Nucleation in Gold Nanoclusters

A Thesis Submitted to the College of

Graduate Studies and Research

in Partial Fulfillment of the Requirements

For the Degree of Masters of Science

in the Department of Chemistry

University of Saskatchewan

Saskatoon, Saskatchewan

by

Eduardo Mendez-Villuendas

Permission To Use

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Chemistry
University of Saskatchewan
110 Science Place
Saskatoon, SK S7N-5C9 Canada

Abstract

The goal of this work is to provide a detailed description of the freezing mechanism in gold clusters. This is accomplished by using constrained Monte Carlo simulations combined with parallel tempering algorithms to evaluate the Free Energy barriers for various temperatures with respect to crystalline order parameters on a 456 atom cluster.

Our simulation results help us to challenge the usual assumption of classic nucleation theory where nucleation starts at the center of a cluster, showing instead that nucleation is favored by freezing started at the surface. We study simplistic phenomenological models for surface freezing and find that the three phase contact line free energy term must be included in order to properly describe the features of the free energy barriers.

Furthermore, we propose an alternative free energy parameter with which we are able to identify a kinetic spinodal temperature where the nucleation barrier disappears and find that the critical cluster size remains finite at the limit of stability of the fluid phase. This result is supported by Molecular Dynamics simulations.

Contents

List of Tables

List of Figures

List of Abbreviations

CNT	Classical Nucleation Theory.
MC	Monte Carlo.
MD	Molecular Dynamics.
EMA	Effective Mass Approximation.
XMC	Monte Carlo code developed for the present thesis work.
MFT	Mean Field Theory.
SC	Simple Cubic lattice.
BCC	Body Centered Cubic lattice.
FCC	Face Centered Cubic lattice.
HCP	Hexagonal Compact lattice.
ICO	Icosahedral.
Dh	Decahedral.
TO	Truncated Octahedral.
$1\ ns$	$1 \times 10^{-9} s$
$1\ ps$	$1 \times 10^{-12} s$
$1\ fs$	$1 \times 10^{-15} s$

List of Symbols

n	Embryo size.
n_{max}	Largest embryo size on a cluster configuration.
n^*	Critical embryo size.
n^*_{max}	Critical largest embryo size.
n_0	Umbrella center.
κ	Umbrella constant.
ϕ	Umbrella bias potential.
ΔG	Change in Gibbs free energy.
ΔF	Change in Helmoltz free energy.
$\Delta F(n^*)$	Free energy barrier to form the critical embryo from the liquid.
$\Delta F(n^*_{max})$	Free energy at the critical embryo respect to the n_{max} parameter.
r_c	Cone algorithm probe raddi ($5.7\mathring{A}$ for gold.).
r_b	Neighboring distance. ($3.85\mathring{A}$ for gold.)
C_{min}	Minimum dot product threshold (0.65 for Gold).
CxN_T	Threshold number of connections per neighbor (0.5 for Gold).
q_i	Six dimensional q_6 vector for particle i.
c_{ij}	Dot product between two neighboring particles i and j.
$A_{\mu,\nu}$	Interface between phases μ and ν (units of area)
$\sigma_{\mu,\nu}$	Free energy contribution for building the interface between phases μ and ν (taken in units of energy per area)
μ_k	Chemical potential of phase k.
$\Delta\mu$	Change in chemical potentials between solid and liquid phases.
J	Rate of nucleation.
T_m	Melting point temperature.
k_B	Boltzmann constant.

Chapter I.

Introduction

Nanometer-sized clusters contain from a few tens to several thousand atoms and exhibit phase transitions and a variety of structures not present in their bulk counterparts that result from quantum mechanical confinement effects, fluctuation of thermodynamic quantities, and large surface to volume ratios. For instance, the difference in chemical environments of the atoms in the surface compared to those in the interior leads to surface reconstruction, a reordering of the surface atoms, which impacts the overall optical properties of these systems [1, 2, 3, 4].

Theoretical interest in the properties of nano-clusters is motivated by fundamental questions regarding the role of system size on the properties of matter and by the underlying desire to harness their special characteristics in the rational design of novel materials and devices [5]. Furthermore, aerosol particles composed of SiO_2, Al_2O_3, NaCl and $(NH_4)_2SO_4$, in the 8nm to 100nm diameter size range, are ubiquitous in the lower troposphere and play a fundamental role in the microphysics of clouds by providing heterogeneous nucleation sites for the formation of liquid droplets or ice crystals, as well as serving as reactants and catalysts in important atmospheric chemical cycles [6]. Recent atmospheric studies on nucleation [7, 8, 9] suggest that the mechanism of freezing in clusters has the effect of changing nucleation rates by orders of magnitude. Understanding how nucleation occurs at the molecular level is a topic of fundamental relevance by its own right. Finally, the crystallization of solids in the bulk phase, starts with the formation of small embryos in the nanometer size range, and hence, their structure determines to great extent the structure of the final solid [10].

In this thesis, we address two important questions:

1. What is the structural mechanism involved in the freezing of nanoparticles? To this end we have used computer simulation techniques to directly calculate the free energy barrier to freezing and examine the properties of the embryos of the solid phase as they form within the fluid nanoparticle. This work represents the first calculations of nucleation free energy barriers for a nanoparticle system.

2. Is there a limit of stability to the fluid phase in a nanometer sized cluster? This is a fundamental question regarding the nature of the fluid phase in general. While it is known that there is a liquid-gas spinodal in single component systems and that fluid mixtures exhibit spinodals with respect to phase separation, the notion that the fluid phase has a limit of stability with respect to a solid phase has not been explored. This work will introduce a new free energy work function to investigate the stability of the fluid in a deeply supercooled nanoparticle.

Gold has traditionally been selected as a preferred system of study mainly because of its stability in colloidal sols and the relative ease with which it can be used to test light scattering theories [11]. Furthermore, the rich variety of methods that have been devised to create gold nanoparticles in the laboratory [12, 13], including nucleation in an inert gas-atmosphere[14] and the deposition of metallic vapor on many substrates [15] makes gold one of the most versatile metals available to the experimentalist.

Properties of gold clusters have been studied extensively by a host of other authors [16, 17], including phenomenology such as solid to solid transformations [18], optics [19], and quantum effects on spectra [20]. *Bartlett et al.* [21] report melting and freezing phenomena of gold nanoclusters of different sizes and test different phenomenological models predicting size dependent melting points, and also calculate dynamic rates of nucleation [22]. In regards to free energies to nucleation, *Nam* *et al.* [23] studied the free energy of a gold cluster with respect to a global order parameter. So far, detailed models have addressed the problem of testing the limits of validity of Classic Nucleation Theory (CNT) in hard sphere colloids [24, 25], and silica [26] to name two examples. Efforts have also been directed at studying nucleation in surfaces [27], however the problem of testing the limits of CNT theory on metallic clusters has not been addressed.

