

where $S(N,V,\{\sigma\})$ measures the entropy associated with the configurations (particle displacements) consistent with the particular structure $\{\sigma\}$. Following [77] this entropy (we will refer to it here as

[5] Of course, there will always be a finite concentration of *localised* defects in any large crystal, but it is the population of *extended* (i.e. macroscopic, and thus structure-defining) stacking faults that concerns us here.

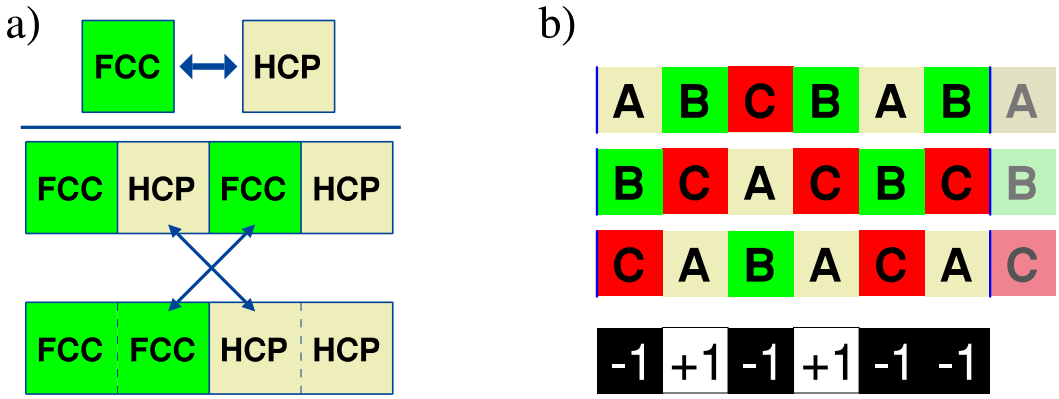


Figure 4.23: (a) Schematic illustration of the simulation cells used by Pronk & Frenkel [81]. Both an *fcc–hcp* lattice-switch algorithm and an Einstein crystal integration were used to evaluate the *fcc–hcp* entropy difference. Furthermore, an lattice-switch algorithm was also applied to an elongated system composed of *fcc* and *hcp* regions (the stacking direction is horizontal here). By switching between two lattices, which have the same volume of *fcc* and *hcp* but composed of either 2 or 4 separate domains, the cost of constructing an *fcc–hcp* interface could be estimated. (b) Any stacking pattern can be mapped (degenerately) onto an Ising-like model ($\sigma = -1$ for *hcp* and $\sigma = +1$ for *fcc*). Periodic images are shown as pale boxes.

'displacement entropy') can usefully be written in the form of an expansion:

$$S(N, V, \{\sigma\}) = N s_0 + N_\perp h \sum_i \sigma_i + N_\perp J \sum_{<ij>} \sigma_i \sigma_j + \dots \ . \tag{4.39}$$

The expansion is effectively ordered in the *range* of the entropic inter-layer 'interactions': the dots (...) represent contributions from interactions (microscopically, displacement-displacement correlation functions) extending over more than 4 layers. The analysis of reference [77] indicates that the series converges quickly, except close to melting. If we neglect the interaction terms altogether we may make the identification

$$h = \frac{1}{N_\perp N_\parallel} \left[S(N, V, \{\sigma = +1\}) - S(N, V, \{\sigma = -1\}) \right] = \frac{\Delta s_{fcc,hcp}}{2} \ , \tag{4.40}$$

and, from eqn. 4.38,

$$\langle \sigma \rangle = \frac{1}{N} \sum_{\{\sigma\},i} P(\{\sigma\}|N, V) \sigma_i = \tanh \left[N_\perp h \right] = \tanh \left[\frac{N_\perp \Delta s_{fcc,hcp}}{2} \right] \ . \tag{4.41}$$

The correspondence with a 1D paramagnet is clear. The familiar competition between orientation energy and entropy is played out here as a competition between displacement entropy and stacking entropy, with $N_\perp$ playing the role of an inverse temperature. Thus, the thermodynamic ideal ($\alpha = 1$) is only attained to the extent that $N_\perp \Delta s$ is large in comparison with unity. The entropy difference Δs is to be used in the units of table 4.3, and is dependent upon density.

To translate this result into the more usual notation of the field [82], we introduce a parameter α measuring the probability that a chosen close-packed plane sits within an *fcc* environment as distinct from

the *hcp* environment.[6] Thus, the value of α is zero for pure *hcp*, one for pure *fcc*, and $^1/_2$ for *rhcp*. The relationship between this measure and that used in eqn. 4.41 is simply,

$$\alpha = \frac{1}{2}(1 + \langle\sigma\rangle) \; . \tag{4.42}$$

The behaviour of α as a function of $N^{\perp}$ is shown in figure 4.24, using the value of Δs calculated for a system of $N = 18^3$ hard-spheres at a fixed density of $\tilde{\rho} = 0.7778$.

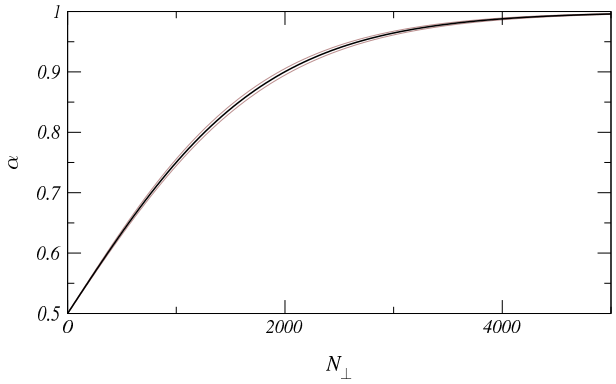


Figure 4.24: Plot of the stacking order α as a function of the number of particles per stacking-planes, for a crystallite composed of hard-spheres at a density of $\tilde{\rho} = 0.7778$. Small crystallites are predicted to display purely random-stacking, and only those with $N_{\perp} \gtrsim 5000$ can be expected to display pure *fcc* stacking behaviour. Pale lines indicate the change in this function due to the errors associated with our estimate of the *fcc–hcp* entropy difference.

4.11 Comparison with experiment

Notwithstanding the simplicity of the model, these results do have implications for experimentally-realizable systems. The immediate relevance to atomic systems is tenuous,[7] but the model has been widely used to account for the behaviour of assemblies of 'hard', 'spherical' colloidal particles [82]. Since the predicted entropy-density difference is so small there are potentially many ways (e.g. residual interactions between the spheres, polydispersity) in which the applicability of the theory may be compromised. But, of these, it seems that the most significant issues to be addressed are to do with *scales* – length and time.

First, the length-scales. In the experiments reported in [82] the colloidal particles have diameters of order 10^{-7} m and the samples comprise crystallites with linear dimensions of order 10^{-5} m. The number

[6] Note the double use of α as both a stacking-fault probability and a structure label.

[7] Ref. [83] shows that some theoretical predictions for the phase diagram of a Lennard-Jones solid depend extremely sensitively on the *fcc-hcp* hard-sphere entropy difference.

of particles in such a crystallite ($N \sim 10^6$) is large compared with those in our simulation, but not sufficiently large that the stacking entropy can be ignored. The obvious implications are qualitatively consistent with the observations reported in [82] which show α values (deduced from Bragg scattering intensities) ranging from 0.5 (signalling essentially random hexagonal close-packing, *rhcp*) through to $\alpha = 0.8$.

The observed spread in α values presumably reflects the issue of timescales. The smallness of the entropy difference (which supplies the kinetic driving force towards the equilibrium state) suggests that the equilibrium behaviour will be observed only in samples which are grown sufficiently slowly and (or) given sufficient time for subsequent annealing [81]. The results of [82] do indeed suggest a correlation between observed α value and the slowness of the growth process. Experiments done in micro-gravity [84], where growth processes are greatly accelerated, yield essentially randomly close-packed crystals.

With these issues in mind, other researchers here at Edinburgh have recently attempted a detailed examination the structural phase behaviour of hard-sphere colloidal systems [85]. These experiments sought a greater understanding of the role of gravity by using two different types of colloidal dispersions; one where the densities of the solvent and the particles were significantly different (and so the effects of gravity were strong) and one where the densities of the two components were very closely matched (referred to here as 'milli-gravity' conditions). The samples were rapidly formed from the dense fluid and contained a wide range of sizes of crystallite, the structures of which were determined using powder-diffraction techniques. When the effect of gravity was weak, no preference for either *fcc* or *hcp* was observed, independent of crystallite size. However, when the gravitational effects were stronger, or if the sample was disturbed (gently shaken, or rhythmically tipped back and forth), the *fcc* structure was found to slowly appear. Unfortunately, in the milli-gravity experiments, the structure was not observed to anneal towards either *fcc* or *hcp*, and so on the basis of the evidence in [85] it is not clear whether the preference for *fcc* in the 'disturbed' systems is because the systems were annealing towards equilibrium, or whether *fcc* is only generated by non-equilibrium shear effects. However, considering the results of both [82] and [85], it seems likely that the shear effects are simply speeding-up the equilibration process, and that *fcc* is indeed the equilibrium structure.

4.12 Discussion

The full agreement between the present work and that of [81] leaves little doubt that the equilibrium entropy difference between the two close-packed structures of hard spheres has finally been established securely and with high precision, at least at one density. Although a small discrepancy with respect to the results of [77] remains, the accord of our close-packed limit results with those established using

pressure difference measurements [70] suggests that the curve in fig. 4.20 provides a relatively complete and trustworthy picture of the density dependence.

These results have shown that the lattice-switch Monte Carlo method for hard-spheres is an accurate and powerful tool for investigating the structural phase behaviour of this system. Since the original publication [1], the LS technique has been applied by a number of different authors [81, 77, 80]. However, if one compares LSMC to its 'closest competitor' (Einstein-crystal integration) on the basis of precision-for-computational-buck there seems to be no clear winner in the hard-sphere studies to date: reference [81] reports calculations using both methods that achieve comparable levels of precision on the basis of comparable computational time. But one should note that the entropy difference ultimately determined is some *four orders of magnitude* smaller than the separate entropies of the two phases, as determined via integration methods. One can see this as a testimony to the care with which the integration procedure has been carried out; or as a strong indicator that more direct approach is called for. There are also two other counts — both somewhat subjective — on which to suggest that the LS approach is superior. First, it seems somewhat illuminating (by comparison with integration methods) to read-off the result for a free energy difference directly from a figure like fig. 4.19 which *shows* what it *means*. Secondly it also seems that LS wins in regard to the transparency of the uncertainties to be attached to its results. The LS error bounds represent purely statistical uncertainties associated with the measurement of the relative weights of two peaks in the $\mathcal{M}$- distribution. The integration-method error bounds have to aggregate the uncertainties associated with different stages of the integration process.

Having successfully constructed a lattice-switch method suitable for the hard-sphere potential, the next step is to broaden the applicability of the technique to include soft potentials. This development forms the basis of chapter 6, where the Lennard-Jones solid will form the focus of our attention.

Chapter 5

Tools of the Trade: Part II

Investigating phase coexistence

For the hard-sphere solid, we have presented evidence to support the claim that for any density (or pressure) the thermodynamically stable phase is *fcc*. As long as the number of particles is reasonably large, there appears to be no density for which the equilibrium phase consists of coexisting *fcc* and *hcp* crystallites. We shall see, however, that this is not the case for a system of Lennard-Jones particles. Chapter 6 will begin with an overview of the evidence which for many years has supported the assertion that; i) at low temperatures *hcp* should be the stable phase over a wide range of pressures, and ii) at higher pressures and temperatures the *fcc* structure should be preferred. Therefore, over some range in the p-T phase diagram, there must lie a boundary along which *fcc* and *hcp* coexist.

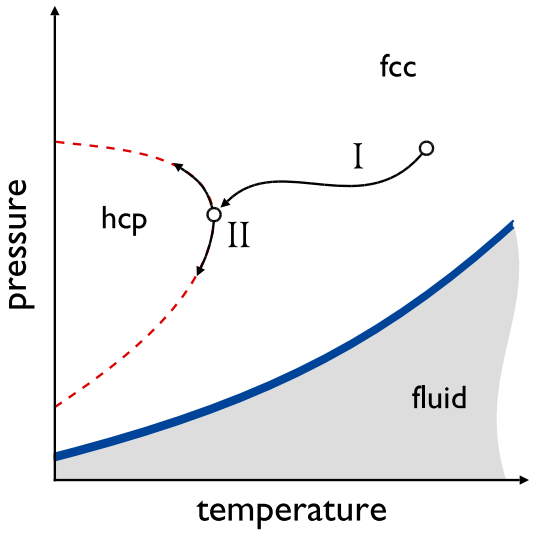


Figure 5.1: Schematic illustration of the two-stage scheme for determining the *fcc–hcp* coexistence curve.

Our task, therefore, demands that we do more than simply determine the free-energy difference between the two structures at one or two state points, as we did for the hard-sphere system. We must attempt to determine the position of the entire coexistence curve, by mapping out state points for which the free-energy difference between the structures is zero. To do this, we break the investigation down into two stages (see fig. 5.1). First, we must track down the location of *some* arbitrary point on the coexistence curve as quickly as possible. Second, we attempt to move along the coexistence curve in order to trace it out in its entirety. Unfortunately, any evaluation of the free-energy difference between *fcc* and *hcp* will take a significant amount of computation time to complete. We must, therefore, attempt to find ways of using the information determined in one simulation (at some specific state point) to extrapolate to other regions of the p-T state-space so that the location of the coexistence curve can be estimated as efficiently as possible. There are a range of techniques available in the literature which are specifically designed to deal with this type of problem. Of these, only histogram reweighting and the Gibbs-Duhem integration method are sufficiently general to warrant detailed examination here. For a broader treatment of coexistence techniques, in a variety of contexts, the reader is directed to ref. [86].

5.1 Stage I: Extrapolating to coexistence

We first assume that, using LSMC, the free-energy difference between *fcc* and *hcp* has been determined for some arbitrary point in the state-space of the Lennard-Jones solid (precisely how this can be achieved will form the core of chapter 6). Unless we have been extremely fortunate, the free-energy difference at this state-point will be significantly different from zero. We should therefore attempt to use the same simulation data to make a prediction as to where the coexistence curve lies by extrapolating toward a different value of the temperature and/or pressure for which the free-energy difference *is* zero. This stage of coexistence-evaluation will be attempted via two techniques, and the quality and efficiency of both methods will be compared in §6.6.2. The first of these is single-histogram extrapolation, which endeavours to use detailed information about the evolution of a simulation to predict what *would* have happened *if* the simulation had been performed at a different temperature and/or pressure. The second method is quite different, and is based on using the results of the simulation in a Newton-Raphson root-finding algorithm. Having explained both of these techniques, we shall examine the ways in which the coexistence curve may be traced out in full.

5.2 Single-histogram extrapolation

The concepts behind single-histogram extrapolation (SHE) have been around for almost as long as the Monte Carlo technique itself, but have only become popular since the 1988 treatment by Ferrenberg & Swendsen [87]. The essential idea is that of *reweighting* the observed microstates from a MC simulation. As the microstate weights depend on the ensemble in which the simulation is performed, we treat the NVT and NPT ensembles separately here.

5.2.1 Canonical SHE

In this case, we wish to extrapolate the results of a simulation performed at one (inverse) temperature (β) to a different temperature (β'). Trivially, the way in which the weight of any microstate ($W(\{\vec{q}\})$) is modified by a shift in temperature from β to β' can be expressed as

$$W(\{\vec{q}\}|\beta') = \frac{W(\{\vec{q}\}|\beta')}{W(\{\vec{q}\}|\beta)} W(\{\vec{q}\}|\beta) \; . \tag{5.1}$$

An MC simulation performed at a temperature β will generate a sequence of $I = 1...N_I$ microstates with weights $W(\{\vec{q}\}_I|\beta)$, and so the canonical average of any observable at a different temperature β' can be estimated as

$$\overline{\mathcal{O}(\beta')} \overset{\text{EB}}{=} \frac{\sum_I \frac{W(\{\vec{q}\}_I|\beta')}{W(\{\vec{q}\}_I|\beta)} \mathcal{O}_I}{\sum_I \frac{W(\{\vec{q}\}_I|\beta')}{W(\{\vec{q}\}_I|\beta)}} \; . \tag{5.2}$$

More explicitly, using the Boltzmann weight in the canonical ensemble (eqn. 2.5) we find,

$$\overline{\mathcal{O}(\beta')} \overset{\text{EB}}{=} \frac{\sum_I \exp\left[-E_I(\beta' - \beta)\right] \mathcal{O}_I}{\sum_I \exp\left[-E_I(\beta' - \beta)\right]} \; . \tag{5.3}$$

An illustrative example of this reweighting process is given in fig. 5.2(a), where the energy of the system is chosen as the observable to be extrapolated.

To apply this technique to a LSMC simulation, we must extend the form supplied above to allow us to deal with multicanonical extended sampling. This requires us to make two assumptions about the nature of the LSMC simulation. Firstly, that it produces some order parameter, $\mathcal{M}$, which measures the 'mismatch' between a pair of candidate structures (based on applying some unspecified lattice-switch transformation to the microscopic particle configuration). Secondly, that this order parameter forms a pathway along which the simulation may be multicanonically biased, and thus that the probability distribution of $\mathcal{M}$ encapsulates the information required to construct a suitable weight function. For example, a canonical LSMC simulation will produce a series of measurements of E and $\mathcal{M}$ (the I'th measurements being E_I, $\mathcal{M}_I$), generated at some chosen temperature β, using a multicanonical weight function $\{\eta\}$. Here, we wish to use this information to predict the required weight function and the free-energy difference between the two structures at some other temperature, β'.

For any observable, the temperature-reweighted extrapolation in the MCMC context can be determined as,

$$\overline{\mathcal{O}(\beta')} \stackrel{\text{EB}}{=} \frac{\sum_I \mathcal{O}_I \exp\left[\eta_I + (\beta - \beta')E_I\right]}{\sum_I \exp\left[\eta_I + (\beta - \beta')E_I\right]} \; . \tag{5.4}$$

The statistical error associated with this estimator can be calculated using the standard block-averaging method described in §3.4. All that has changed is the weighting of each observation. Clearly, if we set $\beta' = \beta$ then we recover

$$\overline{\mathcal{O}(\beta)} \stackrel{\text{EB}}{=} \frac{\sum_I \mathcal{O}_I \exp\left[\eta_I\right]}{\sum_I \exp\left[\eta_I\right]} \; , \tag{5.5}$$

which is our original expression (eqn. 3.36) for recovering an observable from a multicanonically weighted simulation. In order to predict the weight function at some other temperature, we must predict how the entire $P(\mathcal{M})$ curve will change as a function of temperature. To do this, we first use the simulation data to accumulate the following discretized estimator for the (multicanonical) joint distribution of $\mathcal{M}$ and E,

$$P(\mathcal{M}, E|\beta, \{\eta\}) \doteq H_{ij}(\beta, \{\eta\}) \; , \tag{5.6}$$

where i and j index the bins for $\mathcal{M}$ and E values respectively, and $H_{ij}(\beta, \{\eta\})$ represents the number of observations falling in bin ij. The symbol $\doteq$ indicates that we are using a discrete histogram to provide an approximation to a continuous probability distribution, and so the histogram must be normalised in order to recover an estimate of the probability distribution. The multicanonical weight can be folded out using

$$H_{ij}(\beta) = \exp\left[\eta_i\right] H_{ij}(\beta, \{\eta\}) \; , \tag{5.7}$$

yielding an estimation of the joint distribution

$$P(\mathcal{M}, E|\beta) \doteq H_{ij}(\beta) \; . \tag{5.8}$$

The histogram of $\mathcal{M}$ can be determined by summing over all energies for each value of $\mathcal{M}$,

$$P(\mathcal{M}|\beta) \doteq H_i(\beta) \equiv \sum_j H_{ij}(\beta) \; , \tag{5.9}$$

and the change in H_{ij} when extrapolating from β to β' is defined (via eqn. 5.1) as

$$H_{ij}(\beta') \equiv \exp\left[(\beta - \beta')E_j\right] H_{ij}(\beta) \; . \tag{5.10}$$

Therefore we can use

$$H_i(\beta') = \sum_j H_{ij}(\beta') \tag{5.11}$$

$$= \exp\left[\eta_i\right] \sum_j \exp\left[(\beta - \beta')E_j\right] H_{ij}(\beta, \{\eta\}) \; , \tag{5.12}$$

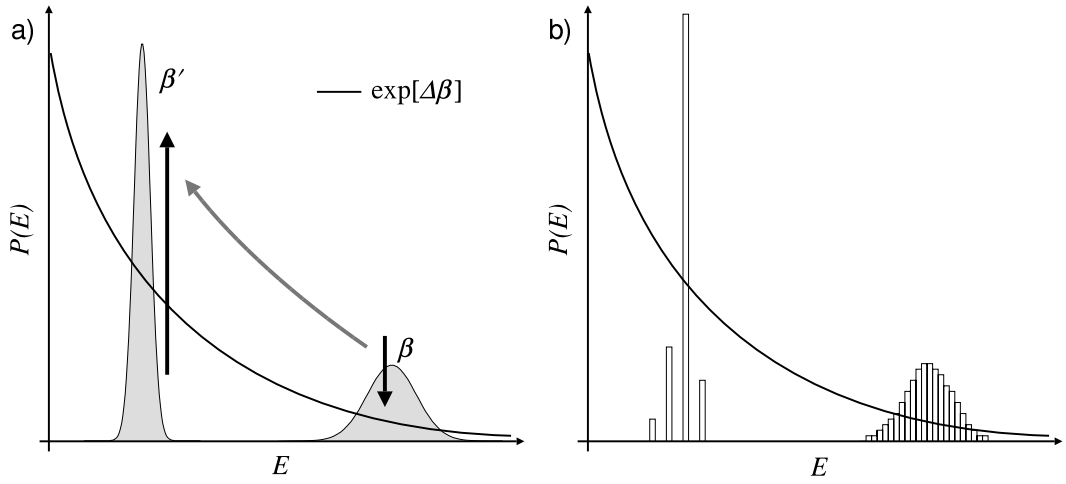


Figure 5.2: Schematic illustration of single-histogram extrapolation. (a) Ideally, given a sufficiently large volume of data at a temperature β, the distribution of the energy for a simulation at β' can be determined by multiplying the distribution $P(E|\beta)$ by $\exp[\Delta\beta]$. This process decreases the statistical weight of energy observations associated with β, and increases the weighting of the energies corresponding to β' (in the tails of the original distribution). (b) However, given that one can only accumulate a finite amount of data in $P(E|\beta)$, the estimate distribution of energy for β' will become (exponentially) worse as the range of the extrapolation ($\Delta\beta$) is increased.

to approximate the $\mathcal{M}$ distribution at some temperature β',

$$P(\mathcal{M}|\beta') \doteq H_i(\beta') \ . \tag{5.13}$$

However, in this form the accuracy of our estimate for $P(\mathcal{M}|\beta')$ will depend on our chosen level of discretization of both E and $\mathcal{M}$. This leads to an undesirable degradation of statistical fidelity, as a consequence of the process of coarse graining over each histogram bin, but this can be rectified. Upon examining the form of equation 5.12, it is clear that any bins for which $H_{ij}(\beta, \{\eta\}) = 0$ will not contribute toward our estimated $H_i(\beta')$, and so can be omitted from the summation. Therefore, if we consider the limit of the energy discretization scale tending toward zero, equation 5.12 can be re-written as

$$H_i(\beta') = \exp\left[\eta_i\right] \sum_{\mathcal{M}_I \sqcup i} \exp\left[(\beta - \beta')E_I\right] \ , \tag{5.14}$$

where the summation is performed over all measurements of $\mathcal{M}_I$ and E_I for which $\mathcal{M}_I$ lies within bin i (denoted as $\mathcal{M}_I \sqcup i$). In this way it is possible to estimate the probability distribution of $\mathcal{M}$ without explicitly discretizing the energy-axis. Eqn. 5.14 can then be used to extrapolate the ratio of the configurational weights of the two structures from the simulation data. If we denote any observations which belong to one of the two phases being compared as $I \in \alpha$ or $I \in \alpha'$, then by separating the data

from each phase we find,

$$\begin{aligned}\mathcal{R}'_{fcc,hcp} &\overset{\text{EB}}{=} \frac{\sum_i H_i(\beta'|\alpha')}{\sum_i H_i(\beta'|\alpha)} \,, \\ &\overset{\text{EB}}{=} \frac{\sum_{I\in\alpha'} \exp\left[\eta_I\right] \exp\left[(\beta-\beta')E_I\right]}{\sum_{I\in\alpha} \exp\left[\eta_I\right] \exp\left[(\beta-\beta')E_I\right]} \,,\end{aligned} \tag{5.15}$$

from which the free-energy difference can be determined,

$$\Delta f = \frac{\Delta F}{N} = \frac{\ln \mathcal{R}'_{fcc,hcp}}{N\beta'} \,. \tag{5.16}$$

Essentially, we have treated the total (MCMC-corrected) 'time' spent in each of the two phases as two separate observables. The ratio of these 'times' is an estimate for the relative probabilities of the phases, and so forms an estimator of the free-energy difference.

5.2.2 Isothermal-isobaric SHE

For the constant pressure ensemble, we can extrapolate to a different pressure (p') as well as a different temperature (β'). The algorithm is essentially the same as outlined above, as all that has changed is the form of the microstate weight. For an extrapolation in temperature and/or pressure, we find

$$\overline{\mathcal{O}(\mathrm{p}',\beta')} \overset{\text{EB}}{=} \frac{\sum_I \mathcal{O}_I \exp\left[\eta_I + (\beta-\beta')(E_I + (\mathrm{p}-\mathrm{p}')V_I)\right]}{\sum_I \exp\left[\eta_I + (\beta-\beta')(E_I + (\mathrm{p}-\mathrm{p}')V_I)\right]} \,, \tag{5.17}$$

where V_I is the Ith measurement of the volume of the system. The Gibbs free-energy difference can be constructed from the ratio of the total number of observations of each structure, just as eqn. 5.15 does in the canonical case.

5.2.3 Application & error analysis

In order to apply the SHE technique, we must record the required information about the microstates a simulation explores. In the NVT case, this means that we periodically measure the energy and the lattice-switch order parameter (E_I & $\mathcal{M}_I$). In the NPT case, we must also record the volume, V_I. This stream of coupled values of the observables can be recorded during the course of a simulation, and then used in the reweighting formulae to extrapolate to different temperatures or pressures. As mentioned above, the reweighting process uses essentially the same as the block-analysis procedure outlined in §3.4, the only modification being the weight associated with each observation.

The frequency with which we take these measurements of $(E, \mathcal{M})$ or $(E, \mathcal{M}, V)$ will be determined by the autocorrelation time of those observables, i.e. the measurement period should be large enough to ensure that the data is not heavily correlated. The block-analysis procedure is of course designed

to eliminate the effects of such correlations on the error estimates, and so it is not strictly necessary to remove the correlated data. However, the sheer volume of data required by the SHE method means that storing the strongly correlated observations would require impractically large amounts of storage space. This scheme allows us to estimate the results of a simulation held at different temperature and/or pressure, using only a tiny fraction of the computer time and space required to perform a full MC simulation. However, the question remains as to how well we might expect the extrapolation to perform.

In principle, it is possible to use the data from a simulation at one state-point to predict the value of an observable at *any* other. However, this is only true if the set of recorded states is very large; for any numerical simulation only a finite set of microstates will be visited and the extrapolation will be limited by this. For example, when using a Boltzmann importance-sampling MC simulation, the exploration of the energy or volume microstates will be exponentially suppressed as we attempts to move away from the region of highest Boltzmann weight. This implies that as we attempt to extrapolate away from the simulation state-point, the 'new' high Boltzmann-weight microstates will be poorly sampled (see fig. 5.2(b)). Consequently, the statistical errors associate with the predicted values of observables will grow rapidly larger the further we attempt to extrapolate. As the size of the system is increased (and the observables of the system become more sharply defined) the range over which SHE produces statistically reliable results becomes smaller. In the thermodynamic limit, this extrapolation technique fails completely, and it is not possible to know *a priori* how far through the state-space the SHE method will allow us to extrapolate successfully.[1] These 'range of reliability' issues of the SHE method will be explored in §6.6.2.

5.3 Newton-Raphson technique

In contrast to the 'data extrapolation' of the SHE method, this technique is based on what may be termed a 'functional extrapolation'. That is, the functional-form of the free-energy difference is approximated using a low-order polynomial, and the roots of this function are used to estimate the points for which the actual free-energy difference is zero. One of the most powerful techniques of this type is the Newton-Raphson root finding method, as lucidly discussed in [88, §9.4]. In its general form, we have a function $y(x)$ for which we wish to determine a root ($y = 0$). If we start by assuming that the function has a root

[1] The range of the SHE extrapolation can be extended by multicanonically weighting the simulation to explore a wide range of macrostates along the extrapolation axis. For example, the work presented in [51] centered around multicanonically weighting an Ising model simulation to explore a wide range of energies, and using SHE reweighting to extract the values of observables over a wide range of temperatures. Here, however, the axis of multicanonical biasing is the $\mathcal{M}$-axis (neither E nor V) and so this cannot be relied upon to assist the extrapolation process.

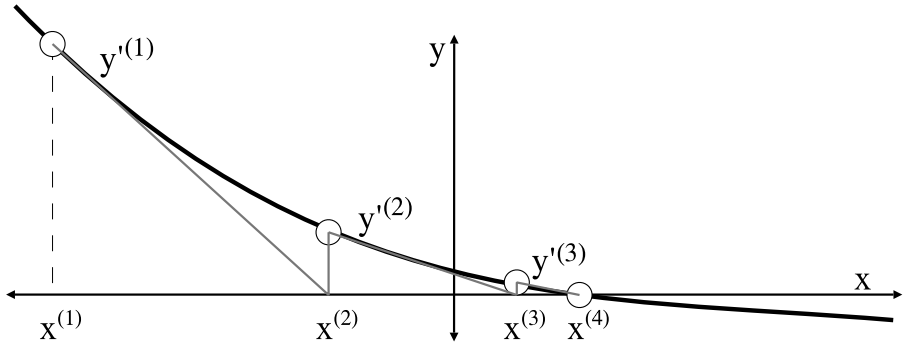


Figure 5.3: An illustration of the Newton-Raphson root finding algorithm. Starting on the far left, we evaluate the function and its gradient at $x^{(1)}$. This information is used to construct a linear approximation to the curve, for which the root ($x^{(2)}$) is trivial to determine. The new value of x is then used to begin the next iteration step. The first four iterations are shown.

not far from some initial guestimate $x = x^{(1)}$, then this guess can be refined via

$$x^{(n+1)} = x^{(n)} - \frac{y(x^{(n)})}{y'(x^{(n)})} \,, \tag{5.18}$$

where $y'(x^{(n)})$ is the gradient $\mathrm{d}y/\mathrm{d}x$ at $x^{(n)}$ (see fig. 5.3). The Newton-Raphson method is powerful (rapidly converging) and accurate if the iterative process is initialised in the neighbourhood of a root. However, in under certain conditions, the iterative process can become unstable or cyclic. For example, if the iteration hits any stationary point in the function ($y'(x^{(n)}) \sim 0$) then the next iteration will "shoot off to outer space" [88]. Should this occur, the only solution is to restart the iteration somewhere else in x, using whatever information we have about the function to direct our choice of $x^{(1)}$.