The content on this thesis work is organized as follows:

In the first chapter we discuss the fundamentals of phase behavior and classic nucleation theories, and then discuss the problem of defining the stability of the liquid phase. This is followed by a review of the most relevant computational methods used to optimize the sampling of states that lead to an efficient calculation of free energy barriers for the crystallization of gold clusters. Finally, we provide an overview of the previous and current research developments in the area of cluster science and in particular the simulation of gold nanoclusters.

The second chapter is devoted to the description of our calculations and models for the nucleation of gold clusters. We start by describing the effective mass approximation (EMA) molecular potential, followed by an assessment of the *n-sized* embryo order parameter, which is tested extensively as to produce an appropriate measure of nucleation. We then proceed to show the calculations involved in the prediction of free energy barriers and compare the result with bulk and surface models of nucleation.

In the third chapter, we introduce a new work function to describe the limit of stability of the liquid clusters. We also use a dynamic calculation of the rate to support our suggestion that there is a spinodal for the liquid phase of the cluster. This thesis is complemented with appendices providing a description of the problem of thermodynamic stability, followed by a detailed justification about the order parameters employed, and at last, a description of the computer algorithms that we implemented.

1.1 Nucleation Theory

The relaxation of a metastable system towards stable equilibrium involves the formation of a new phase and is generally characterized by a nucleation rate which measures the number of growing embryos formed per unit time per unit volume. Localized fuctuations in the metastable phase lead to the formation of small embryos, but because the process is activated, i.e. there is a free energy barrier, and only those embryos greater than the critical size grow spontaneously into the new stable phase while smaller fluctuations dissolve back into the metastable state.

In the usual case scenario, the presence of impurities in the nucleating phase, and the presence of solid interfaces, provides preferential sites for heterogeneous nucleation. In the absence of such nucleating agents, however, homogeneous nucleation becomes the fundamental process for phase transitions [28].

While metastable systems relax via nucleation, unstable systems relax to the new phase by spinodal decomposition. Once the system has reached its limit of metastability and becomes unstable the free energy barrier has disappeared and any fluctuation will cause spontaneous phase separation. Consequently, unstable phase transitions are characterized by the spontaneous growth of long-wavelength fluctuations while metastable phase transitions are characterized by the activated growth of localized embryo growth.

The focus of theoretical studies of nucleation is to understand the molecular mechanisms involved in phase transformations and to relate these to experimentally measured nucleation rates. The different theories used to describe nucleation fall into two categories. Kinetic theories whose aim is to directly evaluate the coefficients governing the growth of embryos in the new phase and phenomenological nucleation theories where nucleation rates are derived from the thermodynamics of embryo formation.

1.1.1 Classical Nucleation Theory

The subject matter of any given nucleation theory is to describe the evolution of the population of embryos in the new forming phase. For example, in a supercooled fluid, the continuous creation and destruction of small embryos results from density fluctuations. These embryos are assumed to shrink and grow in size by gaining or losing single molecules -or atoms-, hence a description of the population change of embryos of size n at time t may be written in the form given by the master equation

$$\frac{\partial f(n,t)}{\partial t} = k^+_{n-1} f(n-1,t) + k^-_{n+1} f(n+1,t) - k^+_n f(n,t) - k^-_n f(n,t), \tag{1.1}$$

where $f(n,t)$ is the number density of embryos containing n monomers at time t, while k^+_n and k^-_n are the rates at which the n-size embryo gains and looses monomers respectively. This equation may be cast in the form:

$$\frac{\partial f(n,t)}{\partial t} = J(n-1,t) - J(n,t), \tag{1.2}$$

where

$$J(n) = k^+_n f(n,t) - k^-_{n+1} f(n+1,t), \tag{1.3}$$

is the net rate at which embryos of size n become embryos of size $n+1$ at time t. In the case of gas phase nucleation, the value of the kinetic constant k^+_n for the attachment of monomers is usually taken from kinetic theory of gasses, while the value of the detachment rate constant k^-_n is more difficult to obtain in an independent way. To avoid this difficulty, classical nucleation theory (CNT) makes use of what is known as the *constrained equilibrium hypothesis* to obtain a relationship between the coefficients k^+ and k^-. By assuming detailed balance at equilibrium, the flux $J(n)$ must vanish and Eq (1.3) becomes

$$J(n) = k^+_n f_{eq}(n) - k^-_{n+1} f_{eq}(n+1) = 0, \tag{1.4}$$

where the equilibrium distribution of n-sized embryos, denoted $f_{eq}(n)$, has replaced the nonequilibrium distribution. Rearranging eqn. (1.4) gives

$$k^-_{n+1} = \frac{f_{eq}(n)}{f_{eq}(n+1)} k^+_n, \tag{1.5}$$

which, when substituted back into eqn. (1.3) yields

$$J(n) = k^+_n f_{eq}(n) \left[\frac{f(n,t)}{f_{eq}(n)} - \frac{f(n+1,t)}{f_{eq}(n+1)} \right]. \tag{1.6}$$

CNT assumes steady state conditions where the populations of different cluster sizes is no longer dependent on time. Consequently, $\partial f(n,t)/\partial t = 0$ so from eqn. (1.2), the flux is neither dependent on the size of the cluster, i.e. $J(n) = J$. The total steady state nucleation rate is then obtained by doing a recurrent sum over eqn. (1.6),

$$J = N_{tot} \left[\sum_{n_{min}}^{n_{max}} \frac{1}{k^+_n f_{eq}(n)} \right]^{-1}, \tag{1.7}$$

where N_{tot} is the total number density of embryos, and the limits of the sum are taken from the smallest n_{min} to the largest embryo in the cluster n_{max}. The properties of the sum are such that for $n \geq n_{min}$, $f(n) = f_{eq}(n)$ and for $n > n_{max}$, $f(n) = 0$. The nucleation rate has been shown to be insensitive to the boundaries of the sum [29]. The equilibrium embryo size distribution $f_{eq}(n)$ is obtained directly from the theory of thermodynamic fluctuations [30]

$$f_{eq}(n) = f_{eq}(0) \exp\left(-\frac{\Delta G(n)}{k_B T} \right), \tag{1.8}$$

where $G(n)$ is formally the work or free energy required to form an n-sized embryo from monomers, k_B is the Boltzmann constant and T is the temperature.

The summation (1.7) is replaced by an integral [1] and the expression (1.8) sub-

---

[1] This is a good approximation for $1/N$ small, where N is the number of atoms in the system and it is strictly valid in the thermodynamic limit

stituted to yield

$$J \approx N_{tot}\left[\int_{n=n_{min}}^{n_{max}} \frac{1}{k_n^+ f_{eq}(1)} \exp\left(\frac{\Delta G(n)}{k_B T}\right) dn\right]^{-1}. \tag{1.9}$$

When the free energy barrier is high, this expression may be approximated by the steepest descent approximation in which the overwhelming contribution to the integral comes from values centered around the location of the critical embryo size n^*. We can then assume that $k^+(n) \cong k^+(n)$ and replace the limits of integration from 0 to ∞. Further, we approximate the free energy around the critical embryo n^* to obtain.

$$\Delta G(n) \approx \Delta G(n^*) + \ \tfrac{1}{2}\tfrac{d^2\Delta G(n)}{dn^2}\Big|_{n^*}(n-n^*)^2. \tag{1.10}$$

The nucleation rate can now be written as

$$J \approx k_{n^*}^+ f_{eq}(1) \exp\left(-\frac{\Delta G(n^*)}{k_B T}\right)\left[\int_0^\infty \exp\,(\frac{-\frac{1}{2}\frac{d^2\Delta G(n)}{dn^2}\Big|_{n^*}(n-n^*)^2}{2k_B T})dn\right]^{-1}. \tag{1.11}$$

Evaluation of the integral in eqn. (1.11) yields the Zeldovich factor [31]

$$Z = \sqrt{\frac{-\frac{\partial^2 \Delta G}{\partial n^2}\big|_{n^*}}{2\pi k_B T}}, \tag{1.12}$$

and the final form of the CNT nucleation rate is given by:

$$J_{CNT} = k_{n^*}^+ Z N_{tot} \exp\left(-\frac{\Delta G(n^*)}{k_B T}\right). \tag{1.13}$$

The key point concerning classical nucleation theory is that, by invoking detailed balance to eliminate one of the kinetic coefficients, the calculation of the nucleation process which is inherently a nonequilibrium process, has been turned into a problem of equilibrium thermodynamics. The main focus of nucleation theories is to develop expressions for $\Delta G(n^*)$, the work of formation of a critical embryo.

In arriving at eqn. (1.13), two important assumptions were required. First, that the steady state is reached quickly. This is usually true but in some specific cases involving crystallization in complex systems, the lag time associated with reaching the steady state can be on the same order of magnitude as the measurement. Second, the need for the barrier to be high means that CNT is useful in the regime where the system is only mildly supercooled, but may become unreliable as the degree of supercooling increases and the nucleation barrier becomes lower.

The limitations of classical nucleation theory revealed by experiments have lead to an extensive search for better models with which free energies and nucleation rates can be properly evaluated. Due to the difficulties of performing experiments accurately and the nature of the assumptions required to make nucleation theory tractable, nucleation is an important and active area of research [32].

1.1.2 The Spinodal Envelope

The spinodal behavior in a system is observed when the conditions of thermodynamic stability are first violated. In the systems subject of the present thesis work, we seek to find the spinodal temperature for the liquid phase, defined as the temperature at which the nucleation barrier goes away. A detailed description of stability criteria is included in Appendix A-I.

To illustrate the physical nature of the spinodal process, let us consider the Pressure-Volume phase diagram of the Van der Waals gas illustrated in fig. (1.1). The critical point is denoted as point c. The isotherm T_c, shown with a dashed line, is a temperature beyond which the liquid phase is not thermodynamically possible. For temperatures below T_c, we show a typical isotherm T_1 defined by the dash dotted line. The collection of all isotherms below the critical point defines a coexistance boundary (binodal line) connecting points bcb'. The spinodal envelope is defined as the boundary where the isothermal compresibility is zero, shown in blue color in the figure and connecting points ecf. The spinodal line is the region where the thermodynamic stability requirement of a positive isothermal compressibility is first violated, i.e. $\kappa = -\frac{1}{V}\left(\frac{dV}{dp}\right)_T < 0$. fc is the supercooled vapor spinodal envelope, and ec is the superheated liquid spinodal envelope. b and b' are the equilibrium states of the binodal. e and f are the spinodal points for isotherm T_1.

In the shadowed region between the spinodal and the coexistence line the system is metastable and the phase transformation occurs by overcoming the nucleation barrier (see fig. 1.1.a) that we described in the preceding section. In the region beneath the spinodal line the vapour phase is in a state of thermodynamic instability and the nucleation occurs through a barrierless and spontaneous mechanism. This phenomena is known as the spinodal decomposition.(see fig. 1.1.b)

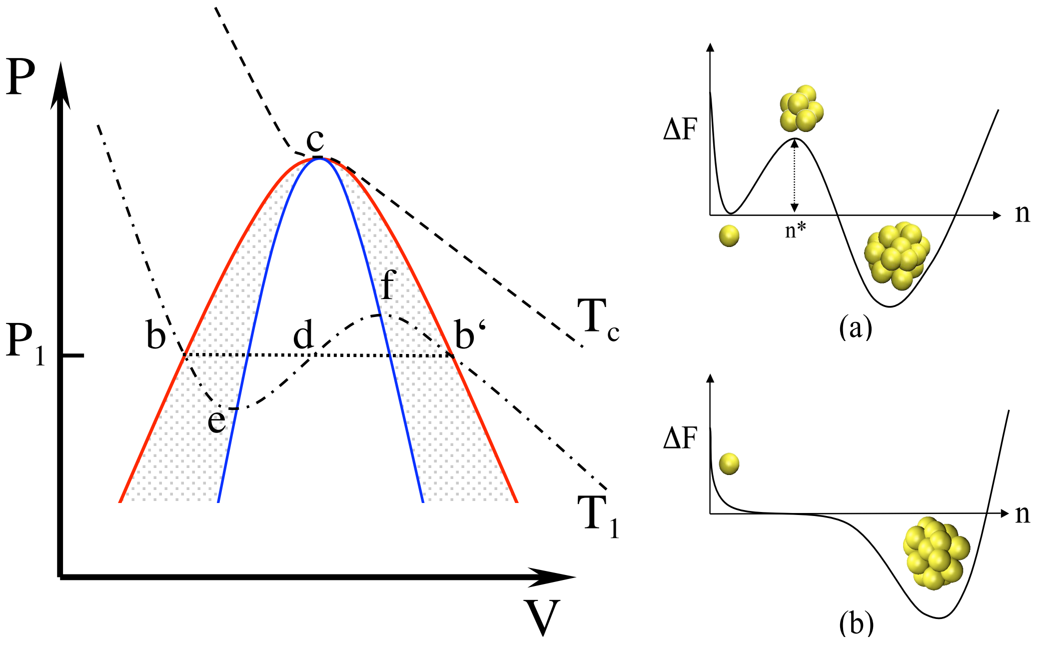


Figure 1.1: Pressure vs. Volume phase diagram for the Van der Waals vapor.