5.3.1 The NVT ensemble

In the canonical case, we wish to find the root of the Helmholtz free energy difference between two structures,

$$\Delta f(\beta) = \frac{F_{\alpha'}(\beta) - F_{\alpha}(\beta)}{NkT} = -\frac{1}{N} \ln \mathcal{R}_{\alpha',\alpha}(\beta) = 0 \,. \tag{5.19}$$

To extrapolate along the inverse-temperature axis, we take equation 5.18 and set $x \equiv \beta$, $y \equiv \Delta f$ to get

$$\beta^{(n+1)} = \beta^{(n)} - \frac{\Delta f(\beta^{(n)})}{\Delta f'(\beta^{(n)})} \,. \tag{5.20}$$

The derivative of Δf with respect to β can be shown to be equal to the average energy difference between the structures (using eqn. A.3),

$$\begin{aligned}\Delta f'(\beta) &= -\frac{1}{N}\frac{d(\ln R_{\alpha',\alpha}(\beta))}{d\beta} \\ &= -\frac{1}{N}\left[\frac{\mathrm{d}(\ln Z_{\alpha'}(\beta))}{\mathrm{d}\beta} - \frac{\mathrm{d}(\ln Z_{\alpha}(\beta))}{\mathrm{d}\beta}\right] \\ &= \frac{1}{N}\left[\overline{E_{\alpha'}(\beta)} - \overline{E_{\alpha}(\beta)}\right] \\ &= \Delta e\ , \end{aligned} \tag{5.21}$$

where Δe is used as a shorthand for the difference in energy per particle between the two structures (note that for any observable, we employ the convention $\Delta\mathcal{O} = \overline{\mathcal{O}_{\alpha'}} - \overline{\mathcal{O}_{\alpha}}$). Using 5.21, eqn. 5.20 becomes

$$\beta^{(n+1)} = \beta^{(n)} - \frac{\Delta f^{(n)}}{\Delta e^{(n)}}\ . \tag{5.22}$$

To use this, we must use the LSMC simulation data to determine the values of Δf and Δe. The free-energy difference is determined in the usual LSMC manner, from the ratio of the configurational weights

$$\Delta f = -\frac{1}{N}\ln \mathcal{R}_{\alpha',\alpha}\ , \tag{5.23}$$

where

$$\mathcal{R}_{\alpha',\alpha} \overset{\text{EB}}{=} \frac{\sum_{I\in\alpha'}\exp\left[\eta_I\right]}{\sum_{I\in\alpha}\exp\left[\eta_I\right]}\ . \tag{5.24}$$

For the energy difference, we have

$$\Delta e = \frac{1}{N}\left(\overline{E_{\alpha'}} - \overline{E_{\alpha}}\right)\ , \tag{5.25}$$

where

$$\overline{E_{\alpha}} \overset{\text{EB}}{=} \frac{\sum_{I\in\alpha} E_I \exp\left[\eta_I\right]}{\sum_{I\in\alpha}\exp\left[\eta_I\right]}\ . \tag{5.26}$$

5.3.2 Measurement & error analysis

There are two distinct ways in which this calculation can be approached. So far, we have assumed that the values of both observables (Δf & Δe) will be determined from a single lattice-switch simulation run. Alternatively, we can determine the value of the free-energy difference by a lattice-switch simulation, and the value of the energy difference by using two Boltzmann importance-sampling simulations (one for each phase). In both cases, the errors associated with the observables can be calculated using block analysis (§3.4). For the three-simulation approach, the errors associated with the two observables are independent and so can be combined as,

$$\frac{\sigma(\beta^{(n+1)})}{\beta^{(n+1)}} = \sqrt{\left(\frac{\sigma(\Delta f)}{\Delta f}\right)^2 + \left(\frac{\sigma(\Delta e)}{\Delta e}\right)^2}\ , \tag{5.27}$$

where the error on the energy difference is determined as,

$$\sigma(\Delta e) = \sqrt{\sigma(E_{\alpha'})^2 + \sigma(E_\alpha)^2} \ . \tag{5.28}$$

However, in the one-simulation case the estimates are not independent, and so the total error must be calculated as

$$\frac{\sigma(\beta^{(n+1)})}{\beta^{(n+1)}} = \left|\frac{\sigma(\Delta f)}{\Delta f}\right| + \left|\frac{\sigma(\Delta e)}{\Delta e}\right| \ , \tag{5.29}$$

using

$$\sigma(\Delta e) = |\sigma(E_{\alpha'})| + |\sigma(E_\alpha)| \ . \tag{5.30}$$

It should be noted, however, that the decision to combine the errors like this is a consequence of being forced to assume that the energies *may* be *strongly* correlated. This is probably somewhat harsh, and so the final error is likely to be an overestimate. Note also that using a single LS simulation means that the energy difference will be less accurately determined than in a dedicated pair of runs. This is simply because the LSMC simulation will spend a reasonable proportion of its time in the low Boltzmann weight configurations which lie on the inter-phase path, whereas an importance sampling simulation will dedicate all of its time to the high-weight configurations. Thus the errors will certainly be larger in the one-simulation case (per unit simulation time). Having said this, it should be noted that performing *extra* simulations to determine Δe is costly in terms of overall time spent, and so both approaches will be implemented and compared in §6.6.2.

5.3.3 The NPT ensemble

For the NPT ensemble, the temperature extrapolation works in much the same way as in the canonical ensemble (c.f eqn. A.5),

$$\beta^{(n+1)} = \beta^{(n)} - \frac{\Delta g}{\Delta h} \ , \tag{5.31}$$

where Δg is the Gibbs free-energy difference (per particle) between the structures, and Δh is the *enthalpy* difference. This latter property is estimated as

$$h_\alpha \overset{\text{EB}}{=} \frac{1}{N}\langle E + \mathrm{p}V \rangle \ , \tag{5.32}$$

using much the same unbiasing and error analysis procedure as for the energy and volume differences in the canonical case (§5.3.2).

The Newton-Raphson approach can also be used to extrapolate to the transition pressure, beginning with

$$\Delta g(\mathrm{p}) = \frac{G_{\alpha'}(\mathrm{p}) - G_\alpha(\mathrm{p})}{NkT} = -\frac{1}{N}\ln \mathcal{R}(\mathrm{p}) \ . \tag{5.33}$$

We identify $x \equiv \mathrm{p}$ and $y \equiv \Delta g$ in eqn. 5.18, leading to

$$\mathrm{p}^{(n+1)} = \mathrm{p}^{(n)} - \frac{\Delta g(\mathrm{p}^{(n)})}{\Delta g'(\mathrm{p}^{(n)})} \,. \tag{5.34}$$

We now use eqn. A.7 to re-express the gradient of the free-energy in terms of the volume difference per particle,

$$\mathrm{p}^{(n+1)} = \mathrm{p}^{(n)} - \frac{\Delta g}{\beta \Delta v} \,. \tag{5.35}$$

Again, the errors from the measurements of Δg and $\beta \Delta v$ are combined to produce the error for $\mathrm{p}^{(n+1)}$ in almost exactly the same way as the errors in Δf and Δe were combined for the canonical algorithm (§5.3.2). The only difference is the presence of the inverse temperature in the denominator of the gradient.

5.3.4 Convergence

The errors calculated via the Newton-Raphson technique can be used as part of a convergence test: the root-finding algorithm will be deemed to have converged when the predicted $\beta^{(n+1)}$ and/or $\mathrm{p}^{(n+1)}$ is statistically indistinguishable from the simulated temperature and/or pressure. Furthermore, this calculation provides a means by which we can determine the error (in β and/or p) associated with the estimation of the coexistence curve's position.

5.4 Stage II: Tracing the coexistence curve

Having identified at least one point on the coexistence curve, we now wish to map out the rest of the curve as efficiently as possible. Three different methods were compared during the course of this work (see §6.6.3), and these are outlined in the following sections.

5.5 Single-histogram extrapolation (reprise)

As explained in §5.2, the SHE method may be used to hunt down an initial point on the coexistence curve. In much the same way, this technique can be used to explore a range of temperatures and pressures in the locale of a simulated (coexistence) state-point, and thus determine (locally) where the free-energy curve lies. This information can then be used as a starting point for the next LSMC simulation, and to predict a suitable weight function for it. Also, as the major restriction of the SHE method is the limit to how far one can successfully extrapolate across the state-space, this approach may in fact be more

useful here. This is simply because the point to which we wish to extrapolate will probably be closer to the original simulation point than in the case of the initial extrapolation to coexistence.

5.6 Gibbs-Duhem integration

In principle, this relatively recently developed technique [89] allows the direct computation of any coexistence curve. It relies on a simple derivation (specific to a particular ensemble) which shows that once a single point on a coexistence curve has been identified, the whole curve may be traced out using only *single-phase simulations*. The method has been successfully applied to a broad range of systems, and a review of the technique and the problems it has been applied to can be found in [90].

The implementation of Gibbs-Duhem Integration (GDI) used here will follow that presented in [89], and to a lesser extent [90]. A broader treatment can be found in either of those references, and some interesting arguments as to the relationship between GDI and other methods (such as histogram reweighting techniques) can be found in [91].

5.6.1 Theory

We restrict the following derivation to two-phase coexistence of a one-component substance for which simulations are to be performed in the NPT ensemble. For two nearby points on the coexistence curve of two phases α and α' in the p-β plane, we may write

$$\begin{aligned} g_\alpha(\mathrm{p},\beta) &- g_{\alpha'}(\mathrm{p},\beta) = 0 \\ g_\alpha(\mathrm{p}+\delta\mathrm{p},\beta+\delta\beta) &- g_{\alpha'}(\mathrm{p}+\delta\mathrm{p},\beta+\delta\beta) = 0\ . \end{aligned} \tag{5.36}$$

These two equations can be combined to reveal

$$g_\alpha(\mathrm{p}+\delta\mathrm{p},\beta+\delta\beta) - g_\alpha(\mathrm{p},\beta) = g_{\alpha'}(\mathrm{p}+\delta\mathrm{p},\beta+\delta\beta) - g_{\alpha'}(\mathrm{p},\beta)\ , \tag{5.37}$$

and by replacing the small differences in g with the total differentials of g,

$$\delta\mathrm{p}\left(\frac{\delta g_\alpha}{\delta \mathrm{p}}\right) + \delta\beta\left(\frac{\delta g_\alpha}{\delta\beta}\right) = \delta\mathrm{p}\left(\frac{\delta g_{\alpha'}}{\delta \mathrm{p}}\right) + \delta\beta\left(\frac{\delta g_{\alpha'}}{\delta\beta}\right)\ . \tag{5.38}$$

Using eqns. A.5 & A.7, this can be re-expressed in terms of the differences in enthalpy and volume between the phases. Then, in the differential limit ($\delta x \to \mathrm{d}x$),

$$\frac{\mathrm{dp}}{\mathrm{d}\beta} = -\frac{\Delta h}{\beta \Delta v}\ , \tag{5.39}$$

or equivalently,

$$\frac{dT}{dp} = kT\frac{\Delta v}{\Delta h}\ . \tag{5.40}$$

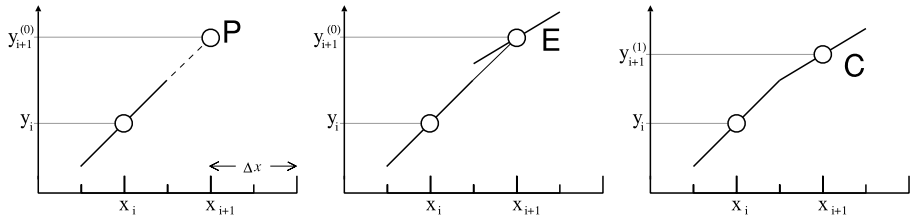


Figure 5.4: A simple illustration of a PECE sequence using the trapezoidal algorithm. First, the algorithm makes a prediction, and the gradient is evaluated at that point. Then, the new gradient is used to correct the original prediction by simply combining the initial and second gradients in a 50:50 ratio. The gradient can then be re-evaluated at this new point, which then determines the next predicted position in the integration.

The gradient of the transition curve at any point along it can thus be estimated by measuring the average enthalpy and volume differences between the phases at that state-point. Therefore, if we have determined the position of a single point on the phase boundary, we can proceed along the phase boundary by measuring the enthalpy and volume differences, constructing the gradient of the curve, and numerically integrating eqn. 5.40. The evaluation of the enthalpy and volume differences can be performed using two separate simulations (one for each phase), and so this approach has the advantage of avoiding the need to bias the simulation into moving between the two phases.

5.6.2 Integration procedure

The numerical integration of eqn. 5.40 will be attempted using two different predictor-corrector algorithms, suitable for any first-order differential equation of the form,

$$\frac{dy}{dx} = \mathcal{G}(x, y) \ . \tag{5.41}$$

There are many different numerical integration schemes which could be applied to this problem (see e.g. [88, Chapter 16]), but predictor-corrector algorithms are the most suitable. This is because the gradient is very expensive to determine, and predictor-corrector schemes allow us to determine a new point on the coexistence curve for every evaluation of $\mathcal{G}(x, y)$. This is in contrast to many other integration schemes (such as the popular Runge-Kutta algorithm [88]), which require more than one evaluation of $\mathcal{G}(x, y)$ for every step of the integration.

A predictor-corrector integration algorithm is primarily defined by two formulae known, unsurprisingly, as the predictor and the corrector. The integration proceeds by stepping along the x-axis by a fixed amount (Δx), evaluating the gradient of the curve every time. The points on the curve which have been determined previously are used to form a prediction of where the next point on the curve lies; the gradient is evaluated at this point, and this information is used to refine (correct) the initial guess (see fig. 5.4). Thus there are three main stages of the calculation, the prediction (P), the evaluation (E) and

the correction (C). Once a PC algorithm has been chosen, there is still some freedom in how the P and C steps should be combined with evaluation steps to get the best result. The "accepted wisdom" [88, §16.7] is to use a predict-evaluate-correct-evaluate scheme (PECE, as in fig. 5.4), unless the gradient is extremely costly to evaluate in which case it should suffice to omit the final evaluation. The work in ref. [89] uses an iterative P(EC)^N algorithm, where running-averages from a pair of long simulation runs are used to refine the corrected value to within some tolerance factor through a series of N EC steps. However, this arguably more accurate scheme[2] requires a large number of gradient estimates, and so can only be applied when the gradient is sufficiently easy to evaluate. Having chosen an algorithm and scheme, we must also consider how the integration process should be initialised. As the P & C formulae often require that a number of previous coexistence points have already been determined, starting such an integration process up in the first place can be awkward.

In this work, two PC methods were compared. The first of these avoids the start-up problem by using the second-order (trapezoidal) PEC scheme (see table 5.1), which only requires a single (LSMC) point in order to proceed. This will be referred to as the 'PEC' scheme. The second is based on the algorithm used in [89], which applies a set of three different algorithms, all of which are presented in table 5.1. The fourth-order Adams scheme is kick-started by using the trapezoidal scheme for the first integration step. This is followed by two Midpoint-algorithm steps, thus producing the four data-points required by the Adams algorithm. This will be called the '3PEC' scheme.

5.6.3 Error analysis

There are three classes of error that the GDI procedure should take into account. These concern (i) the effects of any systematic failure of the integration scheme, (ii) the effect due to the error in the initial coexistence-point evaluation, and most importantly (iii) the effects of the stochastic errors in the gradient (i.e. Δh and Δv) estimation.

i. Systematic errors from the integration procedure

For any PC integration scheme, the step-size Δx must be small enough to allow the algorithm to cope with the rate of change of the gradient. If the step-size is too large then the integration may fail to resolve important features of the curve, and so diverge from the true solution. This source of systematic error can always be remedied by employing a smaller step-size, and can be detected by removing every other data point from the whole set of gradients, and checking that integration procedure gives consistent results.

[2] It should be noted that the authors of Numerical Recipes [88] believe that the time spent iterating the corrector would be better spent using a smaller step-size.

Algorithm	Stage	Formula	# Points
Trapezoid	P	$y_{i+1}^{(0)} = y_i + \Delta x \mathcal{G}_i$	1
	C	$y_{i+1}^{(n+1)} = y_i + \frac{\Delta x}{2}(\mathcal{G}_{i+1}^{(n)} + \mathcal{G}_i)$	1
	σ	$\sigma^2(y_{i+1}^{(n+1)}) = \sigma^2(y_i) + \frac{\Delta x^2}{4}(\sigma^2(\mathcal{G}_{i+1}^{(n)}) + \sigma^2(\mathcal{G}_i))$	as C
Midpoint	P	$y_{i+1}^{(0)} = y_{i-1} + 2\Delta x \mathcal{G}_i$	2
	C	$y_{i+1}^{(n+1)} = y_{i-1} + \frac{\Delta x}{3}(\mathcal{G}_{i+1}^{(n)} + 4\mathcal{G}_i + \mathcal{G}_{i-1})$	2
	σ	$\sigma^2(y_{i+1}^{(n+1)}) = \sigma^2(y_{i-1}) + \frac{\Delta x^2}{9}(\sigma^2(\mathcal{G}_{i+1}^{(n)}) + 16\sigma^2(\mathcal{G}_i) + \sigma^2(\mathcal{G}_{i-1}))$	as C
Adams	P	$y_{i+1}^{(0)} = y_i + \frac{\Delta x}{24}(55\mathcal{G}_i - 59\mathcal{G}_{i-1} + 37\mathcal{G}_{i-2} - 9\mathcal{G}_{i-3})$	4
	C	$y_{i+1}^{(n+1)} = y_i + \frac{\Delta x}{24}(9\mathcal{G}_{i+1}^{(n)} + 19\mathcal{G}_i - 5\mathcal{G}_{i-1} + \mathcal{G}_{i-2})$	3
	σ	$\sigma^2(y_{i+1}^{(n+1)}) = \sigma^2(y_{i-1}) + \frac{\Delta x^2}{576}(81\sigma^2(\mathcal{G}_{i+1}^{(n)}) + 361^2\sigma^2(\mathcal{G}_i) + 25\sigma^2(\mathcal{G}_{i-1}) + \sigma^2(\mathcal{G}_{i-2}))$	as C

Table 5.1: Table of PC algorithm formulae. For each of the three algorithms employed during this work, the predictor (P), corrector (C) and error estimator (σ) formulae are given. The far-right column indicates the number of previous data points which are required by each formula. The error estimator is derived as the geometrical mean of the errors from each term in the corrector formula.

ii. Errors involving the initial state-point

If the estimate for the initial coexistence point is in error (as indeed it always will be) then instead of evaluating the $\Delta g = 0$ curve, we will be evaluating some other curve corresponding to $\Delta g = \epsilon$. Clearly, the rate at which the $\Delta g = \epsilon$ curve diverges from the $\Delta g = 0$ curve will depend on the curvature of the particular P-V-Δg free-energy surface in question. The effect of this initial error is very difficult to estimate beforehand, but there are strategies for dealing with this problem. For example, ref. [90] describes how to determine the preferred direction of integration by measuring how the components of the gradient (the enthalpy and volume) change from step to step. As well as establishing the best direction of integration, this process also allows the error associated with this effect to be estimated as the integration proceeds.

iii. Stochastic errors in the gradient estimation

While there are strategies that can be applied to correct errors of both types (i) and (ii), they are only relevant if this final source of errors has been dealt with. As stated in reference [90], the stochastic error

associated with the evaluation of the gradient is the main source of errors in most GDI procedures, and this is also true for the results to be presented in this thesis. Therefore, this issue must be dealt with in some detail, both in terms of how to estimate the size of these errors, and what the consequences of them may be.

The error analysis proceeds in step with the integration, and the total error is collected at the end of each complete $x \rightarrow x + \Delta x$ step. Only the final correction matters, as the presence of errors in the earlier P/C stages only affect the final correction indirectly (via $\mathcal{G}_{i+1}^{(n)}$) by determining the quality of the convergence. The error-estimation formulae are presented in table 5.1, and are based on the form of the corrector for each scheme, combined with the assumption that all the errors are independent. From the form of these equations, it is clear that the final error over the entire integration is composed of the error on the initial coexistence point ($\sigma(y_0)$), and the cumulative effects of the errors in the gradient estimations ($\sigma(\mathcal{G}_i)$). Of course, we must at least be able to estimate the gradient with an error smaller than itself, otherwise we would not even be able to tell the difference between a positive and negative slope. Here (following [90]), we apply the following 'rule of thumb': the fractional error in the gradient must be at least as small as the desired fractional error in the estimate of the position of the coexistence curve.

As mentioned above, the gradient of the coexistence curve is estimated by using two separate MC simulations to measure the enthalpy and volume of the two different phases. To determine the *error* in the gradient from the two separate sets of simulation data we proceed as follows. For the p-T integration, we map eqn. 5.40 onto 5.41 by stating that $y \equiv T$, $x \equiv \mathrm{p}$ and

$$\mathcal{G}(x, y) \equiv \mathcal{G}(\mathrm{p}, T) \equiv \frac{\Delta v}{\beta^2 \Delta h} \, . \tag{5.42}$$

The accuracy of the gradient estimation depends directly on the fractional errors in the enthalpy and volume differences. These two errors are not independent as they are both determined from the same simulation data, and so the errors must be combined as,

$$\frac{\sigma(\mathcal{G})}{\mathcal{G}} = \frac{\sigma(\Delta v)}{\Delta v} + \frac{\sigma(\Delta h)}{\Delta h} \, . \tag{5.43}$$

These two differences are built from the two separate phase simulation estimates,

$$\frac{\sigma(\Delta v)}{\Delta v} = \frac{\sqrt{\sigma^2(\overline{v_{\alpha'}}) + \sigma^2(\overline{v_\alpha})}}{v_{\alpha'} - v_\alpha} \, , \tag{5.44}$$

and

$$\frac{\sigma(\Delta h)}{\Delta h} = \frac{\sqrt{\sigma^2(\overline{h_{\alpha'}}) + \sigma^2(\overline{h_\alpha})}}{h_{\alpha'} - h_\alpha} \, . \tag{5.45}$$

We require, therefore, that the standard deviation of the volumes and enthalpies is small enough that we can resolve the difference between the structures in a reasonable time. The time required to determine

the gradient to the desired level of accuracy depends strongly on the nature of the two phases being compared, and is very difficult to estimate *a priori*.

If the stochastic errors cannot be ignored, the integration process will begin to fail in a number of ways. If the errors move the integration away from $\Delta g = 0$, then the presence of any adverse curvature of the free-energy surface may lead to a systematic divergence away from coexistence (see point (ii) above). Unfortunately, the error-correction scheme mentioned in (ii) is rendered useless by any significant stochastic errors in $\mathcal{G}$. This is because the fact that the gradient is ill-determined implies that the rate of change of the components of the gradient will be practically impossible to measure.

Also, if the errors are significant on the scale of the gradient itself, then the coexistence curve determined using these gradients will fluctuate about the true curve. The PC integration schemes were intended for use with smooth functions, and so the algorithm (which essentially fits a polynomial to a portion of the coexistence data) may become destabilised by the presence of these fluctuations (c.f. point (i) above).

It should now be clear why the stochastic errors in $\mathcal{G}$ are so important. If we can ignore them, then we can deal with the errors due to the PC algorithm or the curvature of the free-energy surface. However, if the errors in $\mathcal{G}$ are significant, then this makes the more systematic failure of the scheme, via mechanisms (i) and (ii), significantly more likely and impossible to control.

5.7 Predictor-corrector LSMC

This final curve-tracing scheme (developed during this work) is perhaps the simplest of all those considered here. Having performed a LSMC simulation at some point on the coexistence curve, the enthalpy and volume difference between the two structures will be determined from the simulation data. This information is used to estimate the gradient of the coexistence curve (as for GDI), and then used in the trapezoidal predictor equation to predict the transition temperature at $\mathrm{p}' = \mathrm{p} + \Delta\mathrm{p}$,

$$T' = T + \Delta\mathrm{p}\ \mathcal{G}(\mathrm{p}, T)\ . \tag{5.46}$$

A new LSMC simulation is then started at the predicted temperature and pressure, using the weight function from the original simulation. When this simulation is complete, the β-wise Newton-Raphson root finder (§5.3) is then used to correct the predicted transition temperature. This correction process may be applied iteratively if necessary, until the predicted and corrected temperature agree to within their statistical uncertainties. Then we may proceed to determine the next point on the curve. If the gradient is very difficult to measure, then the predictor-step can simply use the current temperature for the prediction, $T' = T$.

The steps in the pressure should be small enough that the weight function does not change too drastically over the course of a single step, as a poor weight function will produce a poor quality estimate of the transition temperature. Apart from this, there is no restriction on the step-size, which may be changed as the simulation proceeds. As long as the correction stage is iterated to convergence, no errors will be propagated along the coexistence curve. Also, by using this method we cannot be misled by any adverse curvature of the free-energy surface, as the iteration will always 'push back' towards $\Delta f = 0$.

5.8 Discussion

The coexistence tools presented in this chapter, combined with the more general MC tools in chapter 3, can now be used to attack the problem of the Lennard-Jones crystal. Once the lattice-switch method has been generalised to deal with soft potentials, the methods presented in this chapter can be compared and contrasted. The best of these techniques will then be used to predict the full *fcc–hcp* coexistence curve of the Lennard-Jones solid.

Chapter 6

The Lennard-Jones Solid

> "If, in some cataclysm, all of scientific knowledge were to be destroyed, and only one sentence passed on to the next generations of creatures, what statement would contain the most information in the fewest words? I believe it is the *atomic hypothesis* (or the atomic fact, or whatever you wish to call it) that *all things are made of atoms - little particles that move around in perpetual motion, attracting each other when they are a little distance apart, but repelling upon being squeezed into one another*." R. P. Feynman [92]

The Lennard-Jones potential is perhaps the simplest mathematical model to fully encapsulate Richard Feynman's interpretation of the atomic hypothesis. However, as for hard-spheres, a problem which may be simply stated is not necessarily simple to solve. In particular, the Lennard-Jones *crystal* has not been satisfactorily dealt with in the literature (see §6.1 below). The simplicity of the model and the incompleteness of the published research record makes the Lennard-Jones solid a perfect arena in which to develop the lattice-switch Monte Carlo technique to deal with soft potentials [3].

6.1 Background

The phase diagram of the Lennard-Jones system is shown in figure 6.1, using data taken from refs. [93, 94]. In fig. 6.1, and throughout this chapter, we use the 'standard' (literature) reduced units for this system: we define a reduced temperature,

$$\frac{1}{\tilde{T}} \equiv \tilde{\beta} \equiv \frac{\epsilon}{kT} \, , \tag{6.1}$$

where the depth of the Lennard-Jones attractive well provides a natural scale on which to consider the thermal energy. The reduced density $\tilde{\rho}$ is defined in terms of the Lennard-Jones 'diameter' σ,

$$\tilde{\rho} \equiv \frac{1}{\tilde{v}} \equiv \rho\sigma^3 \equiv \sigma^3/v \equiv \tilde{\sigma}^3 \tag{6.2}$$

where v is the volume per particle, V/N. Finally, we define the reduced pressure,

$$\tilde{\mathrm{p}} = \frac{\mathrm{p}\sigma^3}{\epsilon} \ . \tag{6.3}$$

In contrast to the hard-sphere work, the temperature has not been folded into the pressure, and so $\tilde{\beta}$ and $\tilde{\mathrm{p}}$ are independent variables. Having chosen ϵ and σ as the natural scales for our dimensionless energies and length-scales, the Lennard-Jones interaction (eqn. 2.4) is re-expressed as follows. We define our scaled coordinates,

$$\vec{\tilde{r}}_i \equiv \frac{\tilde{\sigma}}{\sigma}\vec{r}_i = v^{-\frac{1}{3}}\vec{r}_i \ , \tag{6.4}$$

so that the configurational energy (eqn. 2.4) becomes

$$\tilde{\phi}_{ij} = \frac{\phi_{ij}}{kT} = 4\tilde{\beta}\left[\frac{\tilde{\rho}^4}{\tilde{r}_{ij}^{12}} - \frac{\tilde{\rho}^2}{\tilde{r}_{ij}^{6}}\right] \ . \tag{6.5}$$

The parameters ϵ and σ have been absorbed into the reduced temperature and pressure, and do not appear explicitly anywhere in our calculations.

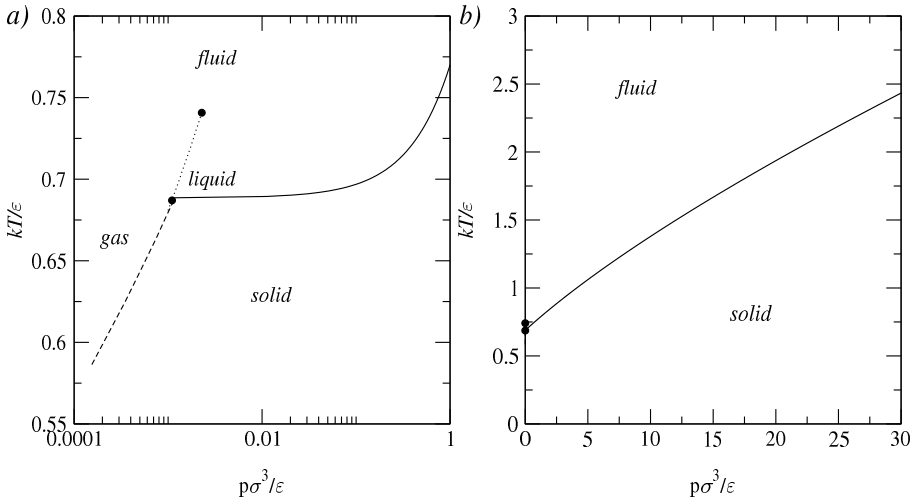


Figure 6.1: The phase diagram of the Lennard-Jones system, as determined in refs. [93, 94]. a) The low-pressure region, using a logarithmic scale for the reduced pressure. The solid-gas, solid-liquid and liquid-gas coexistence curves coincide at the triple-point. The liquid-gas curve terminates at a critical-point, beyond which only a fluid is found. (b) Over larger pressure scales, the solid-fluid coexistence curve dominates the scene.

The articles on which fig. 6.1 is based [93, 94], like many other publications concerning the Lennard-Jones solid, simply *assumed* that *fcc* is the stable crystalline phase. The reason for this common assumption is that in the literature, the Lennard-Jones system is mainly used as a tool with which to investigate

the rare gas solids (i.e. crystalline helium, neon, argon, krypton & xenon). Given the closed outer-electron shells of these elements, which render them largely chemically inert, the interaction between atoms should be approximated reasonably well by the simple, spherically symmetric (electron-electron) repulsion and (van der Waals) attraction encapsulated by the Lennard-Jones potential. As it is these *real* systems which form the true focus of much of the research effort, and as the crystalline phase of these elements is believed to be *fcc* for all temperatures and pressures, it seems quite reasonable to base any theoretical investigations on this empirical observation.

However, if the Lennard-Jones potential is indeed a reasonable model of the interactions between identical rare-gas atoms, then it should be possible to *show* that *fcc* is the stable solid phase of the Lennard-Jones system. Unfortunately, the complete structural phase behaviour of the Lennard-Jones system remains unknown. By considering the role of the ground-state energies alone, it would appear that *hcp* is the stable phase (see §6.3). However, a full *quantum-mechanical* calculation of the Gibbs free-energies (or at least, their differences for the candidate phases) is the only way to settle this issue in the realm of non-zero temperature, and this is by no means a trivial task. This problem of prediction and its consequences for the rare gas solids, as of 1976, is explored in detail in Niebel & Venables' "The Crystal Structure Problem" [95], and we shall return to this issue in §6.7. The work presented here will instead concentrate on determining the phase behaviour of the classical Lennard-Jones solid, and so attempt to clarify the differences between the model's behaviour and that of the rare gases themselves. It may seem surprising that this 'simple' system has not already been exhaustively investigated, but as we shall see during the course of this chapter, this problem is extremely difficult to solve *without recourse to some uncontrolled approximation*. As far we are aware *none* of the many authors who have contributed to the large volume of literature on this topic have successfully determined a full phase-diagram for the solid-state of 'Lennard-Jonesium' *without* being forced into making some uncontrolled approximation.

The earliest work on the Lennard-Jones solid was based on determining the ground-state energies of the candidate structures. A detailed exploration of the ground-state properties is presented in §6.3, while a brief historical overview is given here. In 1925, Lennard-Jones & Ingham [96] compared the *sc*, *bcc* and *fcc* structures, and found *fcc* to have the lowest energy. However, in 1952 Kihara & Koba [97] determined the ground-state energies of the *fcc* and *hcp* structures, and found that for low densities ($\tilde{\rho} \lesssim 2.1$) *hcp* is the preferred phase, and that *fcc* is only favourable above this density. This result has since been confirmed by various authors, to greater levels of accuracy (see [95, §1.2] and references therein). Despite their simplicity, these calculations give reasonably accurate predictions for the measured (low temperature) density and cohesive energy of at least the heavier rare gas solids [9, p. 401].

Later publications involved more sophisticated calculations, based on attempting to determine the non-

zero temperature phase behaviour of the Lennard-Jones solid via the harmonic approximation (described and utilised here in §6.4), which is exact in the limit of low temperature.[1] While this approximation produces an exactly soluble system of equations, performing this calculation is difficult without access to significant computational resources. For example, the 1954 Barron & Domb paper [98] produced a phase diagram for the classical harmonic Lennard-Jones system, but only by recourse to the method of long waves. This partial solution within the harmonic approximation should be valid for low enough temperatures (i.e. below the temperature range for which the harmonic approximation itself is valid), but as with the harmonic approximation itself the question of "how low?" cannot be answered readily without attempting to work under some less severe level of approximation. In 1971, however, a more accurate estimate of the harmonic free-energy difference was given (at the ground-state *densities* alone, see §6.3) by Salsburg and Huckaby [99]. This calculation only included first- and second-neighbour interactions, without examining the role of more distant interactions. However, this calculation does give a prediction for the thermodynamic limit, having taken finite-size effects into account. These early publications all agree that the harmonic contribution to the free-energy favours *fcc* over *hcp*, and this result can be used to predict an approximate *fcc–hcp* transition temperature (although this calculation was not carried out in [99]). Niebel & Venables [95] argued that this temperature is above the sublimation point for all the lightest (less 'classical') rare gas solids, but this statement can only be true at zero pressure, and furthermore their approximate calculations were not based on the most accurate free-energy differences available at the time. Afterwards, the attempts to determine the structural phase behaviour of the rare gas solids, by considering minor changes to the model [100, 101, 95, 102] or considering more subtle effects than that of the bulk free-energy, such as the surface energy [103, 95], have tended to take precedence over any determination of the full phase diagram of the Lennard-Jones system itself.