According to mean field theory, as we approach the limit of stability of the metastable phase we will find a sharp boundary between the unstable and the metastable regimens. If we try to quench the system closer to the unstable regimen, the system decays to equilibrium by condensation of droplets (nucleation), hence it is not possible to reach the spinodal on a finite time. Furthermore, as the barrier goes to zero, the size of fluctuations diverges [33, 34].

The basic underlying principle of Mean Field theories (MFTs) is the replacement of complicated many body interactions with an effective (or mean) field, this simplifies problems greatly by removing the role played by fluctuations [34]. MFTs provide the simplified models with which the study of the spinodal behavior in systems may be undertaken [35].

Although defined in theory, reportedly no realistic system with short range forces have been shown to exhibit spinodal decomposition and instead, systems nucleate through small finite nucleation barriers, with the lifetime of the metastable state decreasing in a monotonic fashion [36].

One of the goals of our simulations is to test if a common system such as a gold cluster on nearly ideal experimental conditions is able to crystallize via spinodal decomposition.

1.1.3 Free energy of embryo formation: Phenomenological Approaches

Phenomenological models capture the intuitive idea that the embryo will have the same basic properties as the bulk phase it is forming. Consequently, the embryo is usually characterized using the densities and surface properties of the bulk system despite the fact that the embryo may only contain a few hundred atoms.

The thermodynamics of embryo formation can be derived using the Gibbs droplet model [33, 37, 38], in which the additional cost of building the interfaces of a cluster is considered by adding surface dependent free energy terms (cf. Gibbs droplet model, appendix A-I). In such a case the minimum work of formation of an *n-sized* embryo in an isothermal process is given by the expression

$$\Delta G_{min} = \sigma A + (P - P')v' + n[\mu'(T, P') - \mu(T, P)], \tag{1.14}$$

where σ is the surface tension, A the interfacial area between the embryo and the liquid phase, $P - P'$ is the difference in pressures between the bulk phase pressure P' and the pressure P inside the embryo, v' is the volume per atom of the solid embryo, and μ' and μ the chemical potentials in the embryo and liquid phase respectively.

For a supercooled vapor away from the critical point, we may consider the case of an incompressible embryo, thus we can write:

$$\mu'(T,P') - \mu'(T,P) = v'(P'-P). \tag{1.15}$$

Here v' is the volume per molecule in the embryo phase. Equation (1.14) then becomes

$$\Delta G_{min} = \sigma A + n[\mu'(T,P) - \mu(T,P)] = \sigma A + n\Delta\mu, \tag{1.16}$$

where the chemical potential difference $\Delta\mu$ between the stable and metastable phases at bulk conditions $\{T,P\}$ is negative valued.

The critical size n^* is defined as the size at which the critical embryo has a maximum, and is obtained from the condition $\frac{\partial\Delta G}{\partial n}|_{n^*} = 0$, while the critical free energy ΔG_{crit} is the energy at which this embryo size is reached.

While embryos smaller than the critical size tend to shrink and reincorporate into the liquid phase, embryos larger than the critical value grow as a result of the subsequent reduction in free energy to form the new phase (see fig. 1.2 left). The critical size is thus a system in unstable equilibrium. For $n < n^*$ work can be recovered during the process, whereas for embryos larger than n^*, the embryo grows spontaneously.

In the case of a crystalline nuclei growing from the liquid phase, eqn. (1.16) takes the form:

$$\Delta G_{min} = \sum_i \sigma_i A_i(n_i) + n\Delta\mu. \tag{1.17}$$

Here the surface free energy density term is given by the sum of the contributions of a model structure with i facets, each with their associated surface area A_i, and free energy density σ_i. An interesting case of such polyhedra is the Icosahedron, which having 12 FCC [111] facets, is a very compact structure.

Eqn. (1.17) is the general expression for homogeneous crystallization in the bulk phase. In the case of heterogeneous nucleation we have to consider crystallization at an interface and therefore need to include additional free energy terms. For example, if a solid embryo forms at a wall interface we can write:

$$\Delta G_{min} = \sum_i \sigma_i^{cv} A_i^{cv} + A^{wc}(\sigma^{wc} - \sigma^{cv}) + \tau L_l - |n\Delta\mu|. \tag{1.18}$$

Where σ_i^{cv} is the crystal-vapor free energy density for facet i, with an associated surface A_i^{vc}, σ^{wc} and σ^{cv} are the crystal-wall, and crystal vapor free energy densities respectively, and τ is the line tension along the 3 phase contact boundary.

Even though it is possible to find the most likely minimal free energy structure by means of the Wulff construction [39] in which the cost of adding high energy facets is reduced by increasing the surface occupied by low free energy ones, it is easier in practice to simply assume that the nucleated crystal is spherically symmetric. The model for a nucleation mechanism with spherical symmetry is a central assumption for the Classical Nucleation Theory (CNT) model. The validity of such an approximation rests in the idea that a sphere is a structure with the minimum surface for a given volume, the validity of such an approximation however, is hampered when the granular effect of arranging atoms in a crystalline pattern inside a volume is further considered. (i.e. the approximation is better in principle, for large systems).

The work of formation of an embryo may be expressed as a function of the number of atoms in it, or in terms of the radius of the embryo:

$$\Delta G = an^{2/3} - bn = cr^2 - dr^3, \tag{1.19}$$

where a, b, c, d are constants. $n^{2/3}$ is the surface term, and n is the volume contribution to the free energy (with respect to the radius of the critical embryo,r^3 is the volume term, and r^2 is the surface contribution). Eqn. (1.19) (see fig. 1.2 right) has a maximum at the critical size n^* where:

$$n^* = \left(\frac{2a}{3b}\right)^3 = \frac{32\pi}{3}\left[\frac{(v')^{2/3}\sigma}{(-\Delta\mu)}\right]^3. \tag{1.20}$$

Alternatively, the critical embryo radius is given by:

$$r^* = \frac{2c}{3d} = \frac{2\sigma v'}{-\Delta\mu}, \tag{1.21}$$

The free energy barrier needed to form an $n^* - sized$ embryo is given by,

$$\Delta G^* = \frac{4a^3}{27b^2} = \frac{16\pi}{3}\left[\frac{v'\sigma^3 2}{-\Delta\mu}\right]^2. \tag{1.22}$$

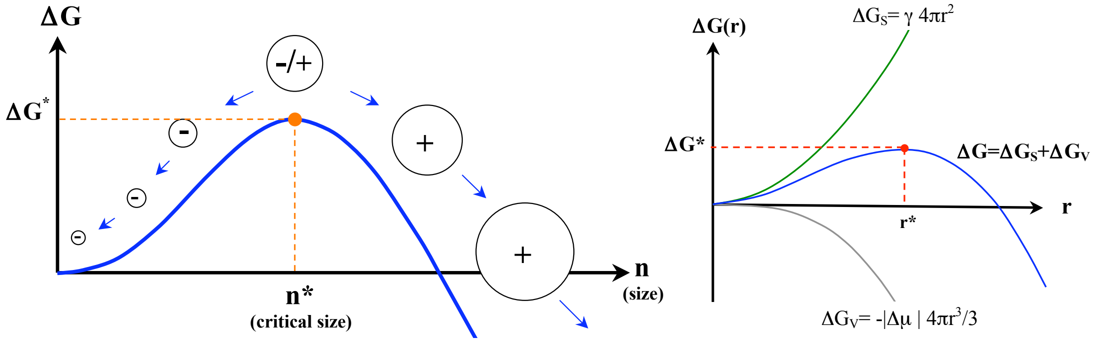


Figure 1.2: liquid-solid transition in the bulk phase. Left: Small embryos tend to decompose back into the native phase, while large embryos tend to increase in size beyond a critical embryo size n^*. Right: According to CNT, at the critical embryo size, there is a balance between the energy gained by changing the chemical potential of a system of particles, and the cost of building an interface.

The free energy cost associated with the formation of the critical embryo decreases as $(-\Delta\mu)^2$ and the size of the nucleus n^* as $(-\Delta\mu)^3$ meaning that as the degree of metastability increases, nuclei get smaller, and the barrier gets lower.

No matter what the model of nucleation is (homogeneous or heterogeneous), a system has to overcome a free energy barrier and form the critical nucleus up to a critical point after which growth becomes spontaneous.

1.1.4 Free energy of embryo formation: Molecular Approach

Whereas phenomenological models define an embryo in terms of bulk properties, molecular theories are a bottom up approach to nucleation theory, and aim to define an embryo based on the local environment of the individual atoms or molecules involved. For example, in the case of freezing, the first step is to identify which atoms appear to be liquid-like and which atoms appear to be solid-like. Then, those solid-like atoms that are all close to each other can be grouped into distinct n-sized embryos. The different criteria for identifying solid-like particles are somewhat arbitrary but are designed to capture the intuitive idea that the local environments around a solid-like atom should be ordered and that this order should be structurally correlated with its neighbours. The details of the embryo criteria used in this thesis are covered in Section (2.4).

Once an *n-sized* embryo can be identified within the cluster, the work required to form the embryo $W = \Delta G(n)$, can be obtained from the probability of its appearance. According to Landau [30], the probability of finding the system in the state defined by an order parameter q, is linked to its free energy via the relation

$$P(q) \propto exp(-\Delta G/k_BT), \tag{1.23}$$

where $P(q)$ is the probability distribution of the system at the state defined by q, $G(q)$ is the corresponding free energy, and c is a constant. The change in free energies (work of formation) from the state q to state q' is given by:

$$\frac{W(q \rightarrow q')}{kT} = \frac{G(q') - G(q)}{kT} = \ln\left[\frac{P(q)}{P(q')}\right]. \tag{1.24}$$

Eqn. (1.23) provides the key connection between the probability of the appearance of a fluctuation, which in this case is the embryo of size n, and its free energy, and forms the basis for developing a molecular approach to nucleation. Frenkel and coworkers [40] introduced an intensive Gibbs free energy function defined as

$$\frac{N_{eq}(n)}{N} \approx \frac{P_n}{N} = \exp\left(-\frac{\Delta G(n)}{k_B T}\right), \tag{1.25}$$

where N is the total number of particles in the system, $\Delta G(n) = G(n) - G(0)$ is the free energy of forming an *n-sized* embryo within the liquid phase, and P_n is the probability associated with the appearance of the embryo. P_n is a quantity which we can obtain from Monte Carlo simulation techniques. Note that this is essentially eqn. (1.8) with $f_{eq}(0) = N$. Later, *Bowles et al.* [41] established the rigourous foundation for this simulation technique, showing that the approximation on the left of eqn. (1.25) only holds in the case where the embryos are rare as follows:

$$P_n = p_n(1) + p_n(2) + p_n(3) + \cdots \approx p_n(1), \tag{1.26}$$

where $p_n(i)$ is probability of observing exactly i embryos of size n. If the formation of different embryos is independent then $p_n(i) = [p_n(1)]^i$ and the higher order terms in eqn. (1.26) can be ignored if $p_n(1) \ll 1$ i.e. if embryos are rare. Similarly, the average number of *n-sized* embryos is

$$N_{eq}(n) = 1p_n(1) + 2p_n(2) + 3p_n(3) + \cdots , \tag{1.27}$$

and again the higher order terms disappear for rare embryos to give $N_{eq}(n) \approx P_n$. A detailed derivation of this approach starting from the partition function is presented in Appendix A-II. The challenge now is to find effective computer simulation techniques that allow the calculation of P_n since these fluctuations are by definition rare and would not be seen in a standard simulation. Section (1.2) introduces the biased Monte Carlo and tempering simulation techniques used in the present work. This molecular approach to calculating the free energy of forming a critical nucleus through eqn. (1.25) is now well established and it has been applied to a variety of systems including the condensation of argon vapour [40], the crystallization of hard-sphere colloids[24] and the crystallization of molten sodium chloride [42]. A comparision between experiment and these simulations suggest mixed results with the calculated rates being some orders of magnitude away from experiment. However, to be fair, experiments measuring nucleation rates are very difficult because the rates are extremely sensitive to external conditions. For example, a variation of a few degrees in temperature may cause a rate to change by several orders of magnitude and so, agreements between experiment and simulation are not as bad as they may first appear to be. Also, it can be argued that molecular potentials may not exactly reproduce realistic systems.

The main problem with the simulation approach is the need for an appropriate order parameter and cluster criteria. What defines an embryo in the simulation is the result of intuition and requires *a priori* assumptions that may not be correct [25]. The role of the order parameter and its effect on barrier calculations is still open to question.