However, in more recent years, a number of authors have attempted to deal with the *fcc–hcp* phase behaviour of 'Lennard-Jonesium'. Parrinello & Rahman [20] used an MD simulation to show that the Lennard-Jones solid preferred a close-packed structure (with stacking faults) to the body-centered cubic lattice for one point in the phase diagram. However, this calculation was not sensitive enough to distinguish between the different possible stacking patterns. Later, Rahman & Jacucci [21] predicted a region of *bcc* stability at high temperatures and pressures in the region of the melting curve (see §6.8 for a more detailed discussion). The 1984 publication of Galashev [104] did include a direct comparison of *fcc* and *hcp* (in an MD simulation) and found that (at low pressure) the temperature of the *hcp* to *fcc* transition is approximately $kT/\epsilon = 0.4$ (although unfortunately, no error estimate was supplied). The works of Jackson & Swol [105] and Choi *et al.* [106, 83] all attempted to produce phase diagrams

[1] Although the approximation is uncontrolled in the sense that the range of temperatures for which it is valid cannot be known *a priori*.

for the *fcc–hcp* behaviour of the Lennard-Jones solid, but these calculations were hampered by uncontrolled approximations. In the case of ref. [83], the authors used a high-temperature perturbation theory, dependent upon the (at that time) inaccurately measured value of the hard-sphere entropy difference at close-packing. However, employing the more recent, more accurate values of the entropy difference (as given in §4.10.7), their results clearly indicated that the predicted transition temperature is so low that it lies beyond the range of accuracy of the perturbation theory. In the case of refs. [105] and [106], the phase-diagram calculations are dominated by the method of truncation of the potential (this will be explored in more detail in §6.3). The high-temperature perturbation theory [83] mentioned above has, however, been used successfully by Choi *et al.* [107] to show that *fcc* is the preferred structure along the solid-liquid melting line. Most recently, Somasi *et al.* [108] compared the *fcc* and *hcp* structures of the Lennard-Jones solid (at one state-point) using a new integration method. Their qualitative result is that the *fcc* structure is stable at $kT/\epsilon = 0.4$ ($\mathrm{p} = 0$), but their quantitative result for the *fcc–hcp* free-energy difference may not be representative of the true behaviour of the Lennard-Jones solid (again, see §6.3).

To summarise, the classical Lennard-Jones system has been found to prefer the *hcp* structure at low temperatures and pressures, but *fcc* has been identified as the stable phase for the crystal both near the melting curve and at high densities ($\tilde{\rho} > 2.1$). The position of the *fcc–hcp* phase transition has been identified as lying near $kT/\epsilon = 0.4$ at zero pressure, but the overall phase boundary has not been traced out in any detail. A number of authors have attempted to determine the full *fcc–hcp* phase diagram, but their results have been inconsistent and inconclusive.

6.2 Truncation schemes

It is not feasible to simulate a system of Lennard-Jones particles using the full $N \times N$ configurational energy for all but the smallest system sizes. The potential *must* be truncated, but the question remains as to how best to perform this truncation in order to minimise its effect upon the free-energies of interest. To this end, two different styles of truncation were compared during this work, referred to as the 'fixed' and 'scaled' truncation schemes. The fixed truncation is defined in units of the Lennard-Jones parameter σ, such that $r_c = \tilde{r}_c\sigma$. In terms of the pair potential, this corresponds to slicing off the tail of the interaction at a fixed point on the Lennard-Jones curve,

$$\phi_f(r_{ij}) = \phi(r_{ij}) - \phi(r_c) \qquad \text{if} \quad r_{ij} < r_c, \tag{6.6}$$

$$= 0 \qquad \text{otherwise.} \tag{6.7}$$

Note that the potential has also been shifted, to suppress a discontinuous change in the energy from occurring as particles pass through the cut-off value.[2] This can occur stochastically, as the crystal evolves, or systematically, as the control parameters are changed, and can introduce unphysical discontinuities into the free-energy. The second truncation scheme employed in this work has more in common with the work performed on the hard-sphere system. At the outset, pairs of particles are identified as either interacting or not, according to the separation of the lattice sites with which they are associated. This style of cut-off is written as $r_c = \tilde{r}_c r_{nn}$, where r_{nn} is the equilibrium nearest-neighbour separation. Thus, a cut-off of $1.1r_{nn}$ will *always* include *only* 1st nearest-neighbour interactions. Similarly, $r_c = 1.5r_{nn}$ will, for *fcc* or *hcp*, include just 1st & 2nd neighbour interactions, as for these lattices the 2nd nearest-neighbours lie at a distance of $\sqrt{2}r_{nn}$ but the 3rd nearest-neighbours lie beyond $1.5r_{nn}$.[3] The interaction is modified such that

$$\phi_s(r_{ij}) = \phi(r_{ij}) \qquad \text{if} \quad R_{ij} < r_c, \tag{6.8}$$

$$= 0 \qquad \text{otherwise}, \tag{6.9}$$

where $R_{ij} = |\vec{R}_j - \vec{R}_i|$ is the separation between lattice sites i and j. The potential-shift used in eqn. 6.7 is not necessary here, as this fixed-neighbour interaction does not allow particles to 'pass through' the cut-off, and so the associated discontinuities in the energy cannot occur.

The major difference between the two truncation schemes can be illustrated by considering the way the interactions are altered as the density is changed. As shown in figure 6.2, a cut-off in units of σ corresponds to a *fixed length*, so that the number of particle-particle interactions can change with the density. In contrast, a cut-off expressed in terms of r_{nn} will always *scale* with the density, so that the number of interactions remains fixed.

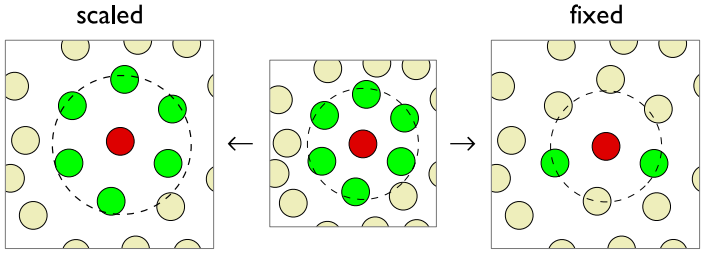


Figure 6.2: Comparison of the fixed-length and scaled truncations under a volume expansion. In the scaled case, the number of interactions remains fixed, whereas for a fixed-length cut-off a sufficiently large expansion can drastically reduce the number of interactions which are included in the calculation.

[2] The shift suppresses the first-order step-discontinuity in the potential, but higher-order discontinuities remain.

[3] The number of interactions within a given cut-off is of course dependent upon the structures involved, but for *fcc* and *hcp* the number (and separation) of 1st & 2nd nearest-neighbours is the same.

The use of a fixed cut-off (specifically, $r_c = 2.5\sigma$) would be consistent with the traditional approach to the fluid phase, and would amount to a reproduction of the work presented in refs. [106, 105] by differing means. This style of truncation has the drawback that the number of neighbour-interactions will increase with the density; over the range of interest ($\tilde{\rho} \sim 1.0...2.0$), the number of neighbours varies from ∼80 to ∼150, and so the CPU time required for each Monte Carlo sweep will double with $\tilde{\rho}$. The scaled cut-off does not have this problem, as the number of interactions is fixed.

However, as we are interested in the behaviour of the 'true' (fully-interacting, large N) Lennard-Jones system, all that should concern us is which of these two methods approximates the $r_c \to \infty$ limit in the most computationally efficient manner. To answer this question, we examine the effect that truncating the potential has on the ground-state and the harmonic free-energies of the Lennard-Jones crystal. These calculations require relatively little computational effort, and allow us to determine how best to truncate the potential before any attempt is made to apply the more computationally intensive lattice-switch Monte Carlo method.

6.3 Ground-state energies & lattice-summation

The phase behaviour of the Lennard-Jones solid at $T = 0$ can be relatively easily determined by calculating the ground-state energies of the candidate lattices. In this regime, the equilibrium structure is simply that which yields the lowest overall energy. For any (stable) crystal structure, the classical ground-state configurations (the lowest configurational-energy microstates) are the perfect-lattice arrangements, where all particles lie precisely upon their lattice sites.[4] For most crystal structures, being in this unique microstate means that each particle 'sees' exactly the same arrangement of neighbouring particles, allowing us to reduce the N^2 configurational energy sum (eqn. 6.5) to N identical N-particle summations.[5]

$$\frac{E_{gs}}{\epsilon} = 2N \left[(\rho^\star)^4 \sum_j \frac{1}{(R^\star_{ij})^{12}} - (\rho^\star)^2 \sum_j \frac{1}{(R^\star_{ij})^6} \right] , \tag{6.10}$$

where each of the summations correspond to a sum over all lattice-sites ($\{\vec{R}\}^\alpha$) neighbouring some given (arbitrary) site, i. Note that both sums are multiplied by $^1/_2N$ in order to avoid double-counting the pair-wise interactions. As is usual in the literature, the lattice-summations have been carried out using a different dimensionless unit system, denoted $\rho^\star$ and $R^\star_{ij}$, where all length-scales are expressed

[4] This can be *shown* to be the case via the harmonic approximation (see §6.4).

[5] While this is automatically true for a Bravais lattice, it need not be the case for any given periodic arrangement (i.e. lattice plus non-trivial basis). However, for the *hcp* structure, the two possible neighbour environments are trivially related by symmetry (one being a reflection of the other) and so, for the structures considered here, the derivation stands.

in units of the nearest-neighbour separation. We identify the lattice-sums as

$$A_n(\alpha) = \sum_j \frac{1}{(R^\star_{ij})^n} \ , \tag{6.11}$$

allowing the ground-state energy (per particle) to be expressed as,

$$\tilde{e}_{gs}(\alpha) = \frac{E_{gs}}{N\epsilon}(\alpha) = 2\left[(\rho^\star)^4 A_{12}(\alpha) - (\rho^\star)^2 A_6(\alpha)\right] . \tag{6.12}$$

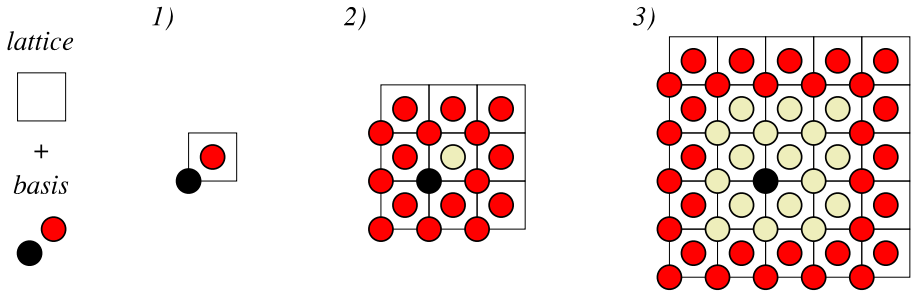


Figure 6.3: Schematic illustration of the lattice-summation calculation for a 2-dimensional square lattice with a simple two-particle basis (far left). Step 1) The $1/r^n$ lattice-sum of a single cell is calculated. Step 2) More cells are added to the system on all sides (in this case, eight), and the new total for the lattice-summation is determined. Step 3) This process is repeated until the lattice-sum (A_n) has converged to a within prescribed tolerance level.

The required lattice-sums (A_{12} & A_6) have been determined for a number of different structures, using an algorithm designed to include more and more distant neighbours until the summation has converged to within some specified level of accuracy. The algorithm is illustrated schematically (for a simple 2-dimensional square lattice) in fig. 6.3, and proceeds as follows. Each structure is specified in terms of a lattice and basis (a 4-site basis in a cubic cell for *fcc*, and a 2-site basis in the usual *hcp* rhomboidal prism), such that the nearest-neighbour separation is equal to unity. Starting with a single cell, the values of $(r^\star)^{-n}$ (for $n = 6$ and $n = 12$) were determined using the distance between a chosen reference site (at the origin) and the neighbouring site-positions (fig. 6.3(1)). The sum of these interactions is determined, and then another layer of lattice-cells is added to the system on all sides (fig. 6.3(2)). The new total for the lattice summation is determined by adding in the values of $(r^\star)^{-n}$ from each of the new cells, and this process is repeated (e.g. fig. 6.3(3)) until the lattice-sum has converged to within $\pm 1 \times 10^{-9}$. This high level of accuracy required the inclusion of $\sim 1 \times 10^8$ and $\sim 1 \times 10^4$ lattice-sites for the $n = 6$ and $n = 12$ summations respectively. The values of the lattice-summations determined in this way (shown in table 6.1) were consistent with (although significantly more accurate than) those published in [9, 96], and the *differences* between the lattice-sums for the *fcc* and *hcp* structures are in agreement with the high-accuracy results of ref. [98]. Table 6.1 also gives $\rho^\star$ in terms of $\tilde{\rho}$, which depends upon the packing fraction and is thus dependent upon the structure.

Structure	$\rho^\star$	$N_n(1)$	Present Work		Literature	
			A_{12}	A_6	A_{12}	A_6
fcc	$\tilde{\rho}/\sqrt{2}$	12	12.131880196	14.453920885	12.13188	14.45392 [96]
hcp	$\tilde{\rho}/\sqrt{2}$	12	12.132293768	14.454897093	12.13229	14.45489 [97]
bcc	$\sqrt{2}\tilde{\rho}$	8	9.114183267	12.253667715	9.11418	12.2533 [96]
sc	$2\tilde{\rho}$	6	6.202149044	8.401923843	6.2021	8.4019 [96]
diamond	$4\tilde{\rho}$	4	4.038904712	5.116771509	-	-

Table 6.1: Comparison of lattice sums, along with the $\rho^\star$ to $\tilde{\rho}$ conversion factor and number of first nearest-neighbours ($N_n(1)$), for a range of crystalline structures. Results of the present work are accurate to within 1×10^{-9}. The literature results were taken from refs. [96] & [97]; no literature results were found for the *diamond* structure. Note that the lattice-sums are all of the order of (but greater than) the number of nearest neighbours for each structure, as one would expect for unit nearest-neighbour separation.

The ground-state energies of the different structures are compared graphically (as a function of density) in figure 6.4. While at low densities, the lowest-energy lattices are in fact the simple-cubic and diamond structures, this actually corresponds (at $T = 0$) to an experimentally inaccessible region of the phase diagram (that of negative pressure). For zero pressure (at $T = 0$), the equilibrium density and structure of the solid phase can be determined by minimising the ground-state energy of the system with respect to the structure and the density (corresponding to the lowest minimum of the set of curves shown in fig 6.4). At higher pressures, the equilibrium density will increase, and the stable phase will be determined by the lowest energy configuration for that pressure. Clearly, the ground-state structure in the non-negative pressure domain is either *fcc* or *hcp*, but the difference between these two curves cannot be shown on the scale of fig. 6.4(a). However, at zero pressure, the density and energy of the structures can be determined by differentiating eqn. 6.12 with respect to the density and setting it equal to zero, yielding

$$\tilde{\rho}_{min}(\alpha) = \sqrt{2}\rho^\star_{min}(\alpha) = \sqrt{\frac{A_6(\alpha)}{A_{12}(\alpha)}}\,. \tag{6.13}$$

Using this expression we find the *fcc* minimum at $\tilde{\rho}_{min}(fcc) = 1.091512667$, with an energy of $\tilde{e}_{gs}(fcc) = -8.610200$, while the (lower-energy) *hcp* minimum lies at $\tilde{\rho}_{min}(hcp) = 1.091530921$, $\tilde{e}_{gs}(fcc) = -8.611070$. The differences in the densities and energies of *fcc* and *hcp* are very small, at $\tilde{\rho}_{min}(fcc) - \tilde{\rho}_{min}(hcp) = -0.000018254$ & $\tilde{e}_{gs}(fcc) - \tilde{e}_{gs}(hcp) = 0.00086955$; five orders of magnitude separate these differences from the corresponding single-phase values. The computed minima are in complete agreement with those published in refs. [109, 110].

Clearly, at least at low temperatures, the structural phase behaviour of the Lennard-Jones solid will be dominated by the *fcc* and *hcp* phases. To compare these structures more directly, we consider the ground-state energy *difference* between them. This can be expressed directly in terms of the lattice

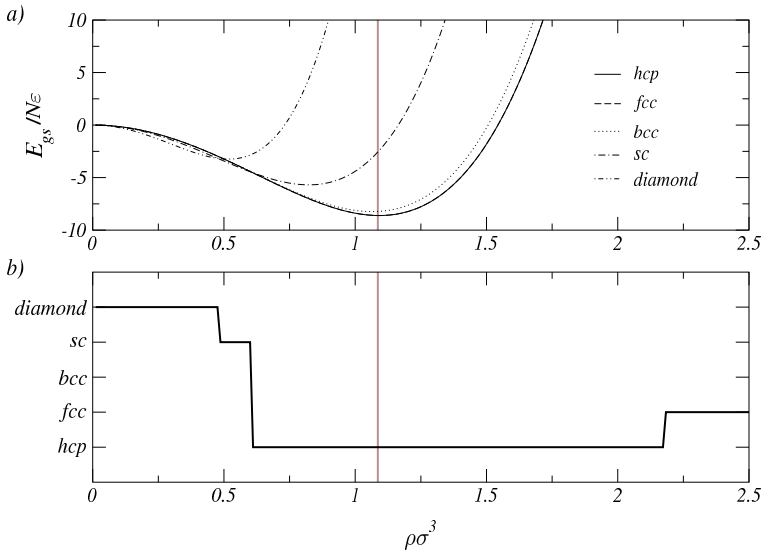


Figure 6.4: (a) Ground-state energy plotted against density for the various structures. The *fcc* and *hcp* curves are indistinguishable on this scale. (b) The graph indicates which structure has the lowest energy at a given density. For both (a) and (b), the grey vertical line indicates the approximate position of the minimum physically accessible ($\tilde{\mathrm{p}} = 0$) density at $T = 0$. Below this density, the (very) low temperature crystal is thermodynamically unstable with respect to fracture.

sums,

$$\begin{aligned} \Delta\tilde{e}_{gs} &= \tilde{e}_{gs}(fcc) - \tilde{e}_{gs}(hcp) \ , \\ &= 2\left[(\rho^\star)^4\Delta A_{12} - (\rho^\star)^2\Delta A_6\right] \ , \end{aligned} \tag{6.14}$$

where

$$\Delta A_n = A_n(fcc) - A_n(hcp) \ . \tag{6.15}$$

Figure 6.5 plots Δe_{gs} as a function of reduced density, showing that while *hcp* has the lowest energy at low density, the preferred phase changes to *fcc* as the density is increased. The exact position of this cross-over can be determined from equation 6.14 by setting $\Delta e_{gs} = 0$, yielding

$$\frac{(\rho_c{}^\star)^4}{(\rho_t{}^\star)^2} = \frac{\Delta A_6}{\Delta A_{12}} \ , \tag{6.16}$$

which in turn gives,

$$\tilde{\rho}_t = 2.172750635 \ . \tag{6.17}$$

This prediction is consistent with the value of $\tilde{\rho} = 2.18$ from ref. [83], as is the form of the exact ground-state energy shown in fig. 6.5. However, these results were *not* found to be in agreement with the results from a number of different implementations of the Lennard-Jones system found in the literature [106, 105]. On closer examination, we have found that the apparent disagreement between the results

presented in these publications and those presented here can be understood by examining the role of the truncation of the potential *alone*.

While (effectively) no truncation has been applied to our ground-state energy results, refs. [106] and [105] both employed a fixed truncation of $r_c = 2.5\sigma$. As well as our 'exact' result, figure 6.5 shows the ground-state energy difference calculated using the same fixed-length cut-off ($r_c = 2.5\sigma$) mentioned above, and also a *scaled* cut-off of $r_c = 2.5r_{nn}$. In the truncated case, the lattice-sum algorithm is essentially the same as before, but now only neighbours within a distance r_c of the origin are included in the summation. Note that these two curves are not only very different in shape to the 'exact' curve, but also that the energies involved are an order of magnitude larger. The results for the 2.5σ truncation show that as the density is increased, this system will undergo three separate structural phase transitions, switching between the *fcc* and *hcp* crystals. This is consistent with the complex and jagged phase diagrams given in refs. [106, 105], but completely unrepresentative of the true ground-state behaviour. The scaled $2.5r_{nn}$ truncation fares little better, as for this interaction there is no ground-state transition at all; the favoured phase is always *fcc*.

In short, our investigations lead us to the conclusion that a large number of interactions (corresponding to a system of at least $N \sim 10^3$ particles) must be included in order to yield low temperature phase behaviour in reasonable agreement with the effectively un-truncated result. This strong dependence on system size may have been overlooked by other authors, and is the probable cause of much of the disagreement between different publications. For example, while Somasi *et al.* [108] focused most of their attention upon relatively large systems ($N \sim 3 \times 10^3$) using a large $\mathcal{O}(N^2)$ truncation, their final result for the *fcc–hcp* free-energy difference was based on a finite-size extrapolation for which one of the two system sizes employed (for each structure) was rather small ($N \sim 3 \times 10^2$). As they employed only two system sizes, it was impossible for them to test the validity of their finite-size analysis. Furthermore, they give no indication of the crucial details of their calculations (e.g. the size and orientation of their lattices, or the truncation scheme) upon which the ground-state energy is so strongly dependent. It was, therefore, impossible to carry out any independent quantitative test of their results during the course of this work.

The primary conclusion to be drawn from all of these results is simple. If we wish to determine the phase behaviour of a system which is in any way representative of the Lennard-Jones model, the ground-state energies *must* be dealt with in their 'exact' (un-truncated) form. For this reason, the truncation of the potential will *only* be applied to the *excitation* energy, over and above the ground-state which will be dealt with 'exactly'. To do this, the configurational energy is modified such that,

$$E(\{\vec{r}\}) = \mathcal{E}(\{\vec{r}\}) + E_{gs} \ , \tag{6.18}$$

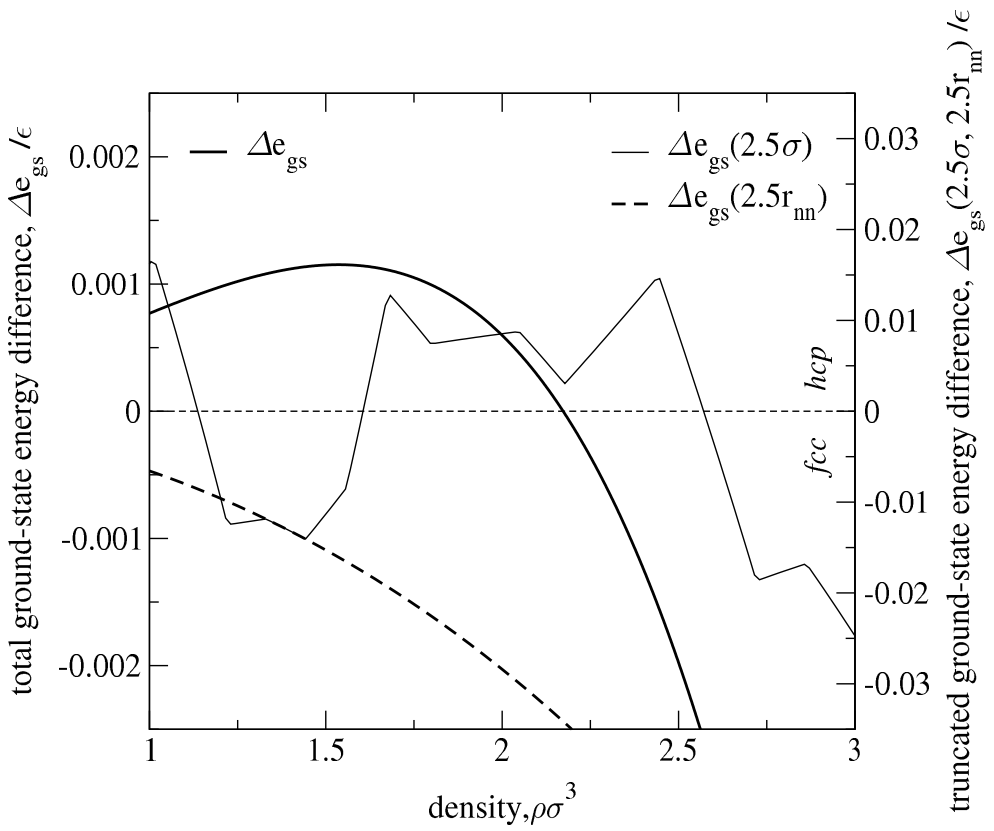


Figure 6.5: The *fcc-hcp* ground-state energy difference as a function of reduced density, calculated using an effectively infinite cut-off (thick line), and using two truncated potentials: a 'fixed-cut and shift' truncation at 2.5σ (thin line), and a 'scaled' truncation at $2.5r_{nn}$ (dashed line). The cut and shift truncation (as used in [106]) causes large fluctuations in the ground-state energy as the density is changed due to the changes in the number of particles within the chosen range of interaction. Note that the vertical scale of the plots for both of the truncated-interaction systems (i.e. the range of the right-hand energy-axis) is an order of magnitude larger than that of the non-truncated system.

where the ground-state energy is calculated from eqn. 6.14, and where the excitation energy $\mathcal{E}$ is calculated over a *restricted* number (N_n) of neighbouring sites within the cut-off radius,

$$\mathcal{E}(\{\vec{r}\}) = \sum_{\langle ij \rangle}^{N_n} \phi_c(r_{ij}) \ . \tag{6.19}$$

The ground-state energy of the *truncated* system is removed from this configurational sum using a modified interaction potential. This modification simply involves subtracting the contribution to the energy associated with the ground-state configuration,

$$\phi_c(r_{ij}) = \phi(r_{ij}) - \phi(R_{ij}) \qquad \text{if} \quad r_{ij} < r_c, \tag{6.20}$$

$$= 0 \qquad \text{otherwise.} \tag{6.21}$$

The manner in which the potential is truncated (either $r_c = X\sigma$ or Xr_{nn}) can now only affect the evaluation of the *temperature-dependent* contribution to the free-energy, whereas the ground-state energies

will always remain exact. However, in order to be able to choose a suitable (fixed or scaled) truncation, one must make *some* estimate of the effect that a restricted interaction has upon the thermal fluctuations.

6.4 The harmonic approximation

As described briefly in §2.4.1, the harmonic approximation provides an exactly soluble system of equations from which the free-energy of a crystal can be determined. Furthermore, this approximation becomes exact in the limit of low temperature, and so provides a means of testing the results of our LSMC simulation. The method is well established, and so only an outline of the technique will be given here; a fuller treatment of the technical details of the harmonic approximation can be found in [111].

The first step is to decompose the positions of the particles ($\{\vec{r}\}$) into lattice sites ($\{\vec{R}\}$) and displacements from those sites ($\{\vec{u}\}$), as described in §4.3 in the context of the lattice switch. The configurational energy of the system is then expanded to second-order in the displacements,

$$E(\{\vec{u}\}) = E_{gs}(\alpha) + \sum_{\langle i,j \rangle} \phi'(R_{ij}) \vec{\eta}_{ij} . \vec{u}_{ij} + \frac{1}{2} \sum_{\langle ij \rangle, v, w} K_{ij}^{vw} u_i^v u_j^w \ , \tag{6.22}$$

where $\vec{\eta}_{ij}$ denotes a site-site unit vector (from site i to j),

$$\vec{\eta}_{ij} = \frac{\vec{R}_{ij}}{R_{ij}} \ , \tag{6.23}$$

and where $u_i^v \ u_j^w$ are the Cartesian components (in the v and w directions) of the displacements of the particles associated with sites i and j; i.e. the second summation is carried out over all combinations of Cartesian directions for all distinct pairs of particles. The matrix K (known as the 'dynamical matrix') is defined as [111]

$$K_{ij}^{vw} \equiv \left[\frac{\phi'(R_{ij})}{R_{ij}} - \phi''(R_{ij}) \right] \eta_{ij}^v \eta_{ij}^w - \frac{\phi'(R_{ij})}{R_{ij}} \delta_{vw} \qquad i \neq j \tag{6.24}$$

$$K_{ii}^{vw} \equiv -\sum_{j \neq i} K_{ij}^{vw} \qquad \text{otherwise.} \tag{6.25}$$

The lowest (zero) order term in eqn. 6.22 describes the energy of the system when all displacements are zero, i.e. the ground-state energy (as calculated in §6.3). The first-order term determines the mean force acting upon a particle when it is vanishingly close to its site, which must be zero for any stable crystalline arrangement. The final, second-order term controls the restoring force on any particle that moves a small distance from its site, and it is this term which describes the dynamics of the particles as harmonic springs. Thus, it is the dynamical matrix which determines the harmonic free-energy (f_h), and when combined with the ground-state energy (Δe_{gs}) yields the total free-energy of the crystal within the harmonic approximation:

$$f(\alpha, \tilde{\rho}) \approx e_{gs}(\alpha, \tilde{\rho}) + f_h(\alpha, \tilde{\rho}) \ . \tag{6.26}$$

It can be shown [111] that (as long as the linear term in eqn. 6.22 is indeed zero), the harmonic contribution to the free-energy density (f_h) of a phase α, at a constant density of $\tilde{\rho}$, can be calculated as

$$\frac{f_h(\alpha,\tilde{\rho})}{kT} = \tilde{f}_h(\alpha,\tilde{\rho}) = \frac{1}{2N}\sum_{d=1}^{3(N-1)} \ln\lambda_d(\alpha,\tilde{\rho}) \ , \tag{6.27}$$

where λ_d $(d = 1...3N-3))$ captures the non-zero eigenvalues of the $3N \times 3N$ dynamical matrix. These eigenvalues are proportional to the squares of the frequencies associated with the oscillations of the $3N$ normal-modes of the crystal, and the three zero-frequency modes describe the centre-of-mass motion. The particle-particle interactions place no constraint on the centre-of-mass position, and these soft modes are just a consequence of the translational invariance of the system.

While the free-energy in eqn. 6.27 is linearly dependent upon the temperature, the dependence upon the density and the structure is not so simple, and separate eigenvalue calculations must be carried out for every density and structure of interest. In some cases, the harmonic approximation will yield *negative* values of λ (imaginary frequencies), indicating that the system is dynamically unstable and that that structure will not be observed, at that density, in the small-displacement (low-temperature) limit. An example of an harmonically unstable system is shown in figure 6.6, along with a way in which that structure may be stabilised.

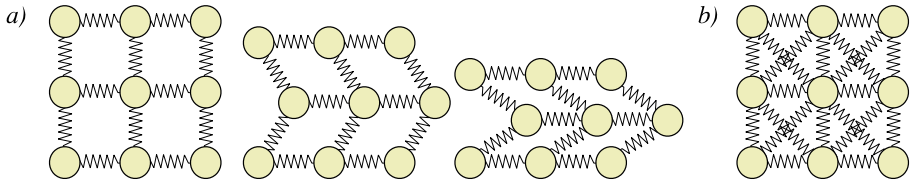


Figure 6.6: (a) A simple square lattice is unstable when the first-nearest neighbour interactions are dominant, and will collapse. This is indicated within the harmonic approximation by the presence of negative eigenvalues in the eigenspectrum of the dynamical matrix. (b) The crystal can be stabilised, however, if the interactions with the second nearest-neighbours are sufficiently strong. These interactions act like the cross-beams in a building's structure, and resist the unstable shear modes of the square lattice.

The mechanism of stabilisation (shown in fig. 6.6) illustrates that the harmonic calculation is also dependent upon the truncation-scheme being invoked (see also fig. 6.7). Furthermore, the choice of truncation not only affects the free-energy (by determining the population of non-zero elements in the dynamical matrix) but can also affect the linear term in eqn. 6.22 in such a way as to render the harmonic calculation meaningless. As stated above, for any stable arrangement of particles, the mean force on a particle should be zero when it is at its site. This is a consequence of the symmetry of the neighbouring lattice sites surrounding each particle, and so will cease to be the case if this symmetry is broken. Any truncation scheme which does not preserve the symmetry of the crystal will destabilise it, and so any truncation schemes for which the mean force on any particle is not zero must be rejected. Here, we use

only spherical truncations, for which we calculate the mean force on the particles and then reject any truncation radii for which the mean force is not zero (to within machine precision).

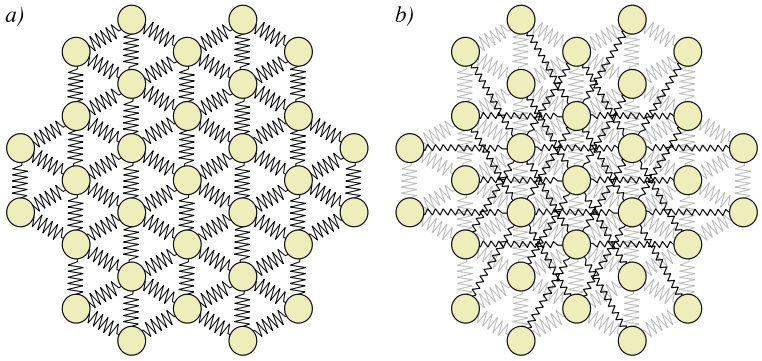


Figure 6.7: Schematic illustration showing a two-dimensional hexagonal crystal with (a) first-nearest and (b) first- & second-neighbour interactions.

For any pair of (stable) candidate structures (α & α'), the overall free-energy difference between them is written as

$$\Delta f_H(\tilde{\rho}) = e_{gs}(\alpha', \tilde{\rho}) - e_{gs}(\alpha, \tilde{\rho}) + kT \left(\tilde{f}_h(\alpha', \tilde{\rho}) - \tilde{f}_h(\alpha, \tilde{\rho}) \right) \ , \tag{6.28}$$

for any density, $\tilde{\rho}$. To determine the *fcc–hcp* phase behaviour of the system within the harmonic approximation, we apply the condition for phase coexistence ($\Delta f_H(\tilde{\rho}) = 0$) to eqn. 6.28. Rearranging this expression, we find that the temperature at which the total free-energy difference is zero (denoted as $\tilde{T}_H$) can be calculated using

$$\tilde{T}_H = \frac{kT_H(\tilde{\rho})}{\epsilon} = -\frac{\tilde{e}_{gs}(fcc, \tilde{\rho}) - \tilde{e}_{gs}(hcp, \tilde{\rho})}{\tilde{f}_h(fcc, \tilde{\rho}) - \tilde{f}_h(hcp, \tilde{\rho})} \ . \tag{6.29}$$

So, given a numerical algorithm that is capable of determining the eigenspectrum of the dynamical matrix of the two structures (at a given density), we can carry out calculations for a wide range of densities and determine the position of the *hcp*→*fcc* transition (i.e. $\tilde{T}_H$) along the way.