Steinhard et al. [43] introduced empirical order parameters as functions of spherical harmonics $Y_{lm}(\theta(\hat{r}), \phi(\hat{r}))$ with the idea of sensing the symmetry of bond orientations regardless of the bond lengths. The scheme starts with choosing a spherical harmonic $Y_{lm}(\theta, \phi)$ with the angles θ and ϕ defined by the radial vectors between atoms i and j, with $Y_{lm}(\theta(\hat{r}), \phi(\hat{r})) = Y_{lm}(\mathbf{r_{ij}})$. A sum is taken over a suitable number of neighbors and averaged over the total number of neighbors $N_{nb}(i)$ for every particle.

$$\hat{q}_{lm}(i) = \frac{1}{N_{nb}(i)} \sum_{j=1}^{N_{nb}(i)} Y_{lm}(\mathbf{r_{ij}}). \tag{1.28}$$

Previous models for nucleation phenomena at the molecular level, notably [23] use the configurational $\hat{q}_{lm}$ vector to construct rotationally invariant order parameters, in particular Q_6 which is defined by:

$$Q_6 = \sqrt{\frac{4\pi}{13} \sum_{m=-6}^{6} |q_{6m}|^2}. \tag{1.29}$$

The disadvantage of such an order parameter however, is that it gives a measure of the global order in all the structure, and hence is incompatible with the idea that a cluster may be defined as a system containing a well ordered embryo, surrounded by disordered atoms, and instead produces an averaged version of this measurement. Other authors use variations of the Q_6 order parameter, defining alternative order parameters by modifying the index in the sum (1.29) to $Q_l = \sqrt{\frac{4\pi}{2l+1}\sum_{m=-l}^{l}|q_{lm}|^2}$.

The choice of order parameter depends on the properties of the system one is aiming to study. Each order parameter has characteristic signature values which are structure dependent. Table (1.1) summarizes the values of these signature values for some common periodic crystals. Fig. (1.3) shows distributions for these functions. Analytical expressions for all these order parameters, can be found in literature [25], their description however, is beyond the scope of the present work.

Structure	N_b	Q_4	Q_6	Q_8	c_{ij}
ICO (bulk)	12	0	0.199	-	0.50
ICO(surf)	6	0	0.207	-	0.50
FCC	12	0.19	0.57	0.40	0.7
BCC	12	0.08	0.54	0.38	-
HCP	12	0.10	0.48	0.32	0.7
SC	6	0.76	0.35	0.72	-
SC	10	0.40	0.02	0.60	-
LIQ	12	0.02	0.03	0.02	0.3

Table 1.1: Characterstic signatures for Liquid, and various crystals. From left to right, number of neighbors around a particle i, corresponding distribution functions for q_4, q_6, q_8 and at last, the dot product between two neighboring atoms $c_{i,j}$ (taken from ref. [25])

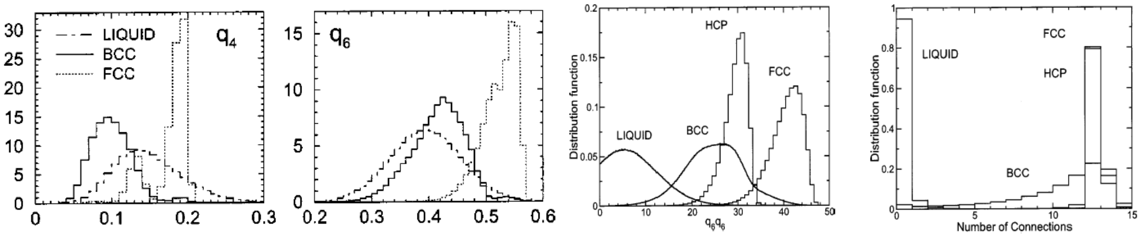


Figure 1.3: Characterstic signatures for Liquid, BCC, FCC and HCP crystals. From left to right, distribution functions for q_4, q_6, c_{ij}, and at last, the number of connections for periodic cells (taken from ref. [25])

1.2 Computer simulation techniques

Studying the formation of the solid phase as the liquid freezes is a formidable challenge from the experimental point of view because it is difficult to identify the critical density fluctuations and the nucleation process is so rapid. However, recent advances suggest that the critical nuclei size can be determined from neutron scattering structure factors [44].

Computer simulation provides a useful tool as it allows us to examine the nucleation process at the molecular level. In particular Monte Carlo (MC) and Molecular Dynamics (MD) methods can be used to calculate energy barriers to nucleation. Other methods that stem from combinations of these methods are described elsewhere in the literature [45].

1.2.1 Monte Carlo Methods

Boltzmann Sampling

Let us introduce the canonical partition function, which uses the NVT ensemble,

$$Q(N,V,T)=\frac{1}{h^{3N}N!}\int\int d\mathbf{p}^N d\mathbf{r}^N exp[-\beta H(\mathbf{p}^N,\mathbf{r}^N)],$$

where $\beta=\frac{1}{k_BT}$, $h =$*Planck's constant*, $\mathbf{r}^N$ are the generalized positions and $\mathbf{p}^N$ are the generalized momenta that represent the phase space. H is the Hamiltonian of the system defined to be equal to the kinetic energy E_k, plus potential energy of the system U.

The probability density is thus given by the expression:

$$\pi(\mathbf{p}^N,\mathbf{r}^N)=\frac{exp[-\beta H(\mathbf{p}^N,\mathbf{r}^N)]}{\int\int exp[-\beta H(\mathbf{p}^N,\mathbf{r}^N)]d\mathbf{p}^N d\mathbf{r}^N}.$$

The expectation value of a thermodynamic quantity A may be cast as:

$$< A >= \int\int \pi(\mathbf{p}^N, \mathbf{r}^N) A(\mathbf{p}^N, \mathbf{r}^N) d\mathbf{p}^N d\mathbf{r}^N \tag{1.30}$$

Eqn. (1.30) is an integral that depends on $6N$ variables for an N particle system. A way to evaluate this intergral is the Monte Carlo method (MC), using the Metropolis criterion. We will now briefly describe the MC scheme.

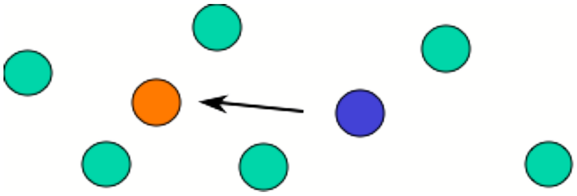


Figure 1.4: Monte Carlo sampling process.

- The energy of the initial configuration U_{old} is first computed.
- A particle is selected at random from this configuration, and given a random displacement (see fig. 1.4), the displacement is given by choosing a random displacement between 0 and 1, such that. $\mathbf{r}_{old} \mapsto \mathbf{r}_{old} + \delta(Rand - 0.5)$. The parameter δ should be chosen in such a way that the sampling is optimal. The Energy of the new configuration is calculated and denoted by U_{new}.
- The test move is accepted with a probability $P_{acc} = min\,[1, exp(-\Delta E/kT)]$, with $\Delta E = U_{new} - U_{old}$.
- If the new configuration is rejected, we reload the old configuration and reset the new energy E_{new} to the old value E_{old}. In either case we measure the property A.

With this in mind, the integral of the expression (1.30) may be computed as an average of the property $A(\mathbf{p}^N, \mathbf{r}^N)$ measured at every configuration $\mathbf{p}^N, \mathbf{r}^N$

$$< A >_{NVT} \approx \frac{1}{N_{MC}} \sum_{i=1}^{N_{MC}} A(\mathbf{p}^N, \mathbf{r}^N), \tag{1.31}$$

where N_{MC} is the total number of measurements, and $A(\mathbf{p}^N, \mathbf{r}^N)$ is the thermodynamic property associated with the configuration $\{\mathbf{p}^N, \mathbf{r}^N\}$.

The errors associated with the evaluation of eqn. (1.31) may be reduced by increasing the sampling, in which the associated error is proportional to the inverse of the square root of the number of measurements of a given property A_i , i.e. $error \propto \frac{1}{\sqrt{N_{MC}}}$.

In the next sections we explain some of the methods with which sampling can be increased in regions of the space phase of our interest.

Statistical Mechanics and Thermodynamics

The link between statistical mechanics and thermodynamics is established by relating the partition function Q and its derivatives to thermodynamic parameters such as the Helmholtz free energy F, internal energy U, and entropy S.

$$\begin{aligned} F &= U - TS = -k_B T \ln Q \\ &= constant + k_B \ln \int\int \pi(\mathbf{p}^N, \mathbf{r}^N) exp(\beta H(\mathbf{p}^N, \mathbf{r}^N)) d\mathbf{p}^N d\mathbf{r}^N \qquad (1.32) \\ &= constant + k_B T \ln < exp(\beta H(\mathbf{p}^N, \mathbf{r}^N)) > . \end{aligned}$$

The internal energy of the system is given by:

$$\begin{aligned} U &= k_B T^2 \left(\frac{\partial \ln Q}{\partial T} \right)_{N,V} \\ &= \int\int exp[H(\mathbf{p}^N, \mathbf{r}^N) \pi(\mathbf{p}^N, \mathbf{r}^N)] d\mathbf{p}^N d\mathbf{r}^N =< H(\mathbf{p}^N, \mathbf{r}^N) > . \qquad (1.33) \end{aligned}$$

The entropy S can be expressed in terms of phase space probabilities $\pi(\mathbf{p}^N, \mathbf{r}^N)$ as:

$$\begin{aligned} S &= k_B \ln Q + k_B T \left(\frac{\partial \ln Q}{\partial T} \right)_{N,V} \\ &= constant - k_B \int \int \pi(\mathbf{p}^N, \mathbf{r}^N) \ln \pi(\mathbf{p}^N, \mathbf{r}^N) d\mathbf{p}^N d\mathbf{r}^N \\ &= constant - k_B < \ln \pi(\mathbf{p}^N, \mathbf{r}^N) > . \end{aligned}$$

An extension to systems that incorporate other variables is straightforward.

Umbrella Sampling

The Umbrella Sampling scheme [46] is usually proposed as an alternative to the Metropolis sampling to handle situations where important contributions to the ensemble average come from configurations which have small Boltzmann factors. The critical nucleus in nucleation appears at the top of a free energy barrier and is intrinsically rare, and a normal Boltzmann sampling would result in measurements with poor statistical accuracy.

If we multiply the ensemble by a weighted probability W that depends on an arbitrary potential, then we can force the system to remain in a region of our interest. In the NVT ensemble, we write the average $<>$ of a thermodynamic property A as:

$$\begin{aligned} < A >_{NVT} &= \frac{\int d\mathbf{r}^N A(\mathbf{r}^N) e^{-\beta U(\mathbf{r}^N)}}{\int d\mathbf{r}^N e^{-\beta U(\mathbf{r}^N)}} \\ &= \frac{\int d\mathbf{r}^N A(\mathbf{r}^N) W(\mathbf{r}^N)^{-1} e^{-\beta U} W(\mathbf{r}^N)}{\int d\mathbf{r}^N e^{-\beta U} W(\mathbf{r}^N)^{-1} W(\mathbf{r}^N)} \\ &= \frac{< A/W(\mathbf{r}^N) >_W}{< W(\mathbf{r}^N)^{-1} >_W}. \end{aligned} \tag{1.34}$$

Here we have assumed a thermodynamic property A, which depends only on the generalized coordinates and not the generalized momenta, and hence are able to avoid the integral with respect to the generalized momenta.

In the case of nucleation, we need to calculate integral (1.34) numerically with respect to the property $A(\mathbf{r}^N) = N_n$, the distribution of embryos of size n in a cluster containing N atoms .This results in the following expression:

$$< N_n >_{NVT} \approx \frac{\sum_i^M [N_n(\mathbf{r}^N)/W(\mathbf{r}^N)]}{\sum_i^M [W(\mathbf{r}^N)^{-1}]}. \tag{1.35}$$

The weighting function is defined for convenience as $W(\mathbf{r}^N) = exp[-\beta\omega(\mathbf{r}^N)]$. Since the probability of the formation of a large embryo is so small, it may be approximated by the probability to find one embryo of certain size in the system. For this reason a bias potential is frequently used to control the size of the largest embryo in the system. A bias potential of the harmonic type is given by:

$$\omega[n(\mathbf{r}^N)] = \frac{1}{2}k_n[n(\mathbf{r}^N) - n_0]^2. \tag{1.36}$$

This harmonic potential is centered around the embryo with largest size n_0. The constant k_n determines the range of sizes sampled in the simulation, whereas the minimum n_0 determines which cluster sizes are sampled most.

Parallel Tempering

Condensed phases generally exhibit a complex potential energy landscape consisting of a large number of potential energy basins associated with the potential energy minimum which are separated by energy basins and saddle points. At low temperatures, a system can become trapped in one or a small number of closely related basins and fails to comply with the condition of ergodicity.

The way to sort out these problems is by creating an extended partition function in such a way that many states are sampled at the same time, with different Thermodynamic parameters. We can then provide additional pathways for the system to escape from a local minima by allowing it to swap between different the different thermodynamic states (see fig. 1.5 left). In the XMC code we developed for the present thesis work (cf. appendix A-III), one processor is used to sample configurations around every umbrella window.