6.4.1 Implementation details

The harmonic-approximation calculation consists of three main stages: Firstly, a crystalline lattice of N sites is constructed. To allow a direct comparison of the lattice switch and harmonic calculations, we use exactly the same system-sizes and geometries (as defined in §4.8.1) in both cases. Having defined the system, we go on to construct the dynamical matrix based on the interactions between particles within a given (scaled or fixed) cut-off. The usual periodic boundary conditions are applied at this stage. However, to determine the eigenvalues of the dynamical matrix *numerically* (using standard matrix algorithms), we must re-express this matrix as a simple $3N \times 3N$ array. The three Cartesian directions

(x, y and z) are identified by an index $v = 1$, 2 or 3. The particle displacements are then expressed as $u_I = u_i^v$, where $I = 3(i-1)+v$ and where I runs from 1 to $3N$. Using the same convention, the dynamical matrix is mapped onto a two-dimensional array W_{IJ},

$$W_{IJ} = K_{ij}^{vw} \tag{6.30}$$

where $I = 3(i-1)+v$ and $J = 3(j-1)+w$. The eigenvalues (λ_d) of the array W_{IJ} can then be determined using standard matrix-eigenvalue algorithms. A number of different algorithms were compared, and the 'best' algorithm was chosen on the basis of the computational efficiency and numerical accuracy of the results. The algorithms were taken from Numerical Recipes (the Jacobi method and the QL algorithm [88, §11.1, §11.3]) and the NAG library (QL algorithms `f02abc` & `f02aac` [112]).

The accuracy of the result is of course limited by the numerical precision of the machine, but more importantly it is also limited by the algorithm itself. The 'quality' of the results was measured by observing how close to zero the three 'zero modes' were actually determined to be.[6] It was found that, for a given level of machine precision (32-bit or 64-bit floating-point arithmetic), all the algorithms gave reasonable results which were in agreement with one another. In terms of CPU time required, however, the different algorithms were somewhat easier to discriminate. While the eigenvalues of the $N = 6^3$ system could be determined in a few seconds by any of the algorithms, the $N = 12^3$ system (requiring the eigenvalues of a 5184×5184 array) could not be solved in a reasonable time using the Jacobi method. The QL algorithms allowed the unnecessary eigen*vector* calculation to be switched off and worked well at 32-bit precision. Of these, the NAG routine (`f02abc`) was found to be slightly faster.

6.4.2 Stability of the crystalline structures

As well as *fcc* and *hcp*, a number of other crystalline structures were also investigated via the harmonic approximation. The simple cubic, diamond and body-centered cubic structures were *not* found to be harmonically stable for *any* range of densities (in agreement with ref. [113]). It appears, therefore, that the Lennard-Jones potential always binds too weakly to all but the nearest-neighbour particles, denying the kind of 'cross-beam' stabilisation described in figure 6.6.

The *fcc* and *hcp* structures however, are stable (i.e. have only real mode-frequencies) over a wide range of densities, from $\tilde{\rho} \approx 0.78$ to at least $\tilde{\rho} \approx 3.40$. This is largely independent of the truncation-scheme employed, although the actual free-energies themselves are not. The upper-limit of 3.40 represents a computational limit upon the accuracy of the calculation; as the density is increased, the strength of

[6] As the zero modes are only approximately zero, they must be clipped out of the frequency sum by using a maximum-tolerance factor. In order to avoid any possible mis-classification of modes, the code was designed to warn the user if any more or less than three modes were counted as zero modes.

the first-nearest neighbour interaction diverges, and this divergence can destabilise the calculation as a consequence of the finite machine precision. The lower limit, 0.78, is a more physical result; the Lennard-Jones potential will not form *fcc* or *hcp* at densities lower than this (in the low temperature limit). As stated in [113], this artifact of the harmonic approximation can be combined with a knowledge of the thermal expansion properties of the solid to produce an estimate of the melting temperature as the temperature at which the density has fallen to the point where the structure becomes harmonically unstable. However, it would be wrong to read too much into this, as any estimate of the melting point *must* consider the free-energies of both the fluid and solid phases [95], and anharmonic contributions to the free-energy become increasingly important as the system approaches melting.

6.4.3 Truncation-scheme dependence

To determine the relative merits of the fixed-length and scaled (fixed neighbour-list) truncation, the harmonic free-energy difference between *fcc* and *hcp* was evaluated using both techniques. Figure 6.8 shows the results of these calculations, using a large cut-off, and two small cut-offs based on fixed length and fixed neighbour-list truncation. It is clear from fig. 6.8 that the fixed neighbour-list truncation (at $1.5r_{nn}$) forms a much better estimate of the (untruncated) free-energy difference than the fixed-length (1.5σ) truncation.

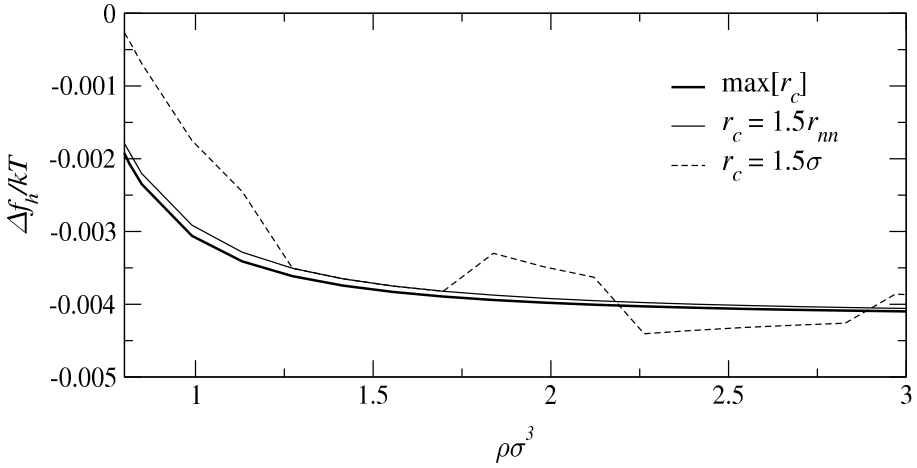


Figure 6.8: The harmonic free-energy difference as a function of density for $N = 6^3$ Lennard-Jones particles, using various truncation schemes. The thick line shows the best harmonic result, based on the maximum cut-off that can be used with this system-size ($r_c = 2.89r_{nn}$). The thin line shows the results taken when only first and second nearest-neighbours are taken into account ($r_c = 1.5r_{nn}$), which follows the max[r_c] curve closely. The dashed line shows the free-energy difference calculated using a fixed-length ($r_c = 1.5\sigma$) cut-off, which generally forms a poorer approximation to the 'best' (max[r_c]) curve.

The argument in favour of the fixed neighbour list method can be strengthened even further by con-

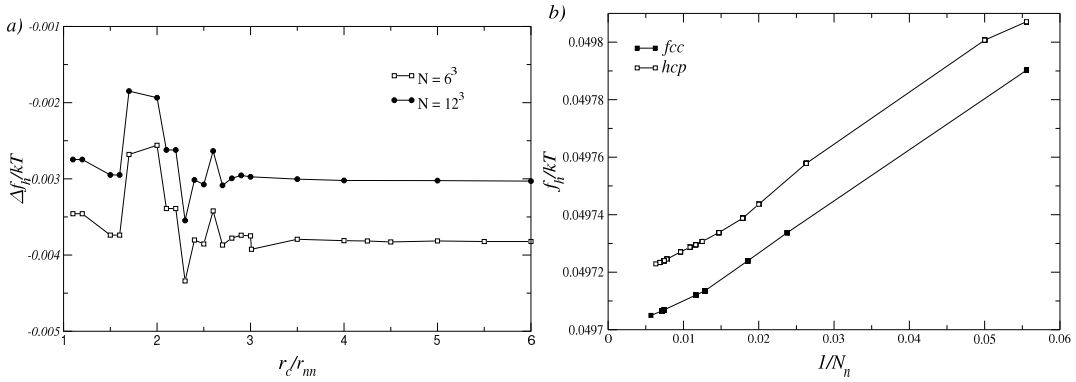


Figure 6.9: a) The dependence of the *fcc-hcp* harmonic free energy difference on the cut-off radius and system size, for a fixed density of $\tilde{\rho} = 1.08762$. The results from $r_c = 1.5\sigma$ are within 3% of those for the 'infinite' truncation radius. (b) The harmonic free-energy versus the (inverse) number of interacting neighbours ($N = 6^3$, $\tilde{\rho} = 1.08762$). Note that all of these results are relevant to both the fixed and scaled truncation schemes, as at fixed density the schemes differ only by a multiplicative (scale) factor in r_c. The jagged form shown in (a) must therefore arise from the fact that the value of N_n changes at different points in r_c for each of the two structures.

sidering the dependence of the free-energy difference upon the range of truncation. This aspect of the calculation is examined in figure 6.9, showing that although the free-energy difference is strongly dependent upon the truncation radius (fig. 6.9(a)), this dependence is very strongly correlated to the number of neighbours that are included by the truncation (fig. 6.9(b)). When the value of r_c picks out *equal numbers of interactions* for *both* structures, the free-energy difference is almost identical. Only the *fixed* neighbour-list truncation scheme can ensure that the number of neighbours included in the calculation is always equal for both structures over the entire range of densities of interest.

These results suggest that the very close ranged 1st or 1st & 2nd nearest-neighbour interactions (for which the number of neighbours is identical in both structures) will be sufficient to encapsulate the behaviour of the harmonic free-energy difference. Indeed, fig. 6.9(a) shows that while 1st-neighbour interactions ($r_c = 1.0r_{nn}$) give a slightly low estimate of the free-energy difference, a truncation of $r_c = 1.5r_{nn}$ (including 1st & 2nd neighbours) forms a good approximation to the large r_c limit.

The harmonic approximation can also be used to help quantify the finite-size effects involved in this problem. Figure 6.9(a) compares the free-energy difference using the $1.5r_{nn}$ truncation for systems of 6^3 and 12^3 sites. In form, the free-energies from the two system sizes follow each other closely, but the shift in the free-energy due to the change in N is significant. Unfortunately, it has not been possible to use the harmonic approximation to determine the free-energy of the largest $N = 18^3$ system, as this requires the determination of the eigenvectors of a 17496×17496 matrix. This is simply not computationally feasible in terms of both memory (>1Gb) and CPU time ($\sim$ two weeks) required.[7]

[7]It should be noted that these requirements are an artifact of the manner in which the harmonic approximation has been implemented. Large (and indeed infinite) systems can be dealt with under the harmonic approximation by using more complex techniques based on integrating over the relevant normal-mode frequencies in Fourier space. Such calculations were not

However, as we shall see later, there is good reason to believe that the finite-size effects for the $N = 12^3$ system are negligible.

6.4.4 Verifying the harmonic calculation

The test-cases used to check the reliability of the harmonic calculation fall into two classes; either simple analytically soluble systems, or comparisons with relevant results for more complex systems taken from the literature. In the former case, the central test systems were a 1-dimensional periodic loop of n harmonic oscillators, and a two-dimensional two-particle periodic (square-lattice) system. Both systems were found to give answers in agreement with the analytical results [9] (to within machine-precision rounding errors). These simple, low-dimensional test systems provide valuable evidence that the results of the significantly more complex three-dimensional systems may be trusted because of the way in which the harmonic approximation was implemented. The algorithm was written so that the dimensionality of the system is just a parameter of the calculation,[8] instead of the more usual approach of having a separate computer program for each. Therefore, a test performed with a system of one spatial dimension is at least a partial test of the harmonic calculation for any number of dimensions.

Of course, there is always the possibility that some dimension-dependent error may creep in, and so published results based upon more realistically complex systems were sought. However, the number of papers dealing with the *fcc–hcp* harmonic free-energy is quite small, and of these very few gave actual *quantitative* predictions against which our calculations could be tested *and* gave enough information about their calculations (especially the truncation scheme) to make any comparison possible. For example, the harmonic results of Choi *et al.* [83] are clearly consistent with the results presented here, but as their data is only presented graphically and no truncation details are provided, it is impossible to perform any kind of rigorous comparison.

However, the $N \to \infty$ results for the harmonic *fcc–hcp* free-energy difference presented in Salsburg & Huckaby [99] do allow a more precise comparison. Using their truncation scheme (1st and 2nd neighbours only), the results from the $N = 6^3$ and $N = 12^3$ systems were combined with an assumed $1/N$ form of asymptotic $N \to \infty$ convergence to produce a free-energy difference of $\Delta\tilde{f}_h = 0.00262$, in complete agreement with this publication.[9] These results also indicate that the free-energy difference for the $N = 12^3$ system is within $9 \times 10^{-5} NkT$ (a few percent) of the result for the thermodynamic limit. As we shall see in §6.6.1 this degree of error is of the order of the statistical uncertainty in our

applied here; see ref. [111] and references therein for more information.

[8] For example, the harmonic modes of a four dimensional hyper-lattice can be calculated using *exactly* the same computer program.

[9] Note that this precise comparison requires care, in that the same two densities for the two different structures as used in [99] (i.e. the differing ground-state densities) must be used in order to recover this result.

LSMC calculations, and so should not significantly compromise our results.

6.4.5 The harmonic phase diagram

As mentioned earlier, the harmonic contribution to the *fcc–hcp* free-energy difference varies linearly with temperature, and can be combined with the ground-state energy difference to produce a phase diagram for the harmonic crystal (eqn. 6.29). This *process* is illustrated in fig. 6.10(a), which shows the shape of the free-energy difference *surface* as a function of both density and temperature. The $\Delta f = 0$ phase-boundary (determined via eqn. 6.29) is shown more clearly in figure 6.10(b), for both the $N = 6^3$ and $N = 12^3$ harmonic systems.

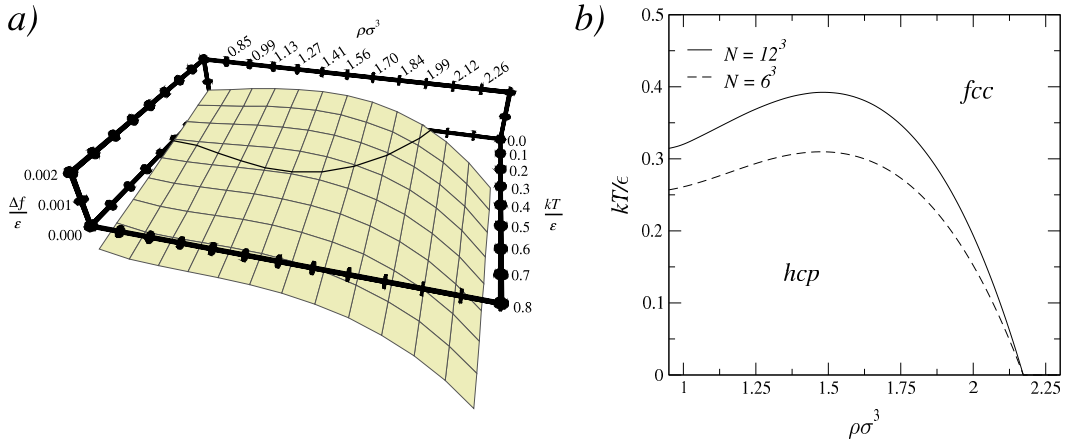


Figure 6.10: Harmonic phase diagrams of the Lennard-Jones solid. (a) Three-dimensional plot showing the variation of the *fcc–hcp* free-energy difference surface with both density and temperature. The dark line indicates the $\Delta f = 0$ phase boundary. (b) Plots of the harmonic $\Delta f = 0$ phase boundary in the $\rho - T$ plane for $N = 6^3$ and $N = 12^3$ Lennard-Jones particles.

6.4.6 Estimating the anharmonic effects

The harmonic approximation, as presented above, began by performing a truncation of the configurational energy expanded in terms of the particle displacement (eqn. 6.22), and produces an approximate expression for the free-energy of a crystal (eqn. 6.26). Trivially, we may write the *true* free-energy of the crystal (for a given structure, density and temperature) in terms of the difference between the actual free-energy and its harmonic counterpart,

$$f(\alpha, \tilde{\rho}, \tilde{T}) = e_{gs}(\alpha, \tilde{\rho}) + f_h(\alpha, \tilde{\rho}, \tilde{T}) + f_a(\alpha, \tilde{\rho}, \tilde{T}) \tag{6.31}$$

where we have introduced the 'anharmonic' free-energy contribution, f_a, which is the free-energy difference between the harmonic approximation and the true free-energy of the system. The ground-state

energy describes the zeroth order (temperature independent) behaviour, the harmonic free-energy is the (linear) first-order term (c.f. eqn. 6.27),

$$f_h(\alpha, \tilde{\rho}, \tilde{T}) = \tilde{T} \tilde{f}_h(\alpha, \tilde{\rho}) \ , \tag{6.32}$$

and the anharmonic free-energy contribution f_a represents all the remaining terms ($\mathcal{O}(\tilde{T}^2)+\mathcal{O}(\tilde{T}^3)+...$) in the expansion. Therefore, at low temperatures, we may approximate the anharmonic free-energy by truncating it to second order,

$$f_a(\alpha, \tilde{\rho}, T) \approx \tilde{T}^2 \tilde{f}_a(\alpha, \tilde{\rho}) \ . \tag{6.33}$$

Both eqns. 6.32 and 6.33 hold separately for any two structures (α) we choose to consider, and so the harmonic and anharmonic free-energy difference between two phases can be written as,

$$\Delta f_h(\tilde{\rho}, \tilde{T}) = \tilde{T} \Delta \tilde{f}_h(\tilde{\rho}) \ , \tag{6.34}$$

$$\Delta f_a(\tilde{\rho}, \tilde{T}) \approx \tilde{T}^2 \Delta \tilde{f}_a(\tilde{\rho}) \ , \tag{6.35}$$

implying that

$$\frac{\Delta f_a(\tilde{\rho}, \tilde{T})}{\Delta f_h(\tilde{\rho}, \tilde{T})} = \tilde{T} \frac{\Delta \tilde{f}_a(\tilde{\rho})}{\Delta \tilde{f}_h(\tilde{\rho})} \ . \tag{6.36}$$

So, for low temperatures, we expect the ratio of the anharmonic and harmonic contributions to the free-energy to vary linearly with temperature. The behaviour of this ratio as a function of *density* is not quite so straightforward. However, as shown in refs. [114] the dependence of the free-energy on the density can be approximated as

$$\frac{\Delta \tilde{f}_a(\tilde{\rho})}{\Delta \tilde{f}_h(\tilde{\rho})} \approx c \frac{[\phi'''(r_{nn})]^2}{[\phi''(r_{nn})]^3} \ , \tag{6.37}$$

where r_{nn} is the nearest-neighbour separation for a density $\tilde{\rho}$, and c is a constant which must be determined empirically. The original truncation of the interaction potential assumes that only those terms which are less than third-order in the displacements need be included, and eqn. 6.37 is essentially a statement that the extent to which the original approximation fails is most significantly controlled by the ratio of the second-order (harmonic) ϕ'' term to the discarded third-order ϕ''' term.

The approximate forms of the temperature and density variation of the free-energy supplied above can be combined to produce an approximation for the form of the $\tilde{\rho} - \tilde{T}$ phase boundary. The harmonic approximation predicts the phase boundary to lie at

$$\Delta e_{gs}(\tilde{\rho}) + \Delta f_h(\tilde{\rho}, \tilde{T}_h) = 0 \ , \tag{6.38}$$

whereas the true phase boundary is defined by

$$\Delta e_{gs}(\tilde{\rho}) + \Delta f_h(\tilde{\rho}, \tilde{T}) + \Delta f_a(\tilde{\rho}, \tilde{T}) = 0 \ . \tag{6.39}$$

Using the scaling forms for the temperature variation (eqns. 6.34 & 6.35), these two equations become,

$$\begin{aligned} \Delta e_{gs}(\tilde{\rho}) + \tilde{T}_h \Delta \tilde{f}_h(\tilde{\rho}) &= 0 \ , \\ \Delta e_{gs}(\tilde{\rho}) + \tilde{T} \Delta \tilde{f}_h(\tilde{\rho}) + \tilde{T}^2 \Delta \tilde{f}_a(\tilde{\rho}) &= 0 \ . \end{aligned} \tag{6.40}$$

Upon subtraction, we find that the difference between the harmonic and the true transition temperatures can be expressed as

$$\Delta\tilde{T} = \tilde{T} - \tilde{T}_h = -\tilde{T}^2 \frac{\Delta \tilde{f}_a(\tilde{\rho})}{\Delta \tilde{f}_h(\tilde{\rho})} = c\tilde{T}^2 \frac{[\phi'''(r_{nn})]^2}{[\phi''(r_{nn})]^3} \ . \tag{6.41}$$

So, given a full knowledge of the harmonic phase diagram, the difference between that and the true phase diagram over the entire density range can be estimated once the value of c has been determined. By using LSMC to measure the position of the phase boundary at one density, we can fix the value of c, and (using eqn. 6.41) approximate the rest of the curve. The quality of this simple parameterisation of the coexistence curve can be ascertained by collecting further numerical estimates for the coexistence behaviour.

6.5 Lattice-switch Monte Carlo for soft potentials

The structure of the LSMC algorithm remains the same as in the case of hard-spheres: There are three possible types of Monte Carlo move; single particle translations, the lattice switch, and (in the constant pressure ensemble) volume dilations. The main difference is that the acceptance probability has changed, and now depends on the magnitude of the energy cost of the move (instead of simply on whether the move is possible or not). As the lattice-switch move is the core idea behind this research, we examine its soft-potential form first, before considering the more standard particle and volume moves in §6.5.5.

The 'physical' nature of the lattice-switch move itself remains unchanged. We use decomposed particle coordinates (eqn. 4.3) to associate each particle of the system with a lattice site, and then attempt to swap one set of lattice sites for another using the mapping shown in fig. 4.6. The *difference* between the hard-sphere and soft-potential formulations of the lattice switch can be clarified by comparing the hard-sphere configuration-space illustration in fig. 3.11 with its softer counterpart shown in fig. 6.11. Instead of the hard boundaries of the hard-sphere system, where only a well-defined subset of configurations will allow the switch to occur, there is always *some* chance that the lattice-switch move will be accepted. The energy cost of the move is determined by our modified configurational energy (eqn. 6.19) as

$$\Delta E_{LS}(\alpha \to \alpha') = \big(E_{gs}(\alpha', \tilde{\rho}) - E_{gs}(\alpha, \tilde{\rho})\big) + \big(\mathcal{E}(\alpha', \tilde{\rho}) - \mathcal{E}(\alpha, \tilde{\rho})\big) \ , \tag{6.42}$$

and from eqn. 3.19, the probability of accepting the lattice-switch move is

$$P_{acc}(\alpha \to \alpha') = \min\Big[1.0, \exp\Big[-\tilde{\beta}\Delta E_{LS}\Big]\Big] \ . \tag{6.43}$$

The lattice switch defined here does not represent the only way in which the lattice switch may be implemented, but on the evidence of the hard-sphere work (§4.10.4), this switch is likely to be the most efficient site-to-site mapping for this system.

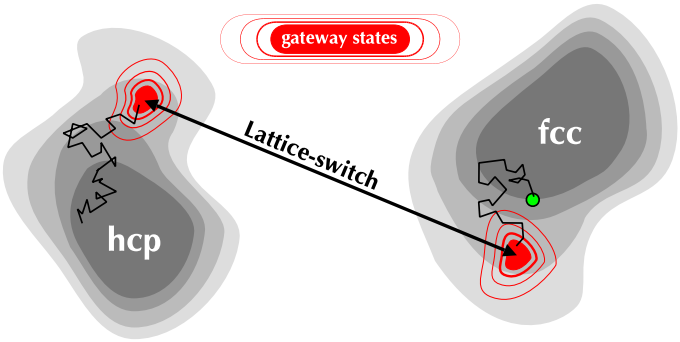


Figure 6.11: Schematic illustration of the lattice-switch operation in the configuration space of a soft-potential system.

As mentioned above, there is always some chance that the lattice switch may occur spontaneously (see §6.6.1). However, like the configurational-energies themselves, the cost of performing the switch is extensive and as the system is made larger, non-Boltzmann sampling techniques must be used to encourage the switch to occur. Thus, as in the case of hard-spheres, we must have some way of measuring 'how close' we are to the configurations for which there is a high probability of accepting the switch.

6.5.1 The lattice-switch order parameter

Compared with the hard-sphere system, where we used the number of pairs of overlapping particles created by the switch, the soft-potential definition of the lattice-switch order parameter is perhaps more natural. We simply re-phrase the energy cost of performing the switch so that it may act as an order parameter, as this is the quantity which defines how likely the switch is to occur, and which we must bias towards zero in order to enhance the switching-probability. We define the order parameter as,

$$\mathcal{M}(\{\vec{u}\}) = \mathcal{E}(\{\vec{u}\}, hcp) - \mathcal{E}(\{\vec{u}\}, fcc) \ , \tag{6.44}$$

where we have chosen to use the energy cost of switching from *fcc* to *hcp*, as opposed to the other way around. For an equilibrium configuration of a given phase, we expect the lattice-switch move to cost a significant amount of energy to perform, i.e. we expect ΔE_{LS} to be large and positive. Therefore the choice of order parameter in eqn. 6.44 should lead to a distribution of $\mathcal{M}$ like that for hard-spheres, with the *fcc* peak in the $\mathcal{M} > 0$ domain and the *hcp* peak in $\mathcal{M} < 0$. However, unlike the hard-sphere case, $\mathcal{M}$ is no longer an order parameter in the strictest sense of the term, because the softness of the interaction blurs the boundary between the structures in the $\mathcal{M} \sim 0$ region. Indeed, it is *possible* for the

simulation to visit either phase at *any* value of $\mathcal{M}$, and so the simulation algorithm must record both the structure and the value of $\mathcal{M}$ in order to determine the time spent in each phase.

Multicanonical weighting

The multicanonical weighting algorithm is also changed by the 'softness' of the order parameter, because the values of $\mathcal{M}$ are no longer naturally discrete. However, as the order parameter is an energy difference, the Boltzmann factor provides a natural choice of discretization scale: $m_0 = \beta^{-1}$ (or in our reduced units, $\tilde{m}_0 = \beta m_0 = 1$). We define $i = 0^{\pm}, \pm 1, \pm 2...$ as an index to a set of bins where $\mathcal{M}$ belongs to a bin i if

$$im_0 \leq \mathcal{M} < (i+1)m_0 \quad \text{when} \quad \mathcal{M} > 0 \ , \text{ and} \tag{6.45}$$

$$(i-1)m_0 \leq \mathcal{M} < im_0 \quad \text{when} \quad \mathcal{M} < 0 \ . \tag{6.46}$$

Having defined our discretization scale, the machinery of multicanonical extended sampling (for any external conditions, $\mathcal{C}$) can be applied exactly as described in the general formulation of this method given in §3.6.1. This allows the distribution $P(\mathcal{M}|\mathcal{C})$ to be sampled uniformly, and by recording the current structure as well as the value of $\mathcal{M}$, the same simulation also allows the distributions associated with each of the two phases ($P(\mathcal{M}|fcc,\mathcal{C})$ & $P(\mathcal{M}|hcp,\mathcal{C})$) to be determined.

6.5.2 The constant volume ensemble

As stated in §2.3, the Helmholtz free energy associated with any particular structure may be written as

$$F(\alpha|N,\rho,\beta) = -\beta^{-1} \ln \mathcal{Z}(\alpha|N,\rho,\beta) \ , \tag{6.47}$$

where $\mathcal{Z}(\alpha|N,\rho,\tilde{\beta})$ is the partition function (or configurational weight) associated with that structure,

$$\mathcal{Z}(\alpha|N,\rho,\beta) = \prod_i \left[\int_\alpha d\vec{r}_i \right] \exp\left[-\beta E(\{\vec{r}\})\right] \tag{6.48}$$

$$= \prod_i \left[\int_\alpha d\vec{u}_i \right] \exp\left[-\beta \mathcal{E}(\{\vec{u}\},\alpha) - \beta E_{gs}(\rho,\alpha)\right] \tag{6.49}$$

As the density is fixed, the ground-state energy term is constant and may be 'brought outside' the integral. This allows the free-energy difference between two structures to be written as

$$\Delta f(\tilde{\rho},\tilde{\beta}) \equiv \frac{1}{N\epsilon}\left[F(\alpha|N,\rho,\beta) - F(\alpha'|N,\rho,\beta)\right] \ , \tag{6.50}$$

$$= \Delta \tilde{e}_{gs}(\tilde{\rho}) - \frac{1}{N\tilde{\beta}} \ln \mathcal{R}_{\alpha,\alpha^\star}(\tilde{\rho},\tilde{\beta}) \ , \tag{6.51}$$

where the ratio of configurational weights is defined to be

$$\mathcal{R}_{\alpha,\alpha^\star} = \frac{\mathcal{Z}_\alpha}{\mathcal{Z}_{\alpha^\star}} = \frac{\prod_i \left[\int_\alpha d\vec{u}_i\right] e^{-\tilde{\beta}\mathcal{E}(\{\vec{u}\},\alpha)}}{\prod_i \left[\int_{\alpha^\star} d\vec{u}_i\right] e^{-\tilde{\beta}\mathcal{E}(\{\vec{u}\},\alpha^\star)}} \tag{6.52}$$

$$= \frac{\sum_{\mathcal{M}} P(\mathcal{M}|\alpha, N, \tilde{\rho}, \tilde{\beta})}{\sum_{\mathcal{M}} P(\mathcal{M}|\alpha^\star, N, \tilde{\rho}, \tilde{\beta})} \, . \tag{6.53}$$

This expression only applies to Boltzmann importance sampling, but the inclusion of the multicanonical weight function alters the calculation only slightly. The bias introduced by the weight function can be unfolded using

$$\frac{\sum_{\mathcal{M}} P(\mathcal{M}|\alpha, N, \tilde{\rho}, \tilde{\beta})}{\sum_{\mathcal{M}} P(\mathcal{M}|\alpha^\star, N, \tilde{\rho}, \tilde{\beta})} = \frac{\sum_{\mathcal{M}} P(\mathcal{M}|\alpha, \vec{\eta}, N, \tilde{\rho}, \tilde{\beta}) \exp\left[-\eta(\mathcal{M})\right]}{\sum_{\mathcal{M}} P(\mathcal{M}|\alpha^\star, \vec{\eta}, N, \tilde{\rho}, \tilde{\beta}) \exp\left[-\eta(\mathcal{M})\right]} \, . \tag{6.54}$$

Apart from having to record both the cost of the switch *and* the current structure, the calculation of $\mathcal{R}$ proceeds in almost exactly the same way as for hard-spheres, using block analysis to estimate the errors associated with the ensemble average. Having determined $\mathcal{R}$, the ground-state energy is folded in using eqn. 6.51, thus yielding the total free-energy difference.

6.5.3 The constant pressure ensemble

In the constant pressure ensemble, the density will fluctuate, and so the ground-state energy cannot be factored out of the integral. We use the full Gibbs free-energy for each phase,

$$G(\alpha|N, \mathrm{p}, \beta) = -\beta^{-1} \ln \mathcal{Z}(\alpha|N, \mathrm{p}, \beta) \, , \tag{6.55}$$

where

$$\mathcal{Z}(\alpha|N, \mathrm{p}, \beta) = \int dv \; \mathcal{Z}(\alpha|N, \rho, \beta) \exp\left[-\beta \mathrm{p} V\right] \, . \tag{6.56}$$

The free-energy difference between two phases α and α' is simply,

$$\Delta g(\tilde{\mathrm{p}}, \tilde{T}) = -\frac{1}{N\tilde{\beta}} \ln \mathcal{R}_{\alpha,\alpha'}(\tilde{\mathrm{p}}, \tilde{\beta}) \, , \tag{6.57}$$

where

$$\mathcal{R}_{\alpha,\alpha'} = \frac{\sum_{\mathcal{M}} P(\mathcal{M}|\alpha, \vec{\eta}, N, \tilde{\mathrm{p}}, \tilde{\beta})}{\sum_{\mathcal{M}} P(\mathcal{M}|\alpha^\star, \vec{\eta}, N, \tilde{\mathrm{p}}, \tilde{\beta})} \, . \tag{6.58}$$

The multicanonical weighting and the measurement procedure are almost exactly the same as for the canonical ensemble; they only differ in the final stage of the free-energy calculation (i.e. by using eqn. 6.57 instead of eqn. 6.51).

6.5.4 The consequences of truncation

Although the evidence presented in §6.4.3 *suggests* that our chosen truncation scheme (eqn. 6.18 with $r_c = 1.5 r_{nn}$) should form a good representation of the true behaviour, it is far from proof that this is the

case. The only way to be sure is to evaluate explicitly the free-energy cost of truncating the potential. As for hard-spheres (§4.8.2), we use the direct-difference method of Bennett [25] (outlined in §2.4.3) to evaluate this free-energy difference. Again, we can expect the particle configurations associated with different ranges of truncation to be very similar, and that the technique can be applied to this problem with no modification. Note that, as was the case for hard-spheres, the formulation of Bennett is exact under the assumption that the configurations explored under Hamiltonian 1, for which the above expectation value is large, are identical to those that would be visited under Hamiltonian 2, as illustrated in fig. 6.12.

To understand how this assumption may fail and produce erroneous results, consider the following. In a crystalline solid, particles are constrained to stay near their lattice sites, caged in by their neighbours. The particles will not swap positions in the lattice on any reasonable timescale. Therefore, the configuration space of the system contains $N!$ identical but distinct fragments, corresponding to the $N!$ different possible ways of ordering the particles within the lattice. However, a Hamiltonian which only uses a subset of the particle interactions may allow particles to pass through one another, thus allowing the simulation to move between the different fragments. Thus, the comparison of a lenient Hamiltonian with a more strict one can lead to poor results if the lenient constraint allows the system to jump between fragments in an unphysical manner. This, however, should not be a problem here, as this swapping of particles between sites would rapidly lead to the breakdown of the entire simulation, and this was only observed at very low densities/high temperatures, i.e. at or above the melting point, well away from the *fcc–hcp* coexistence region.

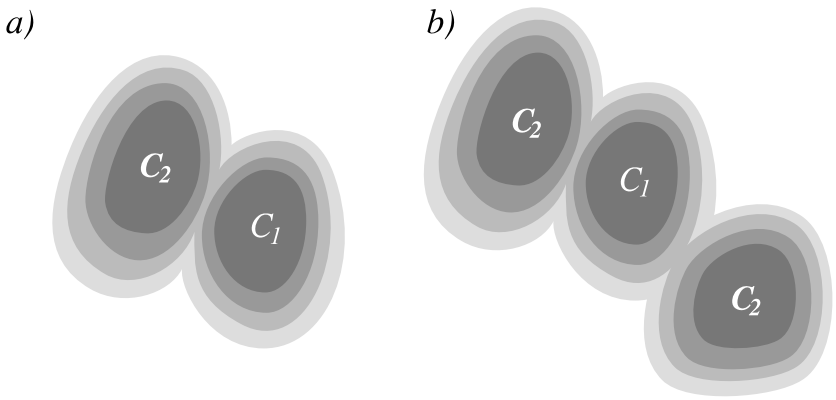


Figure 6.12: As for hard-spheres (see fig. 4.11), neighbouring fragments of phase space may damage the virtual Hamiltonian calculation. (a) If all is well, then the individual fragments of configuration spaces (C_1 & C_2) associated with the two Hamiltonians will overlap significantly with each other on a one-to-one basis. (b) However, if the configurations associated with Hamiltonian 1 serve to 'bridge the gap' between distinct fragments of the configuration space of Hamiltonian 2 (e.g. by allowing particles to swap sites), then the calculation will cease to give valid results.

Here, we use the direct-difference method to determine the free-energy cost of moving from the 1st & 2nd neighbour (scaled) truncation (H_1) to a more 'correct' $\mathcal{O}(N^2)$ calculation (H_2). Two separate importance-sampling MC simulations are used, locked into each of the two different structures. These *fcc* and *hcp* simulations proceed under Hamiltonian 1, i.e. the interactions are truncated at the 1st & 2nd neighbour ($r_c = 1.5\sigma$) level. Periodically (every T_{nn} sweeps) the simulation algorithm calculates the energy of the current microstate under the $\mathcal{O}(N^2)$ Hamiltonian:

$$\mathcal{E}_{nn} = \sum_{i=1}^{n} \sum_{j=1}^{j<i} \phi_{nn}(r_{ij}) \ , \tag{6.59}$$

using

$$\phi_{nn} = \phi(r_{ij}) - \phi(R_{ij}) \qquad \text{if} \quad r_{ij} < r_{max}, \tag{6.60}$$

$$= 0 \qquad \text{otherwise}, \tag{6.61}$$

where r_{max} is the radius of the largest sphere that will fit inside the simulation cell. As in the 1st & 2nd neighbour calculation, the subtraction of $\phi(R_{ij})$ ensures that the ground-state contribution has been removed from the configurational energy, and so for the constant pressure ensemble the true ground-state energy must be added back into both configurational energies,

$$H_1 \equiv \beta E = \beta \left(\mathcal{E}(\{\vec{r}\}) + E_{gs}(\rho) \right) \tag{6.62}$$

$$H_2 \equiv \beta E_{nn} = \beta \left(\mathcal{E}_{nn}(\{\vec{r}\}) + E_{gs}(\rho) \right) \tag{6.63}$$

whereas in the constant-volume simulations, the ground-state contribution is "folded out" throughout the simulation work, and simply folded into the results at the end. In either case, the property of interest (via eqn. 2.20) is

$$D_{nn} = \exp\left[-\beta \left(E_{nn} - E\right)\right] = \exp\left[-\beta \left(\mathcal{E}_{nn} - \mathcal{E}\right)\right] \ . \tag{6.64}$$

During the two simulations, the value of D_{nn} is recorded, along with the energies $\mathcal{E}_{nn}, \mathcal{E}$ and their difference. Then, a separate analysis code is used to produce an estimate of the difference between the free energies of the 'truncated' and 'untruncated' systems, for each structure. The first $\sim 10^4$ MCS are rejected as 'equilibration time', because each simulation is initialised in the perfect-crystal microstate. Block-data analysis (§3.4) is then applied to the remainder of the data, determining

$$f_{nn}(\alpha) \stackrel{\text{EB}}{=} -\frac{1}{N\beta} \ln \langle D_{nn} \rangle_{\alpha} \ , \tag{6.65}$$

for each block. As always, a number of different block-sizes were used to ensure that the calculations were free of significant correlations. This process is carried out for both crystals, and the correction to their free-energy difference (with respect to a truncated-interaction LSMC calculation) is determined as,

$$\Delta f_{nn} = f_{nn}(fcc) - f_{nn}(hcp) \ . \tag{6.66}$$

The calculation is essentially unchanged by moving to the constant-pressure ensemble; the Gibbs free-energy 'cost' of truncation for a given phase is simply,

$$g_{nn}(\alpha) \stackrel{\text{EB}}{=} -\frac{1}{N\beta}\ln\langle D_{nn}\rangle_\alpha \ , \tag{6.67}$$

and the effect on the free-energy difference between the phases is

$$\Delta g_{nn} = g_{nn}(fcc) - g_{nn}(hcp) \ . \tag{6.68}$$

The error associated with the *final* free-energy cost (Δf_{nn} in eqn. 6.66 and Δg_{nn} in eqn. 6.68) is simply calculated as the geometrical mean of the relative (block-analysis) errors of its two independently determined components (e.g. $g_{nn}(fcc)$ & $g_{nn}(hcp)$).

This procedure is certainly not the only way in which the cost of truncation may be estimated. For example, it would be possible to use the ideas behind the lattice switch to perform a full Hamiltonian switch, allowing a simulation to jump between two different truncation schemes by using the energy cost of the switch as the multicanonical order parameter, and evaluating the free-energy difference between the two Hamiltonians from their relative probabilities. However, the virtual-Hamiltonian approach only requires the $\mathcal{O}(N^2)$ energy calculation to be performed periodically, and so this method is significantly more efficient than any method relying on a simulation which implements the full $\mathcal{O}(N^2)$ interaction for a significant proportion of the time.

6.5.5 Implementation details

The lattice, the switch & the cell

As we are still dealing with the same structures, the definition of the lattice, the switch, and the shape of the periodic simulation cell are all exactly as for the hard-sphere work (see §4.8.1).

Single-particle MC moves

As a consequence of the centre-of-mass complications discussed in §4.10.8, only the random walk (RW) particle-move algorithm was employed here. That is, we choose a particle at random, and generate a new displacement vector by adding a uniform cubic random vector $\Delta\vec{u}_i$ to the original displacement $\vec{u}_i$. We calculate the change in energy associated with a move using an algorithm that implements the following local neighbour-list energy calculation,

$$\mathcal{E}_{loc}(i,\vec{u},\alpha) = \sum_j^{N_n}\left[\phi(r_{ij}) - \phi(R_{ij})\right] \ , \tag{6.69}$$

where the summation extends over the N_n neighbouring particles for which $R_{ij} < r_c$. This formulation takes advantage of the fact that the energy of the entire system does not have to be re-calculated from scratch at every step, because each particle only interacts with a subset of the others and so the interactions which do not change need not be taken into account. Using this rule, the energy change for the current phase can be calculated as follows:

$$\Delta\mathcal{E}_\alpha = \mathcal{E}_{loc}(i, \vec{u}_i + \Delta\vec{u}_i, \alpha) - \mathcal{E}_{loc}(i, \vec{u}_i, \alpha) \ , \tag{6.70}$$

that is, the change in total energy is equal to the change in the local energy between the trial displacement vector and the original one. Using this 'running-total' approach to the energy calculation is significantly faster than carrying out a full configurational-energy calculation for each particle move. However, the limited numerical precision of the machine means that this approach allows the *actual* energy of the system to drift away from the running total due to the accumulation of rounding errors. For this reason, the total energy was calculated periodically (every ~ 1000 MCS) and used to correct the running total. If the running total had slipped by more that 1 part in a million, the simulation was deemed to have failed (most likely, melted) and was halted.

A running total was also used to keep tabs on the energy of the current configuration for the 'other' phase ($\Delta\mathcal{E}_{\alpha'}$), by substituting α' for α in eqn. 6.70. From the definition of the lattice-switch order parameter (eqn. 6.44), the *change* in the $\mathcal{M}$ is determined as

$$\Delta\mathcal{M} = \Delta\mathcal{E}_{hcp} - \Delta\mathcal{E}_{fcc} \ . \tag{6.71}$$

Thus, by keeping two running totals, associated with the energies of the two phases, we always know the value of the lattice-switch order parameter. This allows the change in the multicanonical weight to be calculated as

$$\Delta\eta = \eta(\mathcal{M} + \Delta\mathcal{M}) - \eta(\mathcal{M}) \ , \tag{6.72}$$

and each particle move was accepted with a probability given by

$$P_{acc} = \min\left[1.0, \exp\left(-\beta\Delta E_\alpha + \Delta\eta\right)\right] \ . \tag{6.73}$$

Lattice-switch moves

The lattice switch was attempted once per sweep (on average) according to the acceptance rule given by eqn. 6.43. A successful lattice switch does not alter the value of $\mathcal{M}$ (although this success is *constrained* by the value of $\mathcal{M}$), and so the weight function does not (directly) affect the probability of accepting the lattice-switch move. If the move is accepted then the running totals (for $\mathcal{E}(\{\vec{u}\}, \alpha)$ and $\mathcal{E}(\{\vec{u}\}, \alpha')$) are swapped, and the simulation continues in the other phase.

Volume moves

In this case, the fixed aspect-ratio (FVM) algorithm was employed, as described in §3.2.2. A uniform random walk was performed in the volume, attempting one $V \to V + \Delta V = V'$ dilation per sweep (on average). As outlined in ref. [34, p.125], the power-law form of the Lennard-Jones interaction allows the energy change upon a (fixed aspect-ratio) volume dilation to be calculated via a simple scaling rule. Under the FVM scheme, the volume-scaling used by the NPT algorithm (eqn. 3.22) becomes simply

$$\vec{r} = L\vec{s}\,, \tag{6.74}$$

and so the configurational energy of the LJ system can be decomposed such that

$$\mathcal{E}(\{\vec{r}\}) = L^{-12}\mathcal{E}^{(12)}(\{\vec{s}\}) + L^{-6}\mathcal{E}^{(6)}(\{\vec{s}\})\,. \tag{6.75}$$

Using this scaling rule, the change in energy due to a volume dilation ($\Delta\mathcal{E}_v$) can be calculated very quickly indeed. There is an extra overhead, due to the book-keeping required to keep separate running totals for both the twelfth-power and sixth-power contributions to the configurational-energy, but this investment is well worth the effort. Note also that this algorithm is only applicable when a scaled-truncation is being used, which gives that truncation-scheme a significant advantage in the NPT ensemble. However the energy change is determined, the acceptance probability of a volume dilation (using eqn. 3.28) is

$$P_{acc} = \min\left[1.0, \exp\left(-\beta(E_{gs}(\alpha, V') - E_{gs}(\alpha, V)) - \beta\Delta\mathcal{E}_v - \beta \mathrm{p}\Delta V) + N\ln(V'/V) + \Delta\eta\right)\right]\,. \tag{6.76}$$

6.6 Results

This long series of results is organised as follows. Firstly, in §6.6.1, we outline the initial investigation of the system. This involves the tuning of the simulation parameters, the comparison of a number of different weight-evolution schemes, and the testing of some initial LSMC results against the harmonic approximation to check the integrity of our calculations. The second phase of the investigation is concerned with determining which of the techniques described in chapter 5 are the most effective in terms of extrapolating to coexistence (§6.6.2) and tracing the coexistence curve (§6.6.3). Having outlined our preferred techniques, we then use these methods to calculate the phase diagram of the Lennard-Jones solid in §6.6.4. Finally, we examine the reliability of our results in the light of our choice of truncation scheme (§6.6.5), before going on to discuss the possible consequences of our results in §6.7 & §6.8.

6.6.1 Initial investigations

Acceptance rates

As for the hard-sphere system (§4.10.1), we wish to adjust the simulation parameters, i.e. the sizes of the particle and volume moves, in order to maximise the statistical efficiency of the simulation. As before, we use the autocorrelation time of the lattice-switch order parameter as our guide. Figure 6.13 illustrates how both the autocorrelation time (τ) of $\mathcal{M}$ and the acceptance rates (a_r & a_v) depend on the particle and volume move parameters (Δr & Δv, respectively).

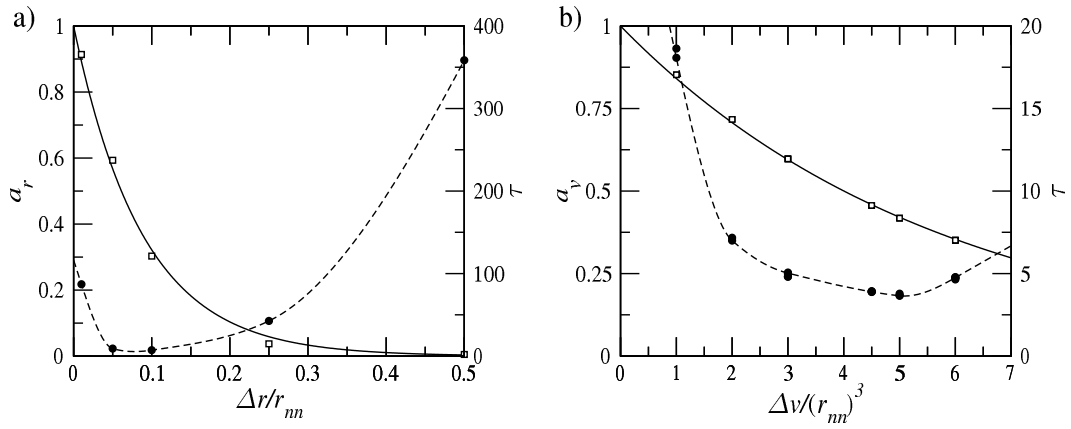


Figure 6.13: Plot of the acceptance rates and autocorrelation times of $\mathcal{M}$ for a single-phase simulation of $N = 6^3$ Lennard-Jones particles at $\tilde{\mathrm{p}} = 0$ and $\tilde{T} = 0.285$. (a) As the particle-move parameter (Δr) is increased, the acceptance rate (a_r) falls exponentially (data and fitted curve shown as open squares and a solid line). The minimum in the autocorrelation time (τ, shown as filled circles with a fitted spline) occurs when the acceptance rate is about one third. (b) Similar plot for the volume-move Δv dependence. In this case, the minimum autocorrelation time of $\mathcal{M}$ corresponds to an acceptance rate of approximately one half.

For the two examples given in fig. 6.13, we find the minima in the autocorrelation times occur for acceptance rates of $a_r \sim$ 30-40% and $a_v \sim$ 45-55%. This result is in broad agreement with those for hard-spheres (§4.10.1), and further investigation has shown that this conclusion appears to be at least approximately valid for the entire range of temperatures and densities/pressures of interest. It is tempting, therefore, to suggest that this may be a quite general property of this class of Monte Carlo simulation. Unfortunately it has not been possible to investigate the validity of this claim in any detail during the course of this work.

However, if we accept that this 'acceptance-rate rule' is at least approximately true, we can use it as an estimator of the statistical efficiency of a simulation. By tweaking the values of Δr and Δv so that the acceptance rates are approximately $a_r \sim 35\%$ and $a_v \sim 50\%$, we can be reasonably sure that the autocorrelation time will be minimised (or at least, acceptably low). This allows our Monte Carlo simulation to be rapidly self-adjusting, automatically homing in on the optimal acceptance rates

without the costs (i.e. the large volumes of data) associated with an autocorrelation-time calculation. Precisely how to do this is also indicated in fig. 6.13, in that the plots of the acceptance rates against the random-walk parameters can be closely approximated by an exponential function of the form,[10]

$$a_X = \exp\left[-\frac{\Delta X}{C}\right] \ . \tag{6.77}$$

This information allows an iterative scheme to be used to home in on the desired acceptance-rate values. While carrying out a number of sweeps (~ 100) using some random-walk parameter ΔX, we can measure the acceptance rate $a_X(\Delta X)$. This can be used to determine the value of C in eqn. 6.77, from which the value of ΔX for which a_X is optimal can be estimated. This automatic adjustment scheme can be repeated until the value of ΔX becomes stable, which usually occurs within one or two iterations. This scheme was found to work very well, quickly yielding acceptable autocorrelation times over the entire range of temperatures, densities and pressures required.

Evolving the weight function

As mentioned in §6.5, the softness of the interaction means that there is always some chance that the lattice-switch move may be accepted without biasing the simulation. Indeed, for the smallest system ($N = 6^3$), it was found that the probability of a spontaneous switch was sufficiently large that a reasonable estimate of the free-energy difference could be determined in a reasonable time in a simple Boltzmann importance-sampling simulation. However, to maximise the efficiency of these simulations, and for the larger system sizes, it is necessary to use extended sampling.

The hard-sphere work (§4.10.5) has shown that the transition-probability methods are significantly more efficient that the visited-states methods. For this reason, VS iteration was not used here. However, a number of different TP algorithms were compared; the multi-initialised (MI), mobile-barrier (MB) and strong-sampling (SS) techniques (as described in §3.6.2). All of these algorithms attempt to ensure that the full range of macrostates will be explored, and produce an initial estimate for the weight function which can then be used to kick-start an MCMC simulation. The efficiency of these weight-generation mechanisms can be determined by comparing the initial estimates for $P(\mathcal{M})$ (and thus, $\eta(\mathcal{M})$) against the 'best estimate' taken via MCMC simulation. As an example, figure 6.14 shows the initial estimates of $P(\mathcal{M})$ and $\eta(\mathcal{M})$ for a system of $N = 6^3$ Lennard-Jones particles generated using the three different techniques.

All three techniques used the same total simulation run-time of 10^6 MCS. For the strong-sampling simulation, a locking-time of $n_{lock} = 1$ MCS was used in one long simulation run. The MI simulation

[10] The exponential form of the acceptance rate has been noted by other authors examining unrelated model systems ([43] and references therein).

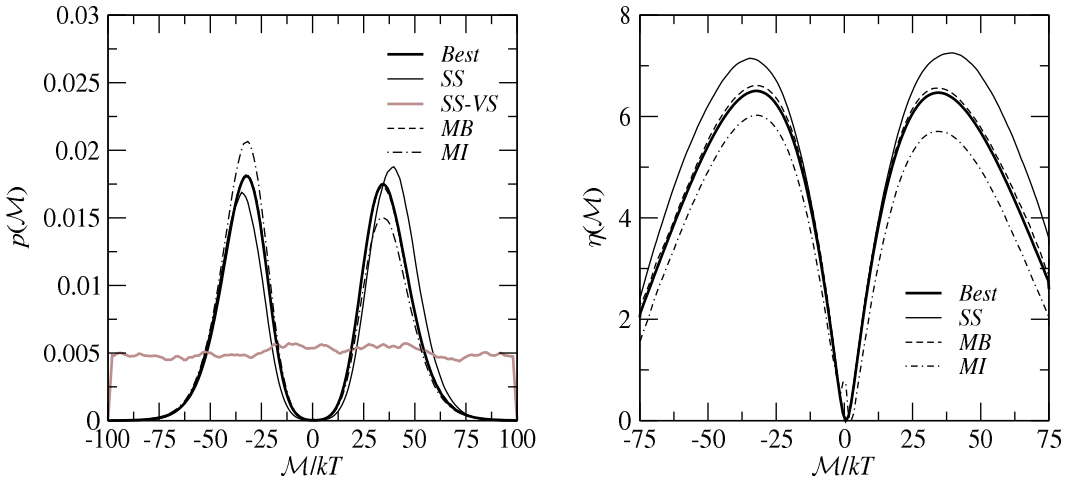


Figure 6.14: Comparison of the probability distributions and weight-functions generated using three different TP techniques. The system consisted of $N = 6^3$ particles, under constant pressure and temperature ($\tilde{\mathrm{p}} = 0$, $\tilde{T} = 0.285$). All estimates are based on runs of 10^6 MCS. The best estimate weight function was determined by using the strong-sampling estimate for $\eta(\mathcal{M})$ to perform a MCMC simulation during which the transition-probability matrix was measured anew. The weight function used a unit discretization in $\mathcal{M}$, and the histogram was 200 bins wide. The grey line (denoted SS-VS) shows the distribution of macrostates actually visited during the strong-sampling calculation, which is indeed very nearly flat.

consisted of 25 runs (all beginning in the 'cold' configuration, where all particles lie upon their lattice sites and thus $\mathcal{M} = 0$), each of 40,000 MCS in length. The MB estimate consists of two runs (one locked into the *fcc* phase, and one in *hcp*), each of length 5×10^5 MCS, with the barrier being moved outward once every 5×10^3 MCS.

Clearly, both the MI and SS weight functions are relatively poor in comparison with the MB estimate. In both cases, this poor estimation appears to stem from the short local-equilibration times allowed by these two tests. In the case of MI, the major problem is the shape of the weight function in the $\mathcal{M} \sim 0$ region, and it is in this region that the system would be most poorly equilibrated with respect to the initial perfect-lattice microstate. For the SS estimate, the *shape* of the function is correct, but the severity of the probability-barrier has been over-estimated. The SS run only allowed ~ 1 MCS for equilibration in each macrostate, which is of the order of the autocorrelation time of the $\mathcal{M}$-exploration. The MB method, however, used a very long blocking time (of $\sim 10^3$ MCS), which allows ample equilibration at every stage. This supposition was confirmed by allowing a significantly longer locking-period in a strong-sampling simulation.

However, this longer locking-time means that in order to bring the strong-sampling estimate up to the quality of the MB exploration, a longer overall run-time must also be used. This is because of the rate at which the strong-sampling technique explores the macrostate-space. The random-walk of the strong-sampling algorithm only explores the macrostates diffusively, whereas the MB algorithm systematically

explores the entire range of $\mathcal{M}$. As the rate of diffusion in the strong-sampling is rather slow, the systematic approach can explore a wider range of $\mathcal{M}$, with ample local-equilibration time, significantly faster than the strong-sampling algorithm.

There is, however, one advantage to using the strong-sampling approach over both of the others; simplicity. Once the algorithm has been implemented, the only free-parameter of the process is the locking-time, and as long as that is set to $\sim 1MCS$ or more, a *reasonable* estimate of the weight function can be found quickly. While the other methods can get better estimates of the weight-function in the same run-time, they both require more careful calibration. For MI, this involves investigating the timescale for full equilibration, to ensure that the length of the individual runs is not so long that the $\mathcal{M} \sim 0$ region is being ignored, but not so short that each run never equilibrates fully. In the MB case, the locking-time must be chosen (which depends on the chosen discretization scale as well as the total run-time), and the estimates of the *fcc* and *hcp* 'sides' of the distribution must be estimated separately and then spliced together. Because SS always works straight away, with essentially no calibration, it was used for all the initial estimate required during the Lennard-Jones work. Also, it was found that the weight function estimated for one set of external conditions was suitable for quite a wide range of temperature and densities/pressures (see below). Therefore, the efficiency with which we can estimate an initial weight function is not crucial to the efficient exploration of the phase diagram.

Temperature & density dependence

The temperature dependence of the weight function is illustrated in figure 6.15, which also compares the results with a weight function from the hard-sphere system for a similar density. It is interesting to note that the positions of the maxima for the $N = 6^3$ hard-sphere and Lennard-Jones systems (at comparable densities) lie in similar ranges of $\mathcal{M}$, although the softness of the Lennard-Jones interaction means that the probability trough between the two phases is significantly shallower. It seems reasonable to assume that for soft potentials, the energy barrier associated with the overlap between two mismatched particle displacements will be of the order of kT, i.e. of the order of unity in the reduced units of $\mathcal{M}/kT$. In the hard-sphere case, the cost of the switch was measured in terms of overlapping pairs of particles, which will also be unity for any mismatched particle-pair. Therefore, it is perhaps not too surprising that for similar densities, the average (equilibrium) value of $\mathcal{M}$ is similar in both phases.

Figure 6.15(b) shows that the inter-phase probability barrier increases with temperature, and that the position of the weight function's central minimum is also temperature dependent. This latter dependency only occurs in the NPT ensemble, and is illustrated in more detail by the close-up in figure 6.16(a), showing that the minimum moves towards the *fcc* side of $\mathcal{M}$ as the temperature is decreased. To un-

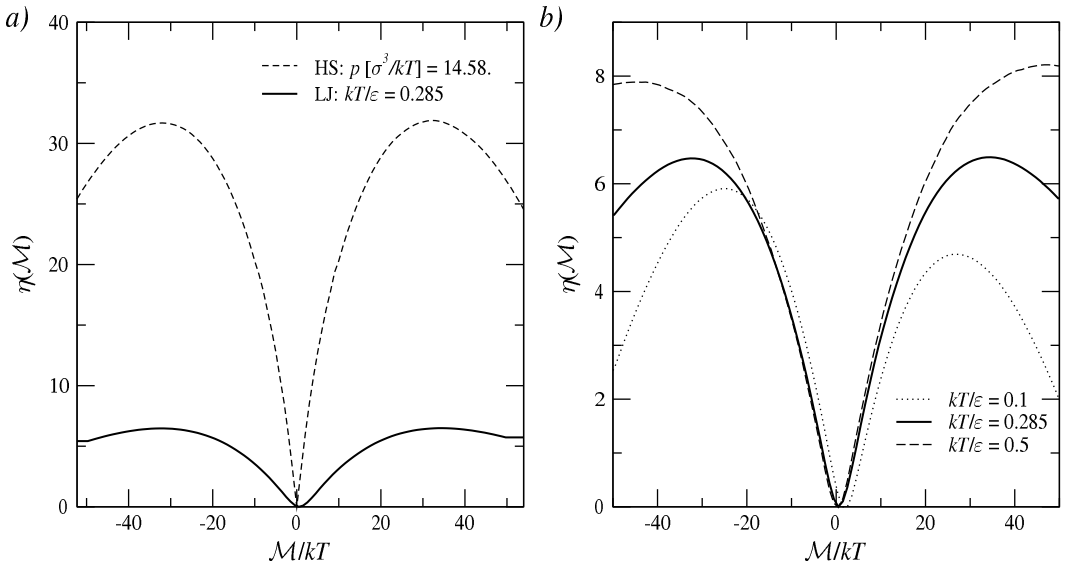


Figure 6.15: Examples of 'best-estimate' weight functions for systems of $N = 6^3$ particles. (a) The zero-pressure, $\tilde{T} = 0.285$ weight function is plotted atop of the weight function of the hard-sphere solid at $\tilde{\mathrm{p}} = 14.38$, for which the densities of the two systems are comparable. (b) Weight functions for the $\tilde{\mathrm{p}} = 0$ Lennard-Jones system at three different temperatures. Note how both the relative weighting and the position of the central minimum vary with temperature.

derstand this, we first recall that our chosen order parameter (eqn. 6.44) is based upon the difference in *excitation* energy between *fcc* and *hcp*, whereas in the NPT ensemble the probability of accepting the move (eqn. 6.43) depends upon both the ground-state and excitation-energy differences. In fact, the probability of accepting the lattice-switch move will reach a maximum (unity) when the difference in excitation energy is equal and opposite to the ground-state energy difference:

$$\langle \Delta E_{LS} \rangle = 0 \,, \tag{6.78}$$

$$\langle \Delta \mathcal{E} \rangle = \langle \Delta E_{gs} \rangle \,. \tag{6.79}$$

The order parameter has been plotted in simulation units, i.e. in units of kT such that

$$\tilde{\mathcal{M}} = \frac{\mathcal{M}}{kT} = \frac{\Delta \mathcal{E}}{kT} \,, \tag{6.80}$$

and so the maximum in the acceptance probability occurs when

$$\langle \tilde{\mathcal{M}} \rangle = \left\langle \frac{\Delta E_{gs}}{kT} \right\rangle \,. \tag{6.81}$$

The ground-state energy difference is positive (favouring *hcp*), and so the maximum acceptance probability occurs in the $\tilde{\mathcal{M}} > 0$ range, and moves to larger values of $\tilde{\mathcal{M}}$ as the temperature tends towards zero. Figure 6.16(b) shows the acceptance rate of the lattice-switch move as a function of $\mathcal{M}$, illustrating that the maximum acceptance rate coincides with the minimum of the weight function. This makes sense, in that one might expect that the most unlikely macrostate, i.e. the macrostate that the algorithm

must 'push hardest' toward in order to facilitate the switch, is also that for which the switch is most likely to occur. However, the precise reason why the minimum in $P(\mathcal{M})$ should correspond to the maximum in the acceptance probability is not clear.

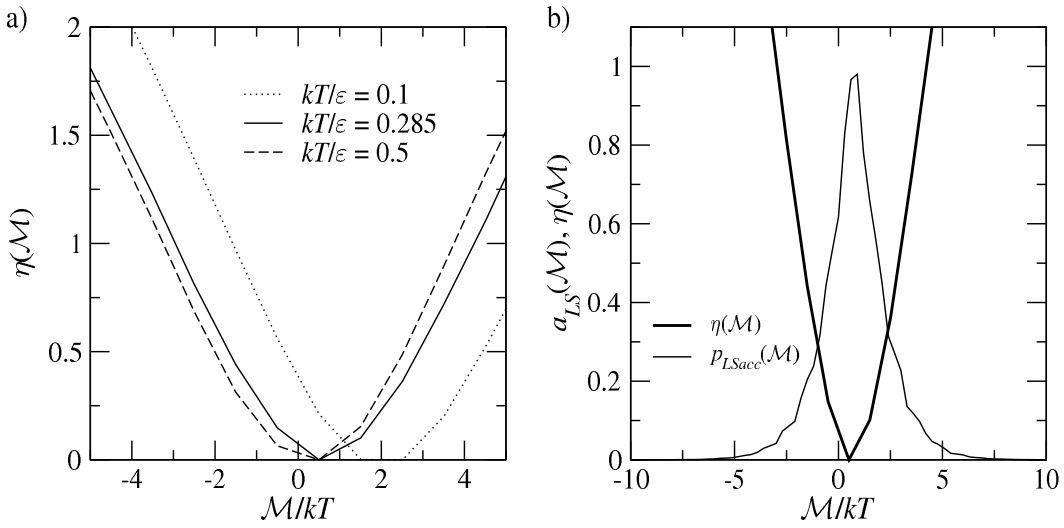


Figure 6.16: Close-up of the switch-zone for the $N = 6^3$, $\tilde{\mathrm{p}} = 0$ Lennard-Jones system. (a) The position of the minimum in the weight function is dependent upon the temperature, tending toward zero as the temperature is increased. (b) A plot of the minimum of the weight function (at $\tilde{T} = 0.285$) superimposed upon the acceptance rate of the lattice-switch move (a_{LS}) as a function of $\mathcal{M}$.

The temperature dependence of the probability distribution $P(\mathcal{M})$ itself is shown in figure 6.17, for which the change in relative weighting of the two structures can be seen more directly. This illustrates the broadening of both the *fcc* and *hcp* peaks as the temperature is increased, as well as the shift in relative weighting as the temperature passes through the coexistence point ($T_c \sim 0.285$).

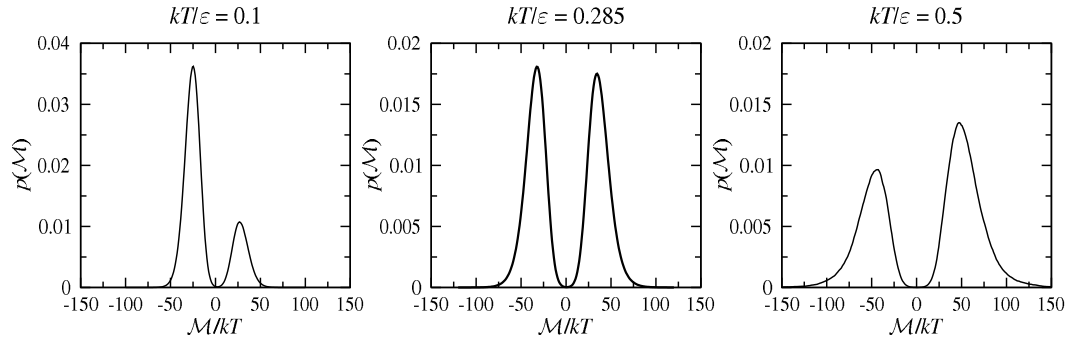


Figure 6.17: Temperature dependence of the probability distribution of $\mathcal{M}$ for a system of $N = 6^3$ Lennard-Jones particles at zero pressure. Note the difference in the vertical scale between the $\tilde{T} = 0.1$ plot and the others.

The temperature and density dependence of the distribution $P(\mathcal{M})$ in the NVT ensemble is illustrated in figure 6.18. Firstly, we note that while the range of accessible temperatures changes significantly between the low and high density systems (the melting temperature changes from $\tilde{T}_m \sim 1$ to $\tilde{T}_m \sim 10$), the

general form of the temperature dependence remains broadly the same. As the temperature is increased, the distribution of $\mathcal{M}$ becomes broader and flatter, but in contrast to the NPT case (figure 6.17) the relative weighting of the phases shown in figure 6.18 appears to change relatively slowly. This is because the ground-state energy difference is not included in the NVT calculations, and illustrates the fact that the relative *excitational* configurational weights of the two phases remains approximately constant (in units of kT) as the temperature is increased, as predicted by the harmonic approximation (eqn. 6.27). Therefore, the rapid change in relative weighting shown in figure 6.17 is primarily due to the competition between the ground-state and temperature-dependent contributions to the free-energy difference.

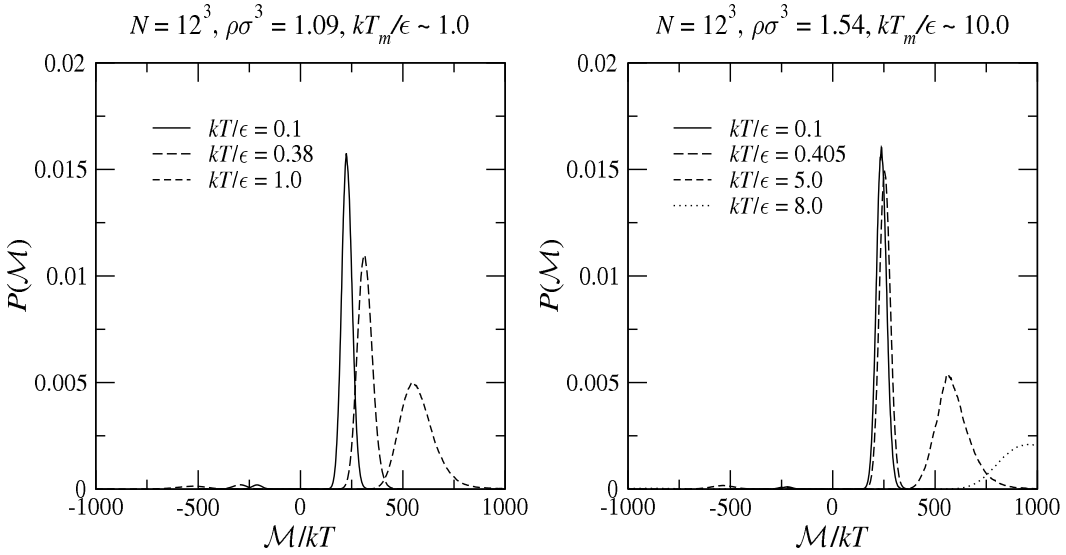


Figure 6.18: Density dependence of $P(\mathcal{M})$ in the NVT ensemble, for two different densities ($\tilde{\rho} = 1.09$ and $\tilde{\rho} = 1.54$). The system consisted of $N = 12^3$ Lennard-Jones particles. As the ground-state energies are not included in the NVT calculations, these figures only shown the temperature-dependent contribution to the free-energy.

Convergence to the harmonic regime

At low temperatures, the harmonic approximation should form an accurate description of the system, and so in this limit, the lattice-switch code should produce the same results. To test this, a series of NVT simulations were performed for a system of $N = 12^3$ particles, at two different densities, for a range of temperatures. The values of the free-energy difference between *fcc* and *hcp* calculated using LSMC are compared against the harmonic predictions in figure 6.19. The low-temperature convergence is clear, and given that the harmonic code has already been extensively tested (§6.4.4), these results give us additional confidence that the lattice-switch calculation has indeed been implemented correctly.

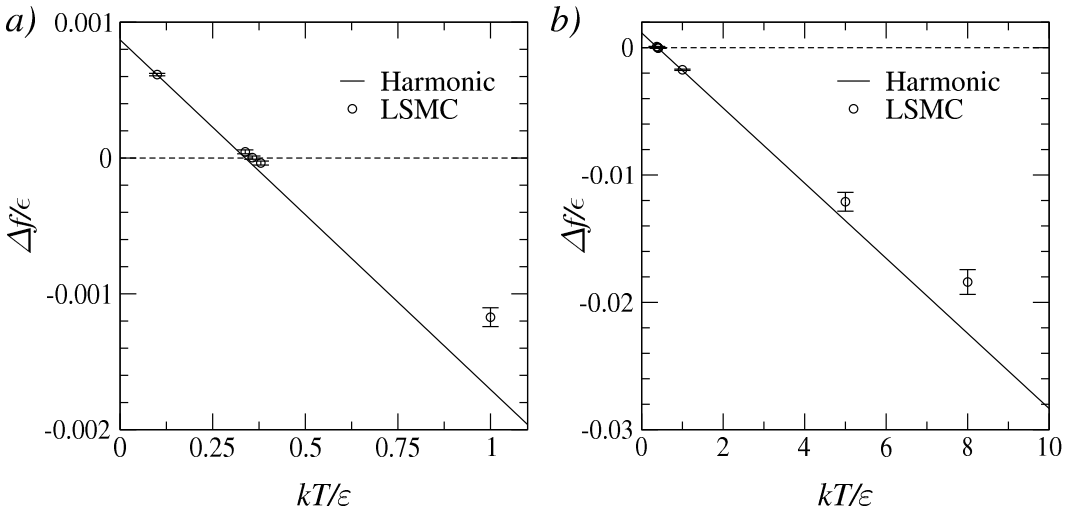


Figure 6.19: The free-energy difference $\Delta f = f(fcc) - f(hcp)$ as a function of reduced temperature at (a) a density of $\rho = 1.092$, close to that corresponding to zero pressure, and (b) a higher density of $\rho = 1.538$, corresponding to the maximum in Δe_{gs} (figure 6.5). The system was composed of $N = 12^3$ particles, and each run was 10^6 MCS in length. The convergence to the harmonic approximation in the low temperature limit is clear.

Finite size effects

To investigate the importance of finite-size effects in this work, the transition temperature was estimated for a (fixed) density of $\tilde{\rho} = 1.09$ for systems of $N = 6^3$, $N = 12^3$ and $N = 18^3$ particles. Figure 6.20 shows the *fcc–hcp* transition temperature as estimated from the free-energy and the enthalpy differences between the structures (from a single LSMC two-phase simulation) using the Newton-Raphson algorithm (§5.3). This plot assumes that the finite-size effects are linear in N^{-1}, as did figure 4.22 for the hard-sphere system. This assumption is very common in the analysis of condensed-matter simulation results, see for example [115, 116], although the rationale behind this particular choice of power-law is not clear. However, for the Lennard-Jones system at least, this assumption appears to be valid to within the accuracy of our simulations. Furthermore, the plot indicates that the $N = 12^3$ results are statistically indistinguishable from those of the $N = 18^3$ system

6.6.2 Extrapolating to coexistence

The problem of extrapolating to coexistence from some arbitrary point in the phase diagram was attacked using the NR and SHE techniques (explained in §5.2 & §5.3). Figure 6.21 gives a typical example of the relative merits of the two techniques. From a single simulation, held at zero pressure and far from the transition temperature, the NR technique gives a good estimate of the position of the transition point. As mentioned in §5.3, is it possible to use two separate simulations (one in each phase) to evaluate the enthalpy difference between *fcc* and *hcp*. In general, this was not found to be necessary, as the

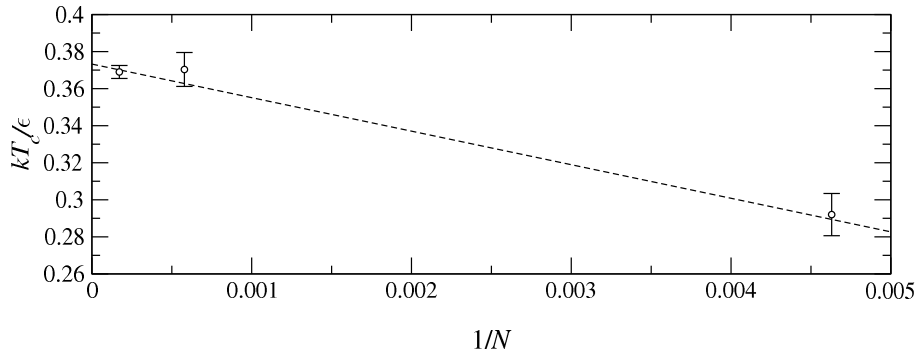


Figure 6.20: System size dependence of the *hcp* to *fcc* transition temperature ($\tilde{T}_c = kT_c/\epsilon$) for a fixed density of $\tilde{\rho} = 1.09$, plotted as a function of N^{-1}.

enthalpy difference inferred from a single (dual-phase) LSMC simulation was found to be sufficiently accurate. Single-histogram extrapolation is, in comparison, extremely poor. The errors associated with this data-wise extrapolation grow rapidly as the range of the extrapolation is increased, and no sensible estimate for the transition temperature could be determined from these results. Generally, we find that the underlying assumption of the NR calculation is good (i.e. that the free-energy difference curve is approximately linear, as indeed it is whenever the harmonic approximation is reasonable), and leads to reliable estimates of the free-energies and thus, the transition temperature T_c.

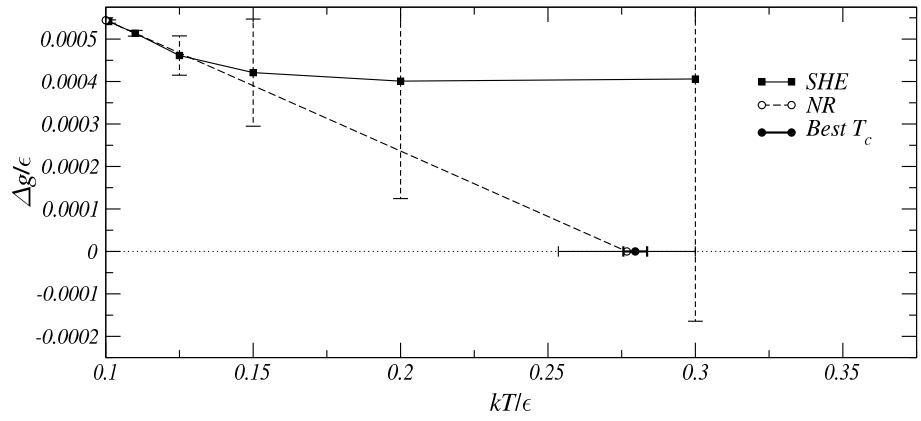


Figure 6.21: Comparison of the NR and SHE extrapolation techniques, in an attempt to predict the zero-pressure transition temperature for a system of $N = 6^3$ Lennard-Jones particles. A single LSMC simulation (of 10^7 MCS in length) was performed at a temperature of $\tilde{T} = 0.1$, and the energy, volume and order parameter ($\mathcal{M}$) of the system were recorded every 10 MCS. The canonical averages of the free-energy and the enthalpy differences between the phases were determined from the sampled data and used for the NR extrapolation (dashed line, open circles). The same data was used to estimate the free-energy difference between the phases via single-histogram extrapolation (solid line, filled squares). Also shown is our best estimate of the true transition temperature for this system (thick line, filled circle).

The contrast between the NR and SHE extrapolations becomes even more pronounced as the size of the system is increased, as the fluctuations in the extensive properties of the system become small on the

scale of the averages of those properties. Broadly speaking, as N increases the size of the errors associated with a given single-histogram extrapolation (of $\Delta\beta$, say) grow as the square-root of N. This sharpening of the enthalpy and free-energy also means that for the same amount of run-time, the estimates for the canonical averages become more accurate. This in turn means that as long as the assumption of the linearity of Δf is reasonable, the NR technique becomes *more* accurate as N increases. For these reasons, NR extrapolation was used for the initial estimation of the Lennard-Jones transition curve.

6.6.3 Tracing the coexistence curve

While the SHE technique is also applicable to this stage of the coexistence evaluation, it was found to be little more reliable here than in the case of extrapolation to coexistence. The range of extrapolation was so poor that no useful predictions could be made using this technique. For example, it was hoped that the SHE approach may allow us to predict a new weight function for a neighbouring point on the coexistence curve. Unfortunately, the (rather noisy) SHE extrapolated weight function was found to give significantly poorer results than simply using the original (un-extrapolated) weight function.

To determine whether Gibbs-Duhem integration is more or less efficient than the LSMC predictor-corrector, the following test-case was used. Firstly, LSMC & NR calculations were used to predict the *fcc-hcp* transition temperature of a $N = 6^3$ Lennard-Jones solid at a pressure of $\mathrm{p}\sigma^3/\epsilon = 771.0$. At this high pressure, the results should be in reasonable agreement with the harmonic approximation, and will therefore provide an independent test of the results. This gave a transition temperature of $T_c = 0.048(1)$, which is indeed in agreement with the harmonic prediction of $T_H = 0.050244$ (based on the assumption that the densities of the two phases are the same at this pressure). Both the GDI algorithms and the lattice-switch predictor-corrector were then used to move along the coexistence curve in steps of $\Delta\mathrm{p} = -25$. The results from these calculations are presented in figure 6.22, along with the coexistence curve from the harmonic approximation. For the high-order (3PEC) integration, the simulation runs for each of the two phases both lasted 5×10^6 MCS. The low-order (PEC) integration was carried out using the gradients determined during the 3PEC integration. The PCLS simulations were 5×10^6 MCS long, and during the course of the integration it was found that just one simulation was sufficient to determine each transition temperature accurately. Therefore, the PCLS method needed only half the run-time of the GDI approach.

There are two main conclusions to be drawn from figure 6.22. Firstly, both GDI algorithms perform badly in comparison with PCLS, which gives smaller errors and is in good agreement with the harmonic prediction. Secondly, of the two GDI calculations, the low-order PEC algorithm performs slightly better than its higher-order (3PEC) counterpart. To understand why, we must examine the observables used

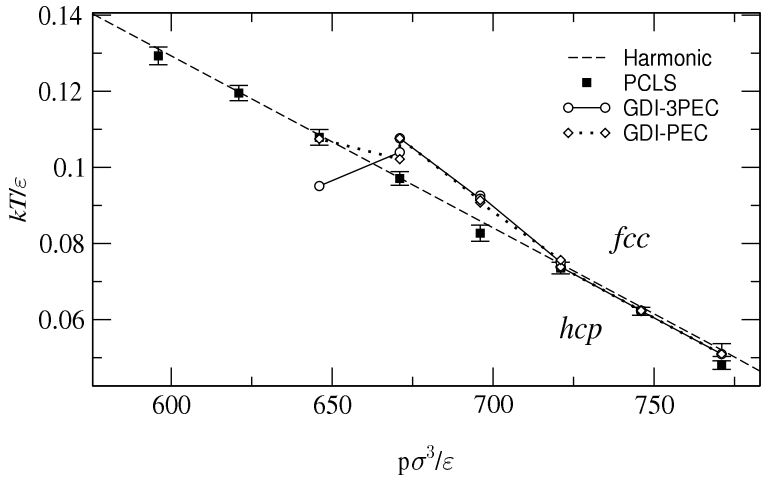


Figure 6.22: A high-pressure, low-temperature portion of the phase-diagram is shown. The coexistence curves predicted via the PEC and 3PEC GDI algorithms, the PCLS calculation, and from the harmonic results are superimposed. The GDI calculations consisted of 4 PEC cycles (requiring 5 gradient evaluations) and ended with a final prediction step for $\tilde{\mathrm{p}} = 646$.

to determine the value of the gradient, i.e. the enthalpy difference and volume difference between the phases (c.f. eqn. 5.40). Figure 6.23 shows the values of Δe and Δv used by the GDI calculations as it moves along the coexistence curve. Figure 6.23(a) also shows the enthalpy difference in the low-temperature limit, where the behaviour is controlled purely by the ground-state energies,

$$\Delta h = \Delta e_{gs}(\tilde{\rho}) \ , \tag{6.82}$$

where we have assumed that because the densities of the two structures are very similar, the contribution to the enthalpy due to the volume difference (in the PV term) is rendered negligible.

Clearly, the value of Δe has been reasonably well determined and, noting that temperature goes down from left to right, agrees well with the approximate form given in eqn. 6.82. This reflects the fact that the enthalpies of the two structures are largely determined by the ground-state energies over this temperature range. However, the Δv plot shows that the errors in the determination of this difference are unacceptably large. At the lower temperatures (right-hand side), the volume difference is reasonably well determined (to within about 10%), but as the temperature increases (from right to left), the errors grow to around 30%. This implies that as the temperature is increased, and the fluctuations in the densities become larger, the volume *difference* between the phases increases too slowly to compensate for these increased fluctuations.

Given that the volume difference is badly measured, but the enthalpy difference is not, the reasons for the success of PCLS and the failure of the two GDI calculations becomes clearer. The higher-order PC algorithm essentially fits a polynomial to the last few gradient evaluations, and the random variation

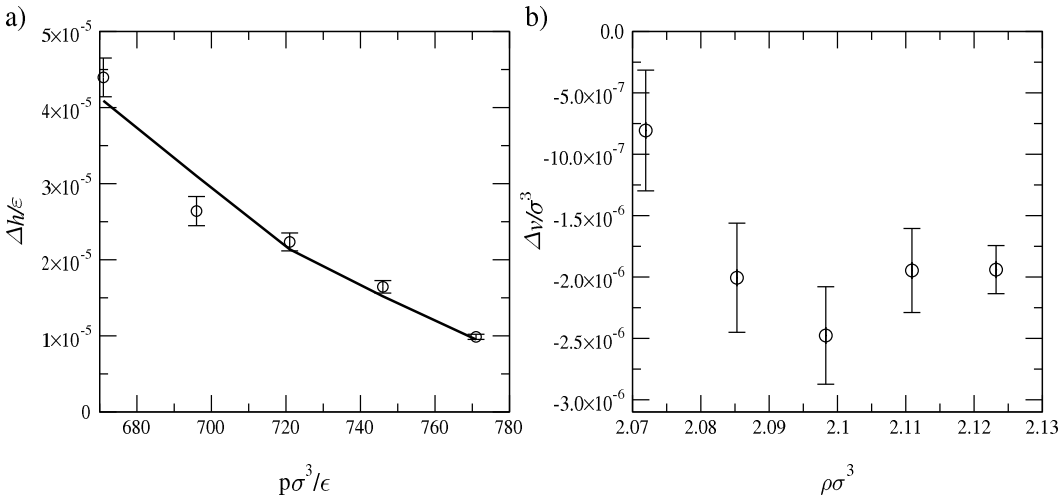


Figure 6.23: The two components of the gradient along the high-pressure low-temperature coexistence curve shown in fig. 6.22. (a) The enthalpy difference plotted against pressure, along with the approximate (low T) analytical form mentioned in the text. (b) The volume difference plotted against density. Note that in both plots, the transition temperature drops from left to right.

of the volume difference (fig. 6.23(b)) has produced an apparent 'trend' downwards in the gradient (upwards in Δv). This explains why the 3PEC algorithm has predicted a change in sign of the gradient (final prediction step of fig. 6.22), whereas the PEC has not. The polynomial extrapolation is based more on the *noise* in the gradient rather than the gradient itself. For this reason, the 3PEC algorithm appears to be a little less stable than the PEC scheme, at least at the predictor stage.

In contrast to the GDI integrations, the PCLS calculation is based on Δe and Δf, and so avoids the errors associated with attempting to determine Δv. Any *predicted* transition temperature that moves away from the *actual* transition temperature will only slow the correction stage, and as both the enthalpy and free-energy differences are relatively well-determined the Newton-Raphson correction algorithm can work accurately. Of course, if the PCLS calculation had instead been set up to integrate along the temperature axis, measuring the volume difference along the way, then the results would have been at least as unreliable as those presented for the Gibbs-Duhem integration. In short, the GDI approach requires that *both* components of the gradient are accurately determined, whereas the PCLS method can be reformulated to work in terms of whichever component is easiest to measure. We have the freedom to choose, and in this case this is most definitely an advantage.

To conclude this comparison of the coexistence-tracing techniques, the effect of the system size should be examined. As we move from $N = 6^3$ to $N = 12^3$ the macroscopic properties become more sharply defined, and measuring the enthalpy and volume difference via two separate simulations becomes easier. However, in an LSMC simulation, the increase in N means that the simulation spends more time in low-Boltzmann weight states (as the two peaks of $P(\mathcal{M})$ become more sharply defined and the gap between

them widens). Note also that for higher N, the weight function changes more rapidly as we moves along the coexistence curve. For $N = 216$, using the same weight function over a pressure range of $\mathrm{p} \rightarrow \mathrm{p} \pm 25$ works acceptably well, but for $N = 1728$ the weight function is only reasonably constant over a range of about $\mathrm{p} \rightarrow \mathrm{p} \pm 5$. Therefore, as N increases, GDI will eventually win. Nevertheless the run-times required by GDI were found to be prohibitively large for both $N = 6^3$ and $N = 12^3$, and so the majority of the calculations were carried out using the PCLS algorithm.

6.6.4 Phase diagram of the Lennard-Jones solid

Having chosen to use the Newton-Raphson method to approach the phase boundary, and the PCLS method to move along it, the shape of the entire transition curve was calculated, the results of which are shown in figure 6.24. The includes results from the $N = 6^3$ and $N = 12^3$ systems, from both the NPT and NVT ensembles, along with the phase curves predicted using the harmonic calculation (under the assumption that the two structures have the same density, as in the NVT LSMC calculations).

The most complete data set was taken for the smaller system-size, as this system could be explored at relatively little computational cost. This showed that the agreement between the NPT and NVT ensembles was good, and that the transition curve rapidly converged onto the harmonic curve as the pressure/density was increased. This knowledge allowed the investigation of a $N = 12^3$ (NPT) system to focus the available computational resources on the low-pressure region, where anharmonic effects are most significant. First, the zero-pressure transition point was determined (via NR), and then the transition curve was traced (using PCLS) until it converged upon the harmonic result. The $\tilde{\mathrm{p}}$-$\tilde{T}$ diagram also includes the melting curve, showing that the *fcc* phase is always stable near melting. Also shown is the fitted transition curve based on the analytical expression (eqn. 6.41) for the magnitude of the anharmonic contribution to T_c given in §6.4.6. The level of agreement between this analytical expression and the $N = 12^3$ NPT LSMC results is explored in more detail in figure 6.25.

These results are in qualitative agreement with those of Barron & Domb [98], and in reasonable quantitative agreement with the results of Galashev [104]. The latter publication estimates the zero-pressure transition to occur at $\tilde{T} = 0.4$ (no error supplied), whereas we find the transition temperature to be $0.361(2)$. Unfortunately, no other quantitatively comparable results were available beyond those already explored during the testing of the ground-state and harmonic calculations. As indicated earlier, this was mainly because those papers which have addressed this problem did not give sufficient information about their simulations to reproduce their results, or used a misleading truncation scheme.

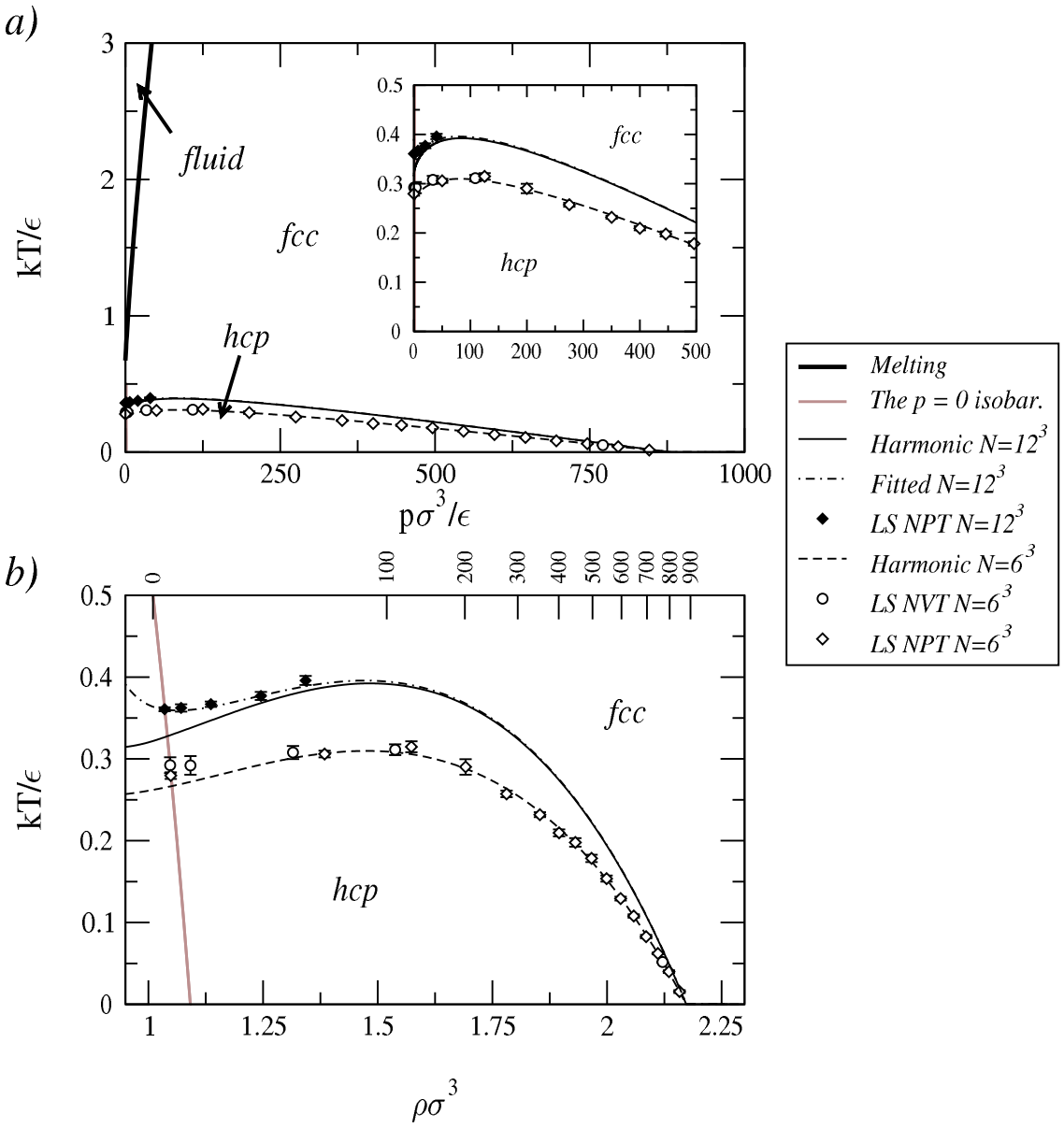


Figure 6.24: Phase diagram of the Lennard-Jones solid, showing the LSMC results and the results of the harmonic calculation. (a) High-temperature phase behaviour in the $\tilde{\mathrm{p}}$-$\tilde{T}$ plane, along with the melting curve take from [93]. The inset shows the low-pressure behaviour in more detail. (b) Low-temperature $\tilde{\rho}$-$\tilde{T}$ behaviour. The grey $\tilde{\mathrm{p}} = 0$ isobar shown in the $\tilde{\rho}$-$\tilde{T}$ plane indicates the lowest physically accessible ($\tilde{\mathrm{p}} \geq 0$) density, which curves toward lower densities as the temperature is increased (reflecting the rate of thermal expansion). The upper horizontal axis of this plot shows the dependence of the coexistence pressure upon the density, reflecting the form of the equation of state along the $N = 12^3$ phase boundary.

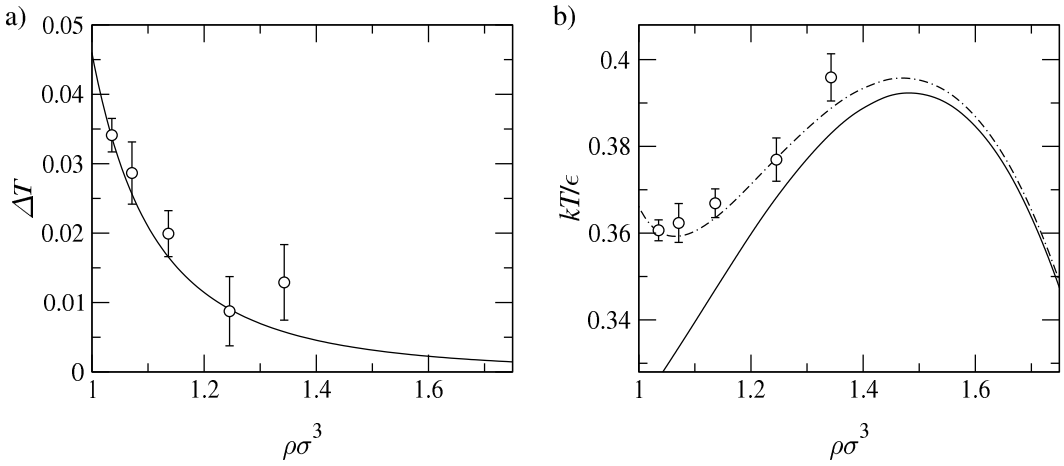


Figure 6.25: Comparison of the estimated form of the anharmonic contribution to the transition temperature with the LSMC results. (a) Plot of the difference between the harmonic and LSMC transition temperatures, along with the fitted analytic curve (eqn. 6.41). (b) A close-up of the anharmonic region of the transition curve (fig. 6.24), showing the fitted curve as the sum of the curve shown in (a) and the harmonic prediction. The single free-parameter of the analytic form was fixed by fitting to the zero-pressure transition point. This fitted form is accurate to within the precision of our estimates for the transition temperature.

6.6.5 The consequences of truncation

As explained in §6.5.4, the evidence from the harmonic approximation does not prove that the free-energy cost associated with truncating the interactions is negligible. Therefore, this free-energy cost was investigated for a range of system parameters, using the 'virtual Hamiltonian' calculation explained in §6.5.4. The results from both the NVT and NPT ensembles are shown in figure 6.26, along with the results for the same free-energy difference determined via the harmonic approximation. Note that as the difference between the densities of the two structures (in an NPT simulation) is so difficult to determine that a comparison with the harmonic approximation is only possible at zero pressure. This is because at low pressure and temperature, the densities of the two structures converge onto the ground-state densities, which are known from the positions of the minima (with respect to density) in the ground-state energy function (see §6.3).

For the smaller system ($N = 6^3$), the high (fixed) density results (denoted as NVT, $\tilde{\rho} = 2$) are in excellent agreement with those from the harmonic approximation. For the $\tilde{\mathrm{p}} = 0$ behaviour, the convergence towards the harmonic approximation in the limit of low temperature is clear. For the larger ($N = 12^3$) system, very similar behaviour is observed. The low-temperature $\tilde{\mathrm{p}} = 0$ results agree with the harmonic prediction, but at higher temperatures the free-energy difference moves away from (and becomes larger than) the harmonic result. This figure also shows the values of Δg_{nn} for a number of different pressures ($\tilde{\mathrm{p}} = 0, 200$ & 400). These results indicate that the free-energy difference associated with the truncation of the potential is largest at zero pressure, and decreases sharply as the pressure is increased.

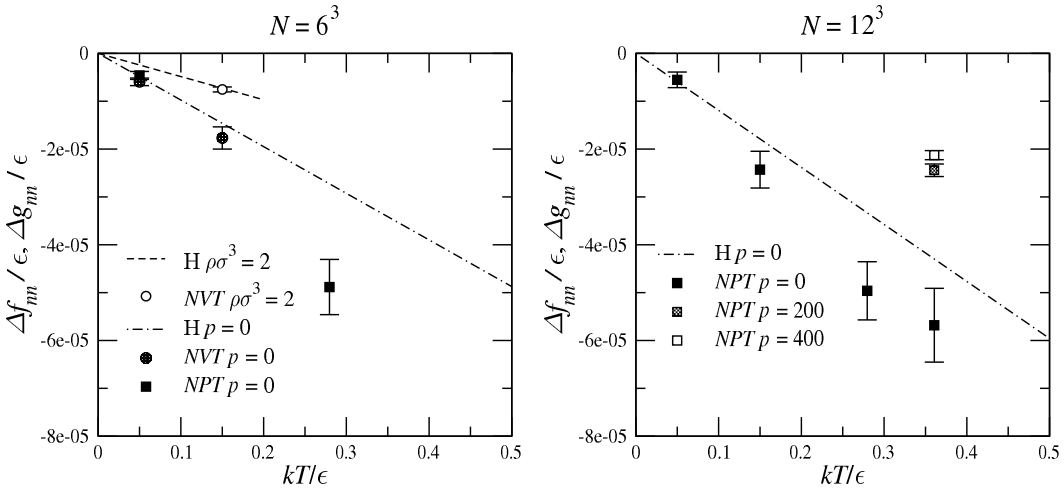


Figure 6.26: A broad selection of results showing how the truncation of the potential affects the free-energy difference between *fcc* and *hcp*. H indicates the results of harmonic calculations, where the ground-state *density* difference (§6.3) has been used to calculate the harmonic estimate of the truncation-dependence at zero pressure.

From the point of view of the validity of our results for the Lennard-Jones system, there is one central point to be drawn from all of these results. No matter what the temperature or pressure, the free-energy difference due to the truncation of the potential is *always* smaller than the statistical errors in the *fcc*–*hcp* free-energy difference upon which the prediction of the Lennard-Jones phase diagram is based (see §6.6.1). Furthermore, the errors associated with the predicted transition temperatures shown in fig. 6.24 are the combined result of the errors in both the enthalpy and free-energy differences between the phases, and so the amount by which the truncation of the potential has shifted the coexistence curve will be somewhat less than the stochastic errors shown there. Therefore, given that the finite-size effects are small (§6.6.1), the phase diagram presented in fig. 6.24 is believed to form a good representation of the phase behaviour of the classical Lennard-Jones solid.

6.7 Comparison with experiment

As mentioned at the start of this chapter, the Lennard-Jones model is believed to provide a reasonable account of the properties of the rare gas solids (RGS) [117], and so it is interesting to compare our predicted behaviour with that observed for these elements. Of course, the analysis of the Lennard-Jones system carried out here has been concerned only with the classical phase behaviour, and so cannot predict the consequences of any quantum-mechanical effects. Strictly speaking, only a full *ab initio* (electron-level) treatment can be relied upon to accurately predict the behaviour of the RGS. However, the Lennard-Jones potential is known to accurately reproduce many of the properties of the RGS [9, p. 401], and it seems reasonable to expect that the experimental results should be broadly consistent

with our predictions in the classical limit (i.e. high temperature and/or high particle mass). In fact, the temperature below which quantum-mechanical effects become significant can be estimated using the results of Salsburg & Huckaby [99]. In that paper, the free-energy difference between *fcc* and *hcp* for the Lennard-Jones system (at low pressure) is estimated (via the harmonic approximation) to be

$$\frac{\Delta f_Q}{kT} \approx \frac{\Delta e_{gs}}{kT} + 2.62 \times 10^{-3} + 4.57 \times 10^{-5} \left(\frac{\Theta}{T}\right)^2 + \ldots \tag{6.83}$$

The first term of this expansion is ground-state energy difference, and the second term is the classical result for the harmonic contribution. The final term describes (approximately) the quantum-mechanical contribution to the *fcc–hcp* free-energy difference (via the Debye temperature, Θ), which favours *fcc* and *increases* as the temperature is *decreased*. Therefore, the Debye temperature indicates the temperature at which the quantum-mechanical contribution becomes significant ($\sim$ a few percent of the harmonic result); above this temperature the behaviour is essentially classical. Table 6.2 shows the value of the Debye temperature for the heavier rare gases, along with their masses, Lennard-Jones parameters, triple-point temperatures and predicted transition temperatures and pressures based on our results for the classical Lennard-Jones solid (T_c and p_c, as defined in the caption, and as drawn in fig. 6.27).

Element	Mass [amu]	ϵ [K]	σ [Å]	Θ [K]	T_t [K]	T_c [K]	p_c [GPa]
Neon	20.2	38.5	2.786	51.5	24.55	13.8	22.1
Argon	39.9	119.8	3.405	36.6	83.85	43.5	37.7
Krypton	83.8	159.9	3.639	25.3	115.95	58.0	41.2
Xenon	131.3	220.9	3.962	20.2	161.25	80.2	44.1

Table 6.2: Table of various properties of the heavier rare gas solids. The triple-point temperatures (T_t) are taken from [118]; the empirically determined Lennard-Jones parameters (ϵ & σ) are taken from [109]. The Debye temperature Θ has been estimated from ref. [99]. Using the phase diagram we have determined for the Lennard-Jones system, we estimate the temperature of the *hcp*→*fcc* transition at low pressure (T_c in Kelvin), and the pressure of the *hcp*→*fcc* transition at $T = 0$ (p_c in Giga-Pascals).

There are two notable omissions from table 6.2, helium and radon. In the latter case, this is simply because this heaviest rare gas is too radioactively unstable for any meaningful exploration of its structural phase behaviour to be performed. In the case of helium, which is four times lighter than neon, the quantum-mechanical effects are so great that any comparison of its phase behaviour with our classical picture is almost certainly meaningless. Curiously, helium does display an *hcp* to *fcc* transition as the temperature is increased [119], but no conclusions should be drawn from this apparent qualitative agreement with the classical picture.

However, for the RGS that are shown in table 6.2 the classical picture may indeed be of some relevance.

By comparing the values of Θ with the $\mathrm{p} = 0$ transition temperature T_c, we see that the *fcc–hcp* transition of the classical system should be present for argon, krypton and xenon. For neon, quantum-mechanical effects cannot be ignored, but the following treatment seems plausible. If we assume that eqn. 6.83 forms a reasonable estimate of the low-pressure behaviour of neon, then we can expect that the quantum-mechanical effects (which favour *fcc*) will wipe out the classical *fcc–hcp* phase transition, and make *fcc* stable for all temperatures.

It should be noted that the role of quantum-mechanical effects is not only significant at low temperature; as the *pressure* is increased, the classical picture can break down completely. The electronic structure of the atoms can suddenly change, producing a metallisation transition which can strongly affect the 'equilibrium' structure. These transitions are predicted to occur at very high pressures, ~ 1 TPa for He going down to ~ 1 GPa for Xe [120]. Experiment observations of RGS have been performed at pressures up to about 100 GPa, but a metallisation transition has only been observed for xenon, at about 10 GPa [119]. Therefore, although the structural phase behaviour predicted from the Lennard-Jones potential cannot deal with this effect, it should only obscure the *fcc–hcp* transition of xenon.

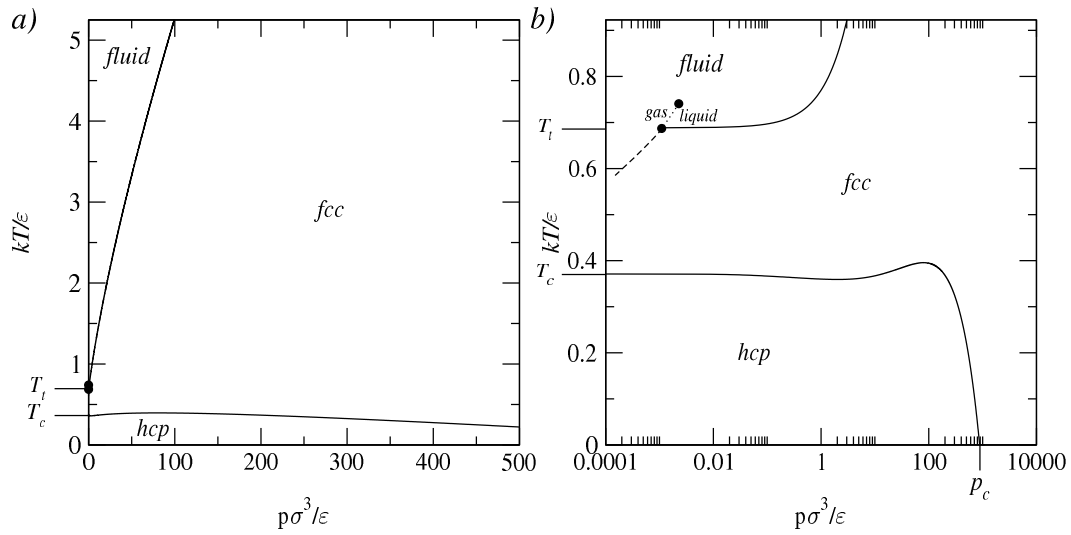


Figure 6.27: Phase diagram of the Lennard-Jones solid, as in figure 6.1, but now including our predicted *fcc–hcp* phase boundary. (a) With the pressure on a linear scale, it is clear that the *fcc* phase dominates the scene. It is always the stable phase close to melting, and indeed at any point above the triple-point temperature. (b) On a logarithmic pressure scale, the relationship between the *fcc–hcp* transition and the low-pressure sublimation curve can be seen. It shows that the structural transition will occur at temperatures below sublimation for all but perhaps the very lowest pressures, far below the $\tilde{\mathrm{p}} = 0.0001$ (~ 0.01atm) lower-limit shown here. This figure also serves to clarify the definition of T_c and p_c, as used in table 6.2.

6.7.1 Experimental evidence

Most of the experimental results from these elements can be split into two categories; high (fixed) temperature ($\sim$ 300 K) experiments where the role of pressure is examined, and low-pressure ($<$ 1atm) experiments where temperature is the controlling factor. For the high-temperature results, such as [119, pp.150–176] & [121], the prediction that *fcc* is the stable phase for all pressures for $T > T_t$ (see fig. 6.27(a)) is completely consistent with the experimental record. However, the low-pressure behaviour (fig. 6.27(b)) is somewhat less clear.

The established view is that *fcc* is always the stable structure of neon, argon, krypton and xenon, a conclusion which can be traced back to the influential Rare Gas Solids vols. 1 & 2 [122] (1976). In Chapter 9 [95], the authors state that

> "Experimentally, it is straightforward to determine the crystal structure by X-ray diffraction. It is well known that all the heavier RGS have the face-centred cubic (fcc) structure, and that there are no phase transitions to other structures in the accessible range of pressure, P, and temperature, T."

However, on closer examination, the evidence presented to back up this categorical statement seems to be less than conclusive. In the same publication, the authors also state that the stability of *fcc* is in fact a matter of opinion, and that it is "not clear" whether *hcp* is stable or not [95, p.583]. Also, from chapter 10 of Rare Gas Solids [123], (italics added)

> "The present authors feel *fairly confident* that the hcp phase of the heavy rare gas solids is always a result of either poor crystal growth, or stabilisation by impurities, or both"

There can be no doubt that the presence of as little as 1% impurities can strongly affect the structural phase behaviour [124, 125]. Note, however, that all of the research into weakly doped rare-gas solids indicates the presence of a transition from *fcc* to *hcp* as the temperature is increased towards the melting point. This is quite distinct from the pure Lennard-Jones result, where the *fcc* structure should be increasingly preferred as the temperature is increased towards melting. This difference may help to distinguish between 'impurity-induced' and 'pure' *fcc–hcp* phase behaviour.

It must also be conceded that no spontaneous structural phase transitions have been *observed* in low-temperature, low-pressure experiments for RGS other than helium. However, given that the free-energy difference is so small, the timescale of the dynamics of the structural phase transition is likely to be long. Furthermore, the dynamics (for a given temperature) can be expected to slow down as the atomic mass increases, and as the strength of the Lennard-Jones interaction increases (see table 6.2). This

expectation is borne out by the observation that near melting, helium crystals anneal much more rapidly than the heavier RGS, and that the rate of annealing drops rapidly as the temperature is lowered [123]. This raises the possibility that a *hcp* to *fcc* transition is present, but is kinetically suppressed for the heavier RGS.

With these ideas concerning the role of impurities and recrystallisation timescales in mind, we review the experimental evidence concerning the equilibrium structure of the heavier rare gases (argon, krypton, xenon). This discussion splits into three main sections, looking at the bulk behaviour of solids frozen from the fluid, and bulk-limit behaviour of both rare-gas films and clusters. Following this, the experimental results for neon will be outlined, before presenting a general summary of the experimental results.

Bulk freezing

There is no doubt that *fcc* is indeed the stable phase near the melting point (see, for example, ref. [126]). However, numerous authors have noted the difficulty in forming pure *fcc* samples under these conditions, and that this task gets even more difficult as the temperature is decreased [127, 117, 95, 123]. For example, the authors of ref. [123] note that the crystals must be annealed at $\gtrsim$60% of the triple-point temperature, before being cooled to low temperatures in order to perform the diffraction measurements. This minimum annealing temperature corresponds to a reduced temperature of $\tilde{T} \gtrsim 0.42$, just above our predicted *hcp–fcc* transition. Other authors [127, 123] have found that bulk stacking-faults are at least metastable at low temperature, and no evidence has been uncovered which shows that the stacking-pattern evolves towards *fcc* at temperatures below the *fcc–hcp* transition temperature predicted for the Lennard-Jones solid.

Thin & thick films

While very thin films ($\sim$ a few atoms thick) cannot be expected to reflect the bulk structure, the recovery of the bulk structure in the bulk (thick-film) limit should be observable. Also, as long as the film is formed at low temperature, slowly, and not annealed (at high temperature) prior to observing the structure, we might expect that such systems can avoid the hysteresis associated with taking the bulk solid through the first-order structural transition, i.e. they should be less likely to become trapped in the metastable non-crystalline states that lie between the *fcc* and *hcp* structures.

The initial (1930) determination of the solid-state structure of the rare gases was performed using this kind of experiment [128, 129, 130, 131, 132]. Thick crystalline films were deposited onto cooled rods,

and then studied via X-ray diffraction. These publications conclude that *fcc* is the stable structure. Curiously, the authors of these papers stated in [130] that they had been forced to ignore three lines in the diffraction pattern in order to make their results consistent with an *fcc* structure. These lines, and also the noted absences of the (200) and (400) reflections, are in fact consistent with a mixed or *hcp*-like stacking pattern, and thus it is not clear that the *hcp* structure was given adequate consideration. Later work [133] appears to confirm the stability of the *fcc* structure (with some stacking faults), but no attempt was made to estimate the level of impurities in their samples.

More recent work has been aware of the problem of impurities, and has attempted to ensure the quality of the films. These publications [134, 135, 136, 137, 138, 123, 139, 140] appear to support at least the metastability of *hcp* or *fcc–hcp* mixtures at low temperatures. Often, the *hcp* phase was *assumed* to be metastable, because when the structure was annealed at a temperature *near the melting point* (i.e. above the predicted *fcc–hcp* transition), the *fcc* structure was recovered. This does not constitute proof that *hcp* is not stable for some lower temperature, as this can only be shown by confirming that a *hcp* region will slowly transform into the *fcc* structure at that temperature. Only ref. [139] appears to present any detailed investigation of this issue, and provides some interesting food for thought. Firstly, the authors observed that xenon and krypton films (10^4–10^5 atoms thick) formed *hcp* at low temperatures, and that *fcc* was preferred near melting. More importantly, they also observed the slow annealing of the xenon crystals at fixed temperature, and found that the fraction of *hcp* grew over time for $T < 87$K (i.e. $\tilde{T} < 0.39$), but that the *fcc*-fraction grew at temperatures above 91K ($\tilde{T} > 0.41$). This is consistent with the presence of a kinetically suppressed *fcc–hcp* phase transition at $\tilde{T} \sim 0.4$, not far from the predicted transition-temperature for Lennard-Jonesium ($\tilde{T} \sim 0.36$). Later work has confirmed that a high temperature is required for the formation of an *fcc* structure in Kr and Xe films [141, 142]. However, these publications have neither confirmed nor denied the observation in [139] that *hcp* is stable at low temperatures, and appear to *assume* that *hcp* is only *meta*stable without providing any evidence that this is indeed the case.

Rare-gas clusters

Much of the most recent work on RGS has focused on the behaviour of clusters of rare-gas atoms. These are directly formed from the gaseous phase at low temperatures and so, like the films, should not be susceptible to the hysteresis associated with any structural phase transition. The aim of the cluster experiments has been to track the evolution of the crystal structure with cluster size, in order to understand the bulk structure more clearly. These clusters are usually formed by high-pressure high-temperature jet expansion of a beam of gaseous argon, injected into a vacuum through a narrow nozzle. The temperature (T_0) and pressure (p_0) of the injection process control the temperature of the individual

clusters (T_{cl}) but as there appears to be no simple relationship between T_{cl} and $\{T_0,p_0\}$, the cluster temperature must be estimated from the diffraction data. Of course, the cluster temperature must be lower than the sublimation temperature for that pressure (or the clusters would not form at all), but as the authors of these publications do not expect the precise value of the temperature to play an important role, an accurate estimate of T_{cl} is often not included.

For smaller ($N \lesssim 500$) clusters, a number of distinct structures are observed which have little to do with the bulk crystal. The clusters are nearly spherical, and appear to possess 5-fold icosahedral symmetry [143]. Most of the analysis of these systems has concentrated on the ground-state energy of the clusters (e.g. [144]), but at least one attempt has been made to include thermal effects [145] (although this treatment did not consider *hcp*-like structures). These treatments have worked well for the smaller clusters, but the route by which the cluster structure tends towards that of the bulk is less clear [146, 147].

In the last few years, apparently contradictory experimental results for the structure of *large* rare-gas clusters have been published by two separate groups: Kovalenko *et al.* [148, 149] and Torchet *et al.* [150, 151]. The former group have found a growing preference of *fcc* with increasing size of argon, krypton and xenon clusters. However, the latter group found that argon clusters tended towards an approximately 50:50 mixture of *fcc* and *hcp* as the cluster size is increased, leading to the conclusion that "argon clusters do not adopt the bulk crystal structure [i.e. *fcc*] as their size increases." and that "Consequently, argon cluster research can no longer be expected to contribute to a solution of the RGS problem." Here, we note that the temperature of the argon clusters was estimated to be 32 $\pm$ 2K in [150] and 35 $\pm$ 4K in [151] (in our reduced units, $\tilde{T} = 0.26(2)$ and $\tilde{T} = 0.29(3)$ respectively). This is sufficiently close to the predicted bulk *fcc–hcp* transition that we might expect that for relatively small systems, no clear preference of *fcc* over *hcp* should be found. Unfortunately, no estimate of cluster temperature was supplied in either [148] or [149], and so it has not been possible to confirm whether the cluster temperature in those experiments was above the bulk *fcc–hcp* transition temperature of the Lennard-Jones system.

Neon

While our prediction that *fcc* should always be the stable phase of neon is consistent with the belief that *fcc* is the stable phase of all the RGS but helium, the experimental evidence is still found to be somewhat lacking. All of the publications examined during this research were based on samples formed directly from the gas onto a cold substrate, and thus the experiments should not have been hindered by the presence of a first-order bulk transition. Most authors have observed *fcc* to be the equilibrium structure [121, 128, 129, 152, 153], but stable *hcp* crystallites have been seen in at least one case [153].

None of these investigations attempted to quantify the level of impurities. Later research has attempted to deal with this issue, but only two such publications have been found for neon. The first, by Kogan *et al.* [154], compared the behaviour of ^{20}Ne and ^{22}Ne isotopes of neon. Both were found to form *fcc* at low temperatures ($\sim$4.2K), but this may be due to the high levels of impurities in the samples (1% and 2% for ^{20}Ne and ^{22}Ne respectively). The only other publication that appears to be sufficiently aware of the problems that impurities may pose is that of Kovalenko & Bagrov [137]. They state that the neon used in the experiment was 99.99% pure (at source), and find *fcc* to be the stable phase over a range of temperatures from $\sim$ 5K to $\sim$ 9K. However, they do not attempt to estimate the level of impurites in the sample, and so the level of *contamination* of the neon (due to the process of loading the low-pressure cell) is not known. Despite these concerns, it seems reasonable to say that the equilibrium structure of neon is very probably *fcc*, but more very careful experimental work is required in order to confirm this.

Summary

As the rare-gas solids were not the original focus of this research, this literature review is not as exhaustive as it could be. However, a significant fraction of the rare-gas solid record has been examined, and very few of these publications appear to contradict our results. Those publications that *do* run against our expectations contain no analysis of impurities, which are known to strongly affect the structural behaviour. In short, there is sufficient evidence to cast doubt on the commonly accepted view that *fcc* is the equilibrium structure of all the RGS, at all temperatures and pressures; a preconception which may be holding back the growth in our understanding of rare-gas solids. However, there are many factors that may allow the observed structure of RGS to differ from that predicted for the classical Lennard-Jones solid:

- The role of quantum-mechanical effects, and the degree to which the Lennard-Jones potential is a good representation of the true inter-atomic interaction.
- The concentration of impurities.
- The conditions under which the solid is formed, especially the rate of crystallisation.
- The annealing rate/temperature.
- The possible (meta)stability of stacking-patterns other than *fcc* and *hcp*, especially twinning.

Therefore, only a much more thorough analysis of the experimental record and/or a new series of experiments performed with these possibilities in mind will be able to decide whether the Lennard-Jones model can accurately predict the behaviour of the rare-gas solids.

6.8 Discussion

The lattice-switch technique has been extended to soft potentials and successfully applied to the long-standing problem of the *fcc–hcp* phase boundary of the Lennard-Jones solid. This success also depended upon finding a suitable truncation scheme, which has been shown to be of crucial importance to the accurate prediction of the structural phase behaviour. This strong dependence arises from the long-range nature of the Lennard-Jones potential combined with the similarity of the *fcc* and *hcp* structures. In order that the generalized LSMC technique may be applied efficiently, a number of different TP weight-evolution schemes have been tested. Of these, the algorithms that force the system to explore the order-parameter space *systematically* are found to be most efficient. Also, we find that while any two-phase LSMC simulation *should* be performed in the NPT ensemble (to ensure that the two structures have the appropriate densities), the NVT ensemble can be used when the two densities are sufficiently similar.

A number of different extrapolation techniques have been tested, in terms of their efficiency at determining the *fcc–hcp* phase boundary. Single-histogram extrapolation was found to perform very badly, and Gibbs-Duhem integration was found to be too dependent on being able to measure both the enthalpy *and* the volume difference between the structures accurately. However, by measuring the enthalpy difference and the *free-energy* difference in a single LSMC simulation, the Newton-Raphson technique and the lattice-switch predictor-corrector were found to be very effective tools for predicting the coexistence behaviour.

These techniques have allowed the Lennard-Jones *fcc–hcp* phase boundary to be accurately determined. The effects of finite size and of the truncation of the interaction have been estimated and appear to be negligible. Note, however, that this may not provide a full picture of the structural phase behaviour of the Lennard-Jonesium. There are indications that there may be a region of *bcc* stability near the melting curve at high temperatures and pressures [21], and as no later work has been able to either confirm or deny this possibility, it would be interesting to apply lattice-switch Monte Carlo to this problem as well. As the two structures are so different, this would probably require incorporating a volume dilation into the lattice-switch move, and should be performed in the NPT ensemble. Although it is by no means certain that there will be a region of *bcc*-stability, researching this issue would deepen our understanding of the Lennard-Jones solid as well as broadening the ‘LSMC portfolio’.

Finally, we have shown our results may be of relevance to the long-standing ‘crystal structure problem’ of the rare-gas solids. Out phase diagram confirms that *hcp* is indeed favoured by the ground-state energy contribution, but the lattice-switch results have shown that as the temperature is increased, *fcc* becomes stable before either sublimation or melting occurs. The predicted transition temperature is sufficiently

high that (for the heavier RGS at least) the classical calculations might just be trustworthy. Some of the experimental evidence appears to support this claim, casting some doubt upon the widely held belief that *fcc* is the stable phase everywhere for argon, krypton and xenon. These results suggest that the structural phase behaviour of rare-gas solids should be re-examined, and ideally that more experimental research should be performed in order to finally solve this long-standing problem.

Chapter 7

Conclusions

7.1 LSMC: The story so far...

Prior to this work, the most widely used technique in the simulation of ($T > 0$K) structural phase behaviour was the Einstein crystal integration method. While the concepts behind this technique are relatively straightforward, it can be extremely complicated to apply it to any given problem, as depicted by the flowchart shown in figure 2.8. In the original lattice-switch publication [1], an alternative scheme was presented which provides a simpler and more direct approach to this class of problem (as illustrated by the LSMC flowchart in figure 7.1). In the work presented here, lattice-switch Monte Carlo has been developed into a broadly applicable numerical technique for studying structural phase behaviour.

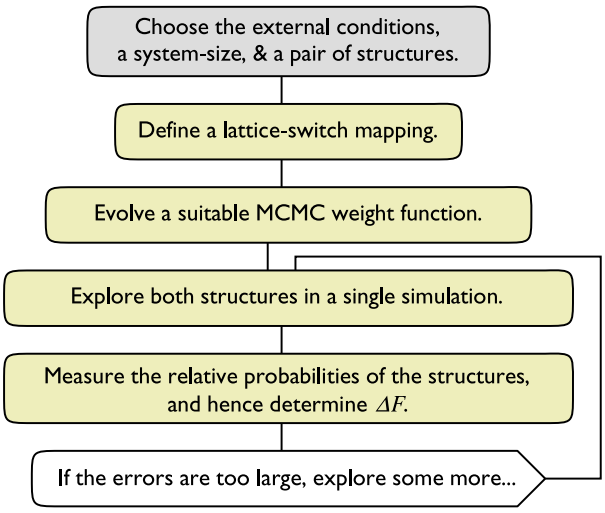


Figure 7.1: Flow diagram of the sequence of steps require when using the lattice-switch Monte Carlo technique to determine the free-energy difference between two structures.

While the issue of how to choose a suitable LS operation for a given problem may not always be obvious (see §7.2 below), and the weight-function estimation requires care, the LSMC approach has one distinct advantage with respect to the way in which an estimate of the errors associated with the free-energy difference may be refined. In the Einstein-crystal integration approach, there are many possible sources of error, and it may not always be clear how best to invest the simulation time required for an improved result. For LSMC however, we can simply run exactly the same simulation for longer. While the two approaches may be approximately equally expensive in terms of simulation time (at least for hard spheres, §4.12), we believe that the directness and simplicity of the lattice-switch method make it an important addition to the toolbox of the computational condensed-matter physicist.

Here, the LSMC technique has been successfully applied to both the hard-sphere and Lennard-Jones solids. As well as providing an arena in which to develop the LSMC method and the MCMC sampling techniques on which it depends, our studies of these two systems have helped to clarify a number of aspects of their structural phase behaviour. For hard-spheres, this work has accurately determined that *fcc* is the equilibrium solid-state structure near melting and in the close-packed limit. This has been confirmed by the results of other authors in the field, allowing us to argue that *fcc* is the stable structure over the entire range of densities for which the solid phase is stable. For the Lennard-Jones system, the combination of a well thought out truncation scheme and LSMC has allowed the *fcc–hcp* free-energy difference of this system to be determined more accurately than any previous attempt. This algorithm was then used as the basis of a new technique (PCLS) for calculating the position of the phase coexistence curve, which was found to be better suited to this particular problem than the more standard techniques available (GDI, SHE). This allowed the full *fcc–hcp* coexistence curve of the Lennard-Jones solid to be accurately determined. These results have led to some interesting questions concerning the structural phase behaviour of the rare-gas solids which, at the very least, would appear to be somewhat more complicated than it is widely understood to be.

7.2 LSMC: Future development

The concept of using a global-coordinate transform to switch between two different phases is much more widely applicable than the work presented here may suggest. During the course of this work, other researchers at Edinburgh have successfully applied these ideas to the solid-*fluid* phase behaviour of the hard-sphere system [155]. This generalised *phase switch* used a cleverly constructed order parameter to facilitate a global-coordinate transform designed to switch between a configuration representative of the fluid-phase and an *fcc* lattice. This in turn allowed a precise determination of the solid-fluid coexistence point.

The formulation supplied for the Lennard-Jones solid may be generalised even further. In §2.4.3, the direct-difference technique of Bennett [25] was introduced as a potentially very powerful means of comparing any two Hamiltonians. Unfortunately, this approach usually fails to produce statistically reliable results because the probability of accepting the Hamiltonian switch is simply too small to allow the free-energy difference to be determined in this way. However, the method supplied in §6.5 uses the energy cost of what is effectively a Hamiltonian switch as the multicanonical order parameter, thus allowing the probability barrier between the two systems to be overcome. This same idea can, in theory, be used to determine the free-energy difference between any two Hamiltonians (see also ref. [156]).

Returning to the lattice-switch, the techniques presented here are immediately applicable to a wide range of systems, including full quantum-mechanical *ab initio* simulation (which is just a more computationally expensive way of evaluating the energy of a given configuration). Some possible examples for future study are the two permutations of the AB_{13} structure of binary hard-sphere systems [157], or perhaps a more precise treatment of the rare-gas solids. For the heavier RGS, it may be sufficient to apply a path-integral approach to the Lennard-Jones solid [158]. However, the most interesting case is probably that of Helium, for which it would be necessary to apply more sophisticated (e.g. *ab initio*) techniques. The structural phase behaviour of ^{3}He and ^{4}He is much more complex than that of Lennard-Jonesium, as both isotopes possess regions of *bcc* stability at very low temperatures as well as the higher temperature *fcc–hcp* transition mentioned in §6.7. There is still some argument over the accurate theoretical prediction of the *bcc* region [159], and very few publications appear to deal with the precise prediction of the *fcc–hcp* phase boundary. This is presumably due to the lack of suitably efficient numerical techniques for dealing with this problem, and so this may prove to be an interesting and useful test-case for any future extension of the lattice-switch technique to what we might call a ‘real system’.

Any further investigation using LSMC must also deal with the question of how best to switch between a given pair of candidate structures. Any arbitrary particle-to-particle displacement mapping *will* work, but as we have seen in §4.10.4, the choice of mapping can strongly affect the efficiency of the simulation. Quite generally, the mapping could be optimised by using something like an MC annealing algorithm to adjust it until some measure of the mapping efficiency has been maximised (such as the inverse of the equilibrium overlap count used in §4.10.4). This procedure is likely to be complicated and probably rather slow, and designing the mapping by hand (using our physical understanding of the relationship between the two structures) is likely to produce better results more quickly. For example, the *fcc* structure can be transformed into the *bcc* structure by stretching the lattice in one direction [21], and it seems likely that this would lead to the most efficient *fcc*-*bcc* mapping.

On this point, it is possible that the efficiency of the *fcc–hcp* mapping could be significantly improved.

Indeed, the smallness of the free-energy difference between *fcc* and *hcp* implies that the regions of configuration space associated with each of the two structures are very nearly the same size. In principle, this means that it should be possible to map the configuration space of one structure directly onto the other, so that nearly all configurations are in fact 'gateway' microstates and that the lattice switch will (almost) always be accepted. This 'ideal mapping' may be distinctly non-trivial, but for soft potentials at least there appears to be at least one possible candidate. The idea is based on taking advantage of the success of the harmonic approximation, and using the normal-modes of the crystal to provide a convenient representation for the switch [160]. This non-local mapping relies on decomposing the particle-displacements into a superposition of the harmonic normal-modes, i.e. using the eigenvectors of the dynamical matrix as the basis set. The normal-modes of the two structures can be matched-up so that the displacements associated with one mode (in one structure) map onto the displacements associated with a different mode (in the other structure). The modes from each of the structures might be matched on the basis of their frequencies (essentially, the energy of the vibrational-mode). With an appropriate scaling of the coordinates, this switch can be *guaranteed* to be accepted to the extent that the harmonic approximation holds, and the harmonic free-energy would be captured through the determinants of the two dynamical matrices (eqn. 4.5). Therefore, at higher temperatures, this mode switch will provide a way of *directly* evaluating the *anharmonic* contribution to the free-energy difference. This development requires care, especially as the harmonic approximation is most easily applicable to the NVT ensemble, whereas LSMC simulations should be performed in the NPT ensemble. Also, the non-local nature of the mapping will significantly slow the simulation (see §6.5), and so for the Lennard-Jones system this approach may be too costly in terms of simulation time. However, this approach could eventually be applicable to *ab initio* simulation, and in this case (where the energy evaluation is so expensive) the normal-mode switch may make all the difference. Work along these lines is currently underway here at Edinburgh.

In short, LSMC represents an original, efficient and potentially very general technique for studying the phase behaviour of many-body systems. The further development of these ideas is already underway, and it seems likely that this technique will go on to be applied to an even wider range of systems, stretching far beyond the research recorded in this thesis.

Bibliography

[1] AD Bruce, NB Wilding & GJ Ackland. 'Free Energy of Crystalline Solids: A Lattice-Switch Monte Carlo Method.' *Physical Review Letters*, **79**(16), pp. 3002–3005 (1997). (e-print @ cond-mat/9706154). iii, 2, 55, 57, 69, 85, 86, 87, 91, 167

[2] AD Bruce, AN Jackson, GJ Ackland & NB Wilding. 'Lattice-switch Monte Carlo method.' *Physical Review E*, **61**(1), pp. 906–919 (January 2000). (e-print @ cond-mat/9910330). vii, 55, 57

[3] AN Jackson & AD Bruce. 'Lattice-switch Monte Carlo for Soft-potentials.' (2001). (to be published). vii, 111

[4] J Maddox. 'Crystals from 1st principles.' *Nature*, **335**(6187), pp. 201–201 (1988). 2

[5] J Keplar. 'Strena.' In *The six-cornered snowflake* (Oxford, 1966) (1611). 6

[6] TC Hales & SP Ferguson. 'The Keplar Conjecture.' (1998). Details of Keplar's conjecture and it's solution can be found at http://www.math.lsa.umich.edu/ hales/countdown/. 6

[7] BJ Alder, WG Hoover & DA Young. 'Studies in Molecular Dynamics. V. High-Density Equation of State and Entropy for Hard-Discs and Spheres.' *Journal of Chemical Physics*, **49**(8), pp. 3688–3696 (October 1968). 6, 56

[8] K Binder & DW Heermann. *Monte Carlo Simulation in Statistical Physics*. Second edition (Springer-Verlag) (1992). 8

[9] NW Ashcroft & ND Mermin. *Solid State Physics* (Saunders College, Philadelphia) (1976). 8, 12, 13, 58, 113, 118, 129, 156

[10] F Reif. *Fundamentals of statistical and thermal physics*. International edition (McGraw-Hill) (1985). 8, 11

[11] M Born & Th von Kármán. *Phys. Z.*, **14**, pp. 15 (1912). 13

[12] M Born & Th von Kármán. *Phys. Z.*, **14**, pp. 65 (1913). 13

[13] RA Cowley. *Journal of Physics C: Solid State*, **6**, pp. 143 (1973). 13

[14] DL Lacks & R C Shukla. 'Molecular dynamics simulations of the effects of truncation of the Taylor expansion of the potential energy on the thermodynamic properties of a crystal.' *Journal of Chemical Physics*, **105**(10), pp. 4185–4190 (September 1996). 13

[15] DA Young & BJ Alder. 'Studies in molecular dynamics. XIII. Singlet and pair distribution functions for hard-disk and hard-sphere solids.' *The Journal of Chemical Physics*, **60**(4), pp. 1254–1267 (1974). 13

[16] J Piasecki & L Peliti. 'Harmonic properties of hard-sphere crystals: a one-dimensional study.' *Journal of Physics A*, **26**, pp. 4819–4825 (1993). 13

[17] D Frenkel & B Smit. *Understanding Molecular Simulation: from algorithms to applications* (Academic Press) (1996). http://molsim.chem.uva.nl/frenkel_smit/. 14, 15, 21, 22, 27, 38, 40, 41, 42, 43, 53, 73

[18] W Kob. 'Computer Simulations of Supercooled Liquids and Glasses.' In ME Cates & RM Evans, editors, *Soft and Fragile Matter*, chapter 10, pp. 259–284 (2000). See also (e-print @ cond-mat/9809268). 15

[19] M Parrinello & A Rahman. 'Polymorphic transitions in single crystals: A new molecular dynamics method.' *Journal of Applied Physics*, **52**(12), pp. 7182–7190 (December 1981). 16

[20] M Parrinello & A Rahman. 'Crystal Structure and Pair Potentials: A Molecular-Dynamics Study.' *Physical Review Letters*, **45**(14), pp. 1196–1199 (October 1980). 16, 114

[21] A Rahman & G Jacucci. 'Relative Stability of f.c.c and b.c.c Structures for Model Systems at High Temperatures.' *Il Nuovo Cimento*, **4D**(4), pp. 357–381 (October 1984). 16, 114, 164, 169

[22] D Frenkel. 'Stability of the High-Pressure Body-Centered Cubic Phase of Helium.' *Physical Review Letters*, **56**(8), pp. 858–860 (Feburary 1986). 16

[23] MA Miller & WP Reinhardt. 'Efficient free energy calculations by variationally optimized metric scaling: Concepts and applications to the volume dependence of cluster free energies and to solid–solid phase transitions.' *Journal of Chemical Physics*, **113**(17), pp. 7035–7046 (2000). 16

[24] GR Smith & AD Bruce. 'A study of the multi-canonical Monte Carlo method.' *Journal of Physics A*, **28**, pp. 6623–6643 (1995). 16, 43, 44, 45, 47, 49

[25] CH Bennet. 'Efficient Estimation of Free Energy Differences from Monte Carlo Data.' *Journal of Computational Physics*, **22**, pp. 245–268 (1976). 17, 70, 136, 169

[26] MC Moody, JR Ray & A Rahman. 'Free energy difference calculations comparing fcc and hcp structures using molecular dynamics computer simulations.' *Journal of Chemical Physics*, **84**(3), pp. 1795–1802 (Feburary 1986). 17

[27] WG Hoover, SG Gray & KW Johnson. 'Thermodynamic Properties of the Fluid and Solid Phases for Inverse Power Potentials.' *Journal of Chemical Physics*, **55**(3), pp. 1128–1136 (August 1971). 20

[28] JG Kirkwood. *Journal of Chemical Physics*, **18**, pp. 380–382 (1950). 20

[29] WG Hoover & FH Ree. 'Melting Transition and Communal Entropy for Hard Spheres.' *Journal of Chemical Physics*, **49**(4), pp. 3609–3617 (October 1968). 20

[30] LV Woodcock. 'Entropy difference between the face-centered cubic and hexagonal close-packed crystal structures.' *Nature (London)*, **384**, pp. 141 (1997). 20, 57, 83, 84

[31] H Ogura, H Matsuda, T Ogawa, N Ogita & A Ueda. 'Computer simulations for the melting curve maximum phenomenon.' *Progress in Theoretical Physics*, **58**, pp. 419–433 (1977). 20

[32] JQ Broughton & Gilmer. 'Molecular dynamics investigation of the crystal-fluid interface. I. Bulk properties.' *Journal of Chemical Physics*, **79**, pp. 5095–5104 (1983). 21

[33] D Frenkel & AJC Ladd. 'New Monte Carlo method to compute the free energy of arbitrary solids. Application to the fcc and hcp phases of hard spheres.' *Journal of Chemical Physics*, **81**, pp. 3188 (1984). 21, 22, 23, 24, 57

[34] MP Allen & DJ Tildesley. *Computer Simulation of Liquids* (Clarendon Press - Oxford) (1987). 27, 40, 69, 140

[35] PK MacKeown & DJ Newman. *Computational Techniques in Physics* (Adam Hilger) (1987). 27, 30

[36] RA Blythe (2000). (private communication). 28

[37] D Mukamel. 'Phase transitions in nonequilibrium systems.' In ME Cates & RM Evans, editors, *Soft and Fragile Matter*, p. 237 (IoP) (2000). 33

[38] JG Kemeny & JL Snell. *Finite Markov Chains* (Nostrand, Princeton) (1960). 33

[39] KL Chung. *Markov Chains with Stationary Transition Probabilities* (Springer, New York) (1967). 33

[40] O Narayan & AP Young. 'Convergence of Monte Carlo Simulation to Equilibrium.' (2000). (e-print @ cond-mat/0008046). 34

[41] VI Manousiouthakis & M W Deem. 'Strict detailed balance is unnecessary in Monte Carlo simulation.' *Journal of Chemical Physics*, **110**(6), pp. 2753–2756 (Feburary 1999). (e-print @ cond-mat/9809240). 34

[42] N Metropolis, A W Rosenbluth, M N Rosenbluth, AH Teller & E Teller. *Journal of Chemical Physics*, **21**, pp. 1087 (1953). 34

[43] R Eppenda & D Frenkel. 'Monte-Carlo study of the isotropic and nematic phases of infinitely thin hard platelets.' *Molecular Physics*, **52**(6), pp. 1303–1334 (1984). 39, 64, 142

[44] FJ Resende & BV Costa. 'Using random number generators in Monte Carlo simulations.' *Physical Review E*, **58**(4), pp. 5183–5184 (1998). 40

[45] Kirkpatrick & Stoll. *Journal of Computational Physics*, **40**, pp. 517 (1981). 40

[46] N Zierler. *Information and Control*, **15**, pp. 67 (1969). 40

[47] JR Taylor. *An introduction to error analysis: the study of uncertainties in physical measurements*. Second edition (University Science Books) (1997). 41

[48] H Cramér. *The Elements of Probability Theory* (Wiley, New York) (1955). 41

[49] K Binder. 'The Monte Carlo Method for the Study of Phase Transitions: A Review of Some Recent Progress.' *Journal of Computational Physics*, **59**, pp. 1 (1985). 42, 49

[50] GM Torrie & JP Valleau. 'Monte Carlo Free Energy Estimates Using Non-Boltzmann Sampling: Application To The Sub-Critical Lennard-Jones Fluid.' *Chemical Physics Letters*, **28**(4), pp. 578–581 (1974). 43

[51] BA Berg & T Neuhaus. 'Multicanonical Ensemble: A New Approach to Simulate First-Order Phase Transitions.' *Physical Review Letters*, **68**(1), pp. 9–12 (1992). 43, 51, 99

[52] GR Smith. *The Measurement of Free Energy by Monte-Carlo Computer Simulation*. Ph.D. thesis, University of Edinburgh (1996). 44

[53] M Fitzgerald, RR Picard & RN Silver. 'Canonical transition probabilities for adaptive Metropolis simulation.' *Europhysics Letters*, **46**(3), pp. 282–287 (May 1999). 48, 49

[54] PMC de Oliveira & TJP Penna. 'Broad Histogram Method.' *Brazilian Journal of Physics*, **26**, pp. 677 (1996). (e-print @ cond-mat/9610041). 51

[55] AR Lima, PMC de Oliveira & TJP Penna. 'A comparison between broad histogram and multicanonical methods.' *Journal of Statistical Physics*, **99**(3-4), pp. 691–705 (2000). (e-print @ cond-mat/0002176). 51

[56] J-S Wang. 'Transition Matrix Monte Carlo Method.' *Computer Physics Communications*, **122**, pp. 22–25 (1999). (e-print @ cond-mat/9810240). 51

[57] J-S Wang, TK Tar & RH Swendsen. 'Transition Matrix Monte Carlo Reweighting and Dynamics.' *Physical Review Letters*, **82**(3), pp. 476–479 (1999). (e-print @ cond-mat/9809181). 51

[58] J-S Wang. 'Flat histogram Monte Carlo method.' *Physica A*, **281**(1-4), pp. 147–150 (2000). (e-print @ cond-mat/9909177). 51

[59] PMC de Oliveira. 'Broad Histogram: An Overview.' *Brazilian Journal of Physics*, **30**(1), pp. 195–211 (2000). (e-print @ cond-mat/0003300). 51

[60] J-S Wang. 'Is the broad histogram random walk dynamics correct?' *European Physical Journal B*, **8**(2), pp. 287–291 (1999). (e-print @ cond-mat/9810017). 51

[61] PMC de Oliveira. 'Broad Histogram Relation Is Exact.' *European Physical Journal B*, **6**(1), pp. 111–115 (1998). (e-print @ cond-mat/9807354). 51

[62] JD Mu noz & H Herrmann. 'Extending the Broad Histogram Method for Continuous Systems.' (1998). (e-print @ cond-mat/9810292). 51

[63] B Grossmann, ML Laursen, T Trappenberg & U-J Wiese. 'A multicanonical algorithm for SU(3) pure gauge theory.' *Physics Letters B*, **293**, pp. 175–180 (1992). 51

[64] NB Wilding. 'Critical-point and coexistence-curve properties of the Lennard-Jones fluid: A finite-size scaling study.' *Physical Review E*, **52**(1), pp. 602–611 (1995). 51

[65] UHE Hansmann & Y Okamoto. 'Finite-size scaling of helix-coil transitions in poly-alanine studied by multicanonical simulations.' *Journal of Chemical Physics*, **110**(2), pp. 1267–1276 (1999). (e-print @ cond-mat/9810358). 51

[66] GR Smith & AD Bruce. 'Multicanonical Monte Carlo study of solid-solid phase coexistance in a model colliod.' *Physical Review E*, **53**(6), pp. 6530–6543 (1996). 51, 52

[67] RJ Speedy. 'Pressure of the metastable hard-sphere fluid.' *Journal of Physics: Condensed Matter*, **9**(41), pp. 8591–8599 (1997). 56

[68] MD Rintoul & S Torquato. 'Metastability and Crystallization in Hard-Sphere Systems.' *Physical Review Letters*, **77**(20), pp. 4198–4201 (1996). 56

[69] NF Carnahan & KE Starling. *Journal of Chemical Physics*, **51**, pp. 635– (1969). 56

[70] RJ Speedy. 'Pressure and entropy of hard-sphere crystals.' *Journal of Physics: Condensed Matter*, **10**, pp. 4387–4391 (1998). 56, 64, 65, 73, 76, 77, 83, 85, 91

[71] BJ Alder, BP Carter & DA Young. 'Crystal Transformation for Hard Spheres.' *Physical Review*, **183**(3), pp. 831–833 (1969). 56, 57

[72] WG Rudd, ZW Salsburg, AP Yu & FH Stillinger. 'Rigid Disks and Spheres at High Densities. III.' *Journal of Chemical Physics*, **49**, pp. 4857–4863 (1968). 57, 65

[73] JL Colot & M Baus. 'The freezing of hard spheres II. A search for structural (f.c.c.–h.c.p.) phase transitions.' *Molecular Physics*, **56**(4), pp. 807–824 (1985). 57

[74] BJ Alder, DA Young & MR Mansigh. 'Hard Sphere Equation of State in the Close-Packed Limit.' *Journal of Computational Physics*, **7**, pp. 361–366 (1971). 57

[75] KW Kratky. 'The stability of fcc and hcp hard-sphere crystals.' *Chemical Physics*, **57**, pp. 167–174 (1981). 57

[76] AD Bruce. (private communication) (1999). 64, 73

[77] Siun-Chuon Mau & David A. Huse. 'Stacking Entropy of Hard-Sphere Crystals.' *Physical Review E*, **59**(4), pp. 4396–4401 (April 1999). 68, 75, 83, 84, 86, 87, 88, 90, 91

[78] LV Woodcock. 'Woodcock replies.' *Nature*, **388**(6639), pp. 236–237 (July 1997). 83, 84

[79] BG Bolhuis, D Frenkel, SC Mau & DA Huse. 'Entropy difference between crystal phases.' *Nature*, **388**(6639), pp. 235–236 (July 1997). 84

[80] KV Tretiakov & KW Wojciechowski. 'Efficient Monte Carlo simulations using a shuffled nested Weyl sequence random number generator.' *Physical Review E*, **60**(6), pp. 7627–7628 (December 1999). 84, 91

[81] Sander Pronk & Daan Frenkel. 'Can stacking faults in hard-sphere crystals anneal out spontaneously?' *Journal of Chemical Physics*, **110**(9), pp. 4589–4592 (March 1999). 83, 84, 85, 86, 88, 90, 91

[82] PN Pusey, W van Megan, P Barlett, BJ Ackerson, JG Rarity & SM Underwood. 'Structure of Crystals of Hard Colloidal Spheres.' *Physical Review Letters*, **63**(25), pp. 2753–2756 (1989). 88, 89, 90

[83] Y Choi, T Ree & FH Ree. 'Phase diagram of a Lennard-Jones solid.' *Journal of Chemical Physics*, **99**, pp. 9917–9919 (1993). 89, 114, 115, 120, 129

[84] J Zhu, M Li, R Rogers, W Meyer, RH Ottewill, WB Russel & PM Chaikin. *Nature*, **387**, pp. 883 (1997). 90

[85] V Martelozzo. Ph.D. thesis, University of Edinburgh (2001). 90

[86] AZ Panagiotopoulos. 'Monte Carlo methods for phase equilibria of fluids.' *Journal of Physics: Condensed Matter*, **12**, pp. R25–R52 (2000). 94

[87] AM Ferrenberg & RH Swendsen. 'Mew Monte Carlo Technique for Studying Phase Transitions.' *Physical Review Letters*, **61**(23), pp. 2635–2638 (1988). 95

[88] WH Press, SA Teukolsky, WT Vetterling & BP Flannery. *Numerical Recipies in FORTRAN: The Art of Scientific Computing*. Second edition (Cambridge University Press) (1992). 99, 100, 105, 106, 126

[89] DA Kofke. 'Direct evaluation of phase coexistence by molecular simulation via integration along the saturation line.' *Journal of Chemical Physics*, **98**(5), pp. 4149–4161 (March 1993). 104, 106

[90] DA Kofke. 'Semigrand canonical Monte Carlo simulation; Integration along coexistence lines.' *Advances in Chemical Physics*, **105**, pp. 405–441 (1999). 104, 107, 108

[91] FA Escobedo. 'Tracing coexistence lines in multicomponent fluid mixtures by molecular simulation.' *Journal of Chemical Physics*, **110**(24), pp. 11999–12010 (1999). 104

[92] RP Feynman. *Six Easy Pieces* (Addison-Wesley) (1965). 111

[93] R Agrawal & DA Kofke. 'Thermodynamic and structural properties of model systems at fluid-solid coexistence II. Melting and sublimation of the Lennard-Jones system.' *Molecular Physics*, **85**(1), pp. 43–59 (1995). 111, 112, 154

[94] MA van der Hoef. 'Free energy of the Lennard-Jones solid.' *Journal of Chemical Physics*, **113**(18), pp. 8142–8148 (2000). 111, 112

[95] KF Niebel & JA Venables. 'The Crystal Structure Problem.' In ML Klein & JA Venables, editors, *Rare Gas Solids*, volume 1, chapter 9, pp. 558–589 (Academic Press) (1976). 113, 114, 127, 159, 160

[96] JE Lennard-Jones & AE Ingham. *Proceedings of the Royal Society*, **A107**, pp. 636 (1925). 113, 118, 119

[97] T Kihara & S Koba. 'Crystal Structures and Intermolecular Forces of Rare Gases.' *Journal of the Physical Society of Japan*, **7**(4), pp. 348–354 (1952). 113, 119

[98] THK Barron & C Domb. 'On the cubic and hexagonal close-packed lattices.' *Proceedings of the Royal Society*, **A227**, pp. 447–465 (1955). 114, 118, 153

[99] ZW Salsburg & DA Huckaby. 'Free Energy of the fcc and hcp Lattices under First- and Second-Neighbour Harmonic Interactions.' *Journal of Computational Physics*, **7**, pp. 489–502 (1971). 114, 129, 157

[100] DK Neb, RK Singh & SP Sanyal. 'On the Stability of Crystal Structures of Rare Gas Solids.' *Physica Status Solidi (B)*, **134**, pp. K97–K100 (1986). 114

[101] VF Lotrich & K Szalewicz. 'Three-Body Contribution to Binding Energy of Solid Argon and Analysis of Crystal Structure.' *Physical Review Letters*, **79**(7), pp. 1301–1304 (1997). 114

[102] BJ Alder & PH Paulson. *Journal of Chemical Physics*, **43**, pp. 4172 (1965). 114

[103] B Borden & C Radin. 'The crystal structure of the nobel gases.' *Journal of Chemical Physics*, **75**(4), pp. 2012–2013 (1981). 114

[104] AE Galashev. 'Thermodynamic stability of hexagonal close packed (hcp) crystals in the molecular-dynamic model.' *Journal of Structural Chemistry*, **25**(5), pp. 728–733 (1985). 114, 153

[105] G Jackson & F van Swol. 'Peturbation theory of a model hcp solid.' *Molecular Physics*, **65**(1), pp. 161–173 (1988). 114, 115, 117, 120, 121

[106] Y Choi, T Ree & FH Ree. 'Hard-sphere radial distribution functions for face-centered cubic and hexagonal close-packed phases: Representation and use in a solid state peturbation theory.' *Journal of Chemical Physics*, **95**(10), pp. 7548–7561 (November 1991). 114, 115, 117, 120, 121, 122

[107] Y Choi, T Ree & FH Ree. 'Crystal stability of heavy-rare-gas solids on the melting line.' *Physical Review B*, **48**(5), pp. 2988–2991 (August 1993). 115

[108] S Somasi, B Khomami & R Lovett. 'Simulation of the third law free-energies of face-centered-cubic and hexagonal-close-packed Lennard-Jones solids.' *Journal of Chemical Physics*, **113**(10), pp. 4320–4330 (2000). 115, 121

[109] RGD Valle & E Venuti. 'Quasi Harmonic Lattice Dynamics and Molecular Dynamics calculations for the Lennard-Jones solids.' *Physical Review B*, **58**(1), pp. 206–212 (1998). (e-print @ cond-mat/9805048). 119, 157

[110] C Kittel. *Introduction to solid state physics*. 7th edition (Wiley) (1996). 119

[111] AA Maradudin. 'Elements of Lattice Dynamics.' In GK Horton & AA Maradudin, editors, *Dynamical Properties of Solids*, volume 1, chapter 1, pp. 18–23 (North-Holland Publishing Company) (1974). 123, 124, 129

[112] *The Numerical Algorithms Group (NAG) Library, C version*. www.nag.com. 126

[113] ER Dobbs & GO Jones. 'Theory and properties of solid Argon.' *Reports on Progress in Physics*, **20**, pp. 516–564 (1957). 126, 127

[114] RJ Hardy. 'Temperature and Pressure Dependence of Intrinsic Anharmonic and Quantum Corrections to the Equation of State.' *Journal of Geophysical Research*, **85**(B12), pp. 7011–7015 (1980). 131

[115] WG Hoover. 'Entropy for Small Classical Crystals.' *Journal of Chemical Physics*, **49**, pp. 1981–1982 (1968). 148

[116] JM Polson, E Trizac, S Pronk & D Frenkel. 'Finite-size corrections to the free energies of crystalline solids.' *Journal of Chemical Physics*, **112**(12), pp. 5339–5342 (1999). (e-print @ cond-mat/9909162). 148

[117] GK Horton. 'Rare Gas Solids-A Century of Excitement and Progress.' In Klein & Venables [122], chapter 1. 156, 160

[118] 'Chemical Elements.com.' http://www.chemicalelements.com/. 157

[119] DA Young. *Phase Diagrams of the Elements* (University of California, Berkeley) (1991). 157, 158, 159

[120] AK McMahan. 'Structural transitions and metallization in compressed solid argon.' *Physical Review B*, **33**(8), pp. 5344–5349 (1986). 158

[121] RJ Hemley, CS Zha, AP Jephcoat, HK Mao, LW Finger & DE Cox. 'X-ray diffraction and equation of state of solid neon to 110GPa.' *Physical Review B*, **39**(16), pp. 11820–11827 (1989). 159, 162

[122] ML Klein & JA Venables, editors. *Rare Gas Solids*, volume 1 & 2 (Academic Press) (1976). 159, 177

[123] JA Venables & BL Smith. 'Crystal growth and crystal defects.' In Klein & Venables [122], chapter 10. 159, 160, 161

[124] T Bricheno & JA Venables. 'The free energy difference between fcc and hcp forms of solid argon from Ar-O_2 vapour pressure measurements.' *Journal of Physics C: Solid State*, **9**, pp. 4095–4108 (1975). 159

[125] L Meyer. 'Phase transitions in van der Waal's lattices.' *Advances in Chemical Physics*, **16**, pp. 343–387 (1969). 159

[126] K Olszewski. 'The Liquefaction and Solidification of Argon.' *Philosophical Transactions of the Royal Society of London*, **A186**, pp. 253–259 (1895). 160

[127] CS Barrett & L Meyer. 'X-Ray Diffraction Study of Solid Argon.' *The Journal of Chemical Physics*, **41**(4), pp. 1078–1081 (1964). 160

[128] J de Smedt, WH Keesom & HH Mooy. 'On the crystal structure of neon.' *Communications from the physical laboratory of the University of Leiden*, **18**(203E), pp. 39–44 (1930). 160, 162

[129] J de Smedt, WH Keesom & HH Mooy. 'On the crystal structure of neon.' *Proceedings of the Academy of Sciences of Amsterdam*, **33**, pp. 255–257 (1930). 160, 162

[130] WH Keesom & HH Mooy. 'On the Crystal Structure of Krypton.' *Proceedings of the Academy of Sciences of Amsterdam*, **33**, pp. 447–449 (1930). 160, 161

[131] WH Keesom & HH Mooy. *Nature*, **126**(3172), pp. 243 (1930). 160

[132] WH Keesom & HH Mooy. *Nature*, **125**(3163), pp. 889 (1930). 160

[133] LH Bolz, HP Brioda & HS Peiser. 'Some observations on growing crystals of argon.' *Acta Crystallographica*, **15**, pp. 810–811 (1962). 161

[134] R Bullough, HR Glyde & JA Venables. 'Stacking-fault energy and many-body force effects in solid argon.' *Physical Review Letters*, **17**(5), pp. 249–250 (1966). 161

[135] I Lefkowitz, K Kramer, MA Shields & GL Pollack. 'Dielectric and Optical Properties of Crystalline Argon.' *Journal of Applied Physics*, **38**(12), pp. 4867–4873 (1967). 161

[136] SI Kolalenko & NN Bagrov. 'Electron-diffraction investigations of the structure of thin solid argon films.' *Soviet Physics - Solid State*, **9**(10), pp. 2396–2397 (1968). 161

[137] SI Kolalenko & NN Bagrov. 'Structure of argon and neon polycrystalline thin films.' *Soviet Physics - Solid State*, **11**(10), pp. 2207–2212 (1970). 161, 163

[138] JA Venables & CA English. 'Thin films of condensed gases.' *Thin Solid Films*, **7**, pp. 369–389 (1971). 161

[139] Y Sonnenblick, ZH Kalman & IT Steinberger. 'Growth and crystal structures of solid xenon and krypton.' *Journal of Crystal Growth*, **58**, pp. 143–151 (1982). 161

[140] JZ Larese, QM Zhang, L Passel & JM Hastings. 'Layer-by-layer growth of solid argon films on graphite as studied by neutron diffraction.' *Physical Review B*, **40**(7), pp. 4271–4275 (1989). 161

[141] RJ Keyse & JA Venables. 'Crystal growth and microstructure of solid Krypton and Xenon.' *Journal of Crystal Growth*, **71**, pp. 525–537 (1985). 161

[142] M Maruyama. 'Growth and roughening transition of rare gas crystals.' *Journal of Crystal Growth*, **89**, pp. 415–522 (1988). 161

[143] BW van de Waal. 'Icosahedral, decahedral, fcc, and defect-fcc structural models for Ar_N clusters, $N \gtrsim 500$: How plausible are they?' *Journal of Chemical Physics*, **98**(6), pp. 4909–4919 (1992). 162

[144] RH Leary & JPK Doye. 'Tetrahedral Global Minimum for the 98-atom Lennard-Jones Cluster.' *Physical Review E*, **60**(6), pp. R6320–R6322 (1999). (e-print @ cond-mat/9908165). 162

[145] JPK Doye & F Calvo. 'Entropic effects on the Size Evolution of Cluster Structure.' *Physical Review Letters*, **86**(16), pp. 3570–3573 (2001). (e-print @ cond-mat/0011466). 162

[146] BW van de Waal. 'Size-Dependent Icosahedral → fcc Structural Transition in Rare-Gas Clusters.' *Physical Review Letters*, **76**(7), pp. 1083–1086 (1996). 162

[147] BW van de Waal. 'Structural analysis of large argon clusters from gas-phase electron diffraction data: some recent results.' *Journal of Molecular Structure*, **485–486**, pp. 269–279 (1999). 162

[148] SI Kovalenko, DD Solnyshkin, ET Verkhovtseva & VV Eremenko. 'Experimental detection of stacking faults in rare gas clusters.' *Chemical Physics Letters*, **250**, pp. 309–312 (1996). 162

[149] SI Kovalenko, DD Solnyshkin, EA Bondarenko, ET Verkhovtseva & VV Eremenko. 'Electron diffraction study of the rise of crystal phase in rare-gas and nitrogen clusters.' *Journal of Crystal Growth*, **191**, pp. 553–557 (1998). 162

[150] M-F de Feraudy & G Torchet. 'Deformation and twin faults in large argon clusters.' *Journal of Crystal Growth*, **217**, pp. 449–457 (2000). 162

[151] BW van de Waal, G Torchet & M-F de Feraudy. 'Structure of rare clusters Ar_N, $10^3 < n < 10^5$: experiments and simulations.' *Chemical Physics Letters*, **331**, pp. 57–63 (2000). 162

[152] DG Henshaw. 'Atomic Distribution in Liquid and Solid Neon and Solid Argon by Neutron Diffraction.' *Physical Review*, **111**(6), pp. 1470–1475 (1958). 162

[153] O Bostanjoglo & R Kleinschmidt. 'Neue Kristallstruktur von festem Neon.' *Zeitschrift für Naturforschung*, **21A**, pp. 2106 (1966). 162

[154] VS Kogan, BG Lazarev & RF Bulatova. 'Different lattice constants of solid neon isotopes.' *Soviet Physics JETP*, **13**(1), pp. 19–20 (1961). 163

[155] NB Wilding & AD Bruce. 'Freezing by Monte Carlo Phase-Switch.' *Physical Review Letters*, **85**(24), pp. 5138–5141 (2000). (e-print @ cond-mat/0009062). 168

[156] C Jarzynski. 'Targeted free energy peturbation.' *(private communication, to be published)* (2001). 169

[157] MD Eldridge, PA Madden & D Frenkel. 'The stability of the AB13 crystal in a binary hard-sphere system.' *Molecular Physics*, **79**(1), pp. 105–120 (1993). 169

[158] MH Müser, P Nielaba & K Binder. 'Path-integral Monte Carlo study of crystalline Lennard-Jones systems.' *Physical Review B*, **51**(5), pp. 2723–2731 (1995). 169

[159] D Frenkel. 'Stability of the High-Pressure Body-Centered-Cubic Phase of Helium.' *Physical Review Letters*, **56**(8), pp. 858–861 (1986). 169

[160] GJ Ackland. (private communication). 170

Appendix A

Free Energy Derivatives and Identities

Here a few free-energy identities are derived and presented in a form useful to the main body of the text.

A.1 The Helmholtz Free Energy

From the definition of the Helmholtz free-energy ($\tilde{f} = \frac{F}{NkT}$) for a given structure α (eqn. 2.12), it is straightforward to determine its derivatives with respect to β,

$$\begin{aligned}
\tilde{f}_\alpha &= -\frac{1}{N}\ln \mathcal{Z}_\alpha && \text{(A.1)}\\
\frac{\delta \tilde{f}_\alpha}{\delta\beta} &= -\frac{1}{N}\frac{\delta}{\delta\beta}\ln \mathcal{Z} && \text{(A.2)}\\
&= -\frac{1}{N}\,\frac{1}{\mathcal{Z}}\frac{\delta}{\delta\beta}\prod_i\left[\int_\alpha \mathrm{d}\vec{r}_i\right]\exp\left[-\beta E\right]\ ,\\
&= -\frac{1}{N}\,\frac{1}{\mathcal{Z}}\prod_i\left[\int_\alpha \mathrm{d}\vec{r}_i\right]\frac{\delta}{\delta\beta}\ \exp\left[-\beta E\right]\ ,\\
&= -\frac{1}{N}\,\frac{1}{\mathcal{Z}}\prod_i\left[\int_\alpha \mathrm{d}\vec{r}_i\right] - E\ \exp\left[-\beta E\right]\ ,\\
&= \frac{\overline{E_\alpha}}{N} = \overline{e_\alpha}\ , && \text{(A.3)}
\end{aligned}$$

where $\overline{e_\alpha}$ is the mean energy per particle for the structure α.

A.2 The Gibbs Free Energy

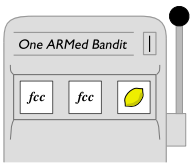
For the Gibbs free-energy ($\tilde{g} = \frac{G}{NkT}$, eqn. 2.16), the derivative with respect to β is derived is much the same way as for A.3,

$$\begin{aligned}
\frac{\delta \tilde{g}_\alpha}{\delta \beta} &= -\frac{1}{N}\frac{\delta}{\delta\beta}\ln \mathcal{G} && \text{(A.4)}\\
&= -\frac{1}{N}\frac{1}{\mathcal{G}}\frac{\delta}{\delta\beta}\prod_i \left[\int_\alpha \mathrm{d}\vec{r}_i\right]\int \mathrm{d}V\ \exp\left[-\beta E - \beta \mathrm{p}V\right]\ ,\\
&= -\frac{1}{N}\frac{1}{\mathcal{G}}\prod_i \left[\int_\alpha \mathrm{d}\vec{r}_i\right]\int \mathrm{d}V\ \frac{\delta}{\delta\beta}\ \exp\left[-\beta E - \beta \mathrm{p}V\right]\ ,\\
&= -\frac{1}{N}\frac{1}{\mathcal{G}}\prod_i \left[\int_\alpha \mathrm{d}\vec{r}_i\right]\int \mathrm{d}V\ -(E+\mathrm{p}V)\ \exp\left[-\beta E - \beta \mathrm{p}V\right]\ ,\\
&= \frac{\overline{E_\alpha - \mathrm{p}V_\alpha}}{N} = \overline{h_\alpha}\ , && \text{(A.5)}
\end{aligned}$$

where $\overline{e_\alpha}$ is the mean energy per particle for the structure α. Similarly, differential of g with respect to the pressure is found to be

$$\begin{aligned}
\frac{\delta \tilde{g}_\alpha}{\delta \mathrm{p}} &= -\frac{1}{N}\frac{\delta}{\delta\mathrm{p}}\ln \mathcal{G} && \text{(A.6)}\\
&= -\frac{1}{N}\frac{1}{\mathcal{G}}\prod_i \left[\int_\alpha \mathrm{d}\vec{r}_i\right]\int \mathrm{d}V\ -\beta V\ \exp\left[-\beta E - \beta \mathrm{p}V\right]\ ,\\
&= \frac{\beta\overline{V_\alpha}}{N} = \beta\overline{v_\alpha}\ , && \text{(A.7)}
\end{aligned}$$

where $\overline{v_\alpha}$ is the mean volume per particle for a particular phase.


One ARMed Bandit
fcc
fcc