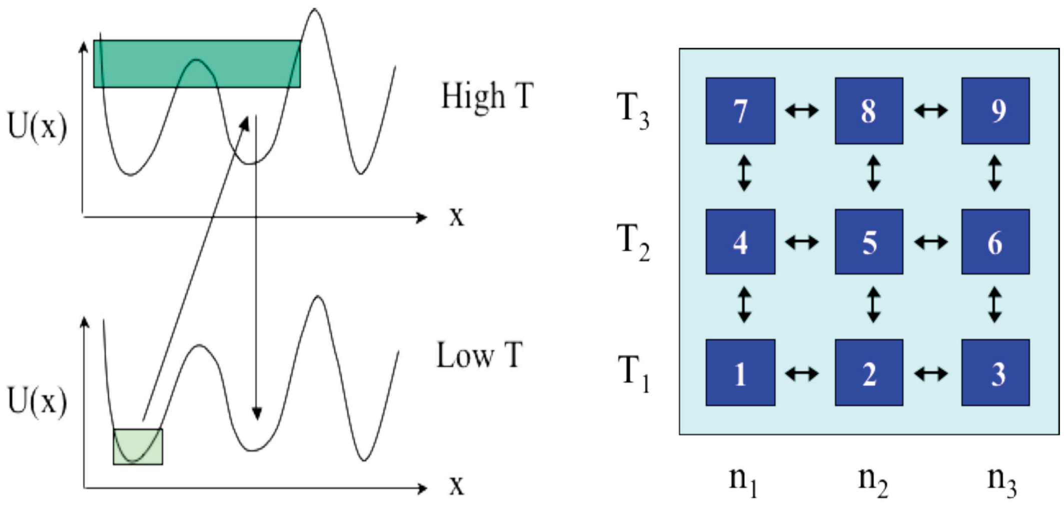


Figure 1.5: Parallel Tempering Scheme: Left, interchanging sampling parameters helps to improve sampling.Right, umbrella center and temperature parallel tempering grid.

Let us define a system i with temperature T_i. A collection of k of these systems ordered according to an increasing temperature scale, $T_1 < T_2 < T_3 \cdots < T_k$ form a system whose partition function Q_{ext} is defined as the product of the individual NVT_i ensembles:

$$Q_{ext} = \prod_{i=1}^{N} Q_{NVT_i} = \prod_{i=1}^{N} \frac{1}{\Lambda_i^{3N} N!} \int \cdots \int dr^N exp[-\beta_i u(r_i^N)],$$

where r_i^N denotes the positions of N particles in system i. In order to sample this ensemble, it is in principle sufficient to perform simulations in all individual ensembles. But it is also feasible to introduce a MC move consisting in the swapping of configurations between two given ensembles with an acceptance rule for this swap obeying the *detailed balance condition* (cf. appendix A-IV). If the configuration of system i is denoted as $\mathbf{i} = \mathbf{r_i^N}$, this condition may be written as:

$$\begin{aligned} \frac{acc[(\mathbf{i},\beta_\mathbf{i}),(\mathbf{j},\beta_\mathbf{j}) \rightarrow (\mathbf{j},\beta_\mathbf{i}),(\mathbf{i},\beta_\mathbf{j})]}{acc[(\mathbf{i},\beta_\mathbf{j}),(\mathbf{j},\beta_\mathbf{i}) \rightarrow (\mathbf{i},\beta_\mathbf{i}),(\mathbf{j},\beta_\mathbf{j})]} &= \frac{exp[-\beta_i U(\mathbf{j}) - \beta_j U(\mathbf{i})]}{exp[-\beta_i U(\mathbf{i}) - \beta_j U(\mathbf{j})]} \\ &= exp(\beta_i - \beta_j)[U(\mathbf{i}) - U(\mathbf{j})]. \end{aligned}$$

Computationally, swap moves are very inexpensive, involving only the interprocessor communication time required to interchange parameters in the simulation, rather than configurations. To sumarize, a chain of MC moves for every configuration is accepted according to

$$P_{acc}(o \rightarrow n) = min[1, exp(-\beta(U(n) - U(o)],$$

while swapping between i and j configurations is made according to

$$P_{acc}(i \rightarrow j) = min[1, exp((\beta_j - \beta_i)(U(i) - U(j))].$$

The parallel tempering scheme is especially useful for cases where many local minimum are present. Swaping between a state in the system with higher energy and that of a lower energy, will allow for more smooth statistics, since simulations may avoid being confined to definite regions in the space of configurations. In the context of cluster nucleation for example, a low temperature T_{low} can be trapped on a region where the characteristic embryo size is $n_{max} = 0$. Interchanging configurations with a system equilibrated at a higher temperature T_{high} is going to allow the sampling to proceed at another local energy minima. (see fig. 1.5 right.)

1.2.2 Molecular Dynamics Methods

Molecular Dynamics (MD) is a scheme for studying the natural time evolution of a system of N particles in a volume V. In classical molecular dynamics forces are calculated through Newton's Laws. If F_i is a force acting on particle i, with a mass equal to m_i then the force is given as $F_i = m_i a_i$, with a_i the acceleration of the particle. MD is a deterministic scheme unlike MC which we have already discussed.

In order to evolve the system in time, equations of motion have to be integrated. One such scheme is the velocity Verlet algorithm [47], given by:

$$\mathbf{r}_i(t+h) = \mathbf{r}_i + h\mathbf{v}_i(t) + \frac{h^2}{2}\frac{\mathbf{F}_i}{m_i}, \tag{1.37}$$

$$\mathbf{v}_i(t+h) = \mathbf{v}_i + h\left[\frac{\mathbf{F}_i(t) + \mathbf{F}_i(t+h)}{2m_i}\right]. \tag{1.38}$$

Here $\mathbf{v}_i$, $\mathbf{F}_i$, $\mathbf{r}_i$ are the velocity, force and position corresponding to the i_{th} particle, the time step h, is in the order of a few femtoseconds which is typical for a liquid, and in all cases should be chosen in such a way that the total energy is conserved on each iteration step. This integration scheme, conserves the energy of the system and is known as a constant energy molecular dynamics, and appropriate for simulations in the $\{N, V, E\}$ ensemble.

The expression for the forces $\mathbf{F}_i$ is obtained through the relationship between a force and the gradient of the potential $\mathbf{F}_i = -\nabla V_i$, where the interaction potential V_i for our system is that described by the EMA potential [48, 49].

Since the integration of the Verlet equations (1.38) depends on the time increment h with an error size of the fourth order, i.e. $O(n^4)$ we have to find an appropriate constant h. If this increment in time is too small we will need more time steps to produce longer simulation times, which will require too many calculations, in the other hand increasing this number too much, may lead to considerable errors.

1.3 Freezing in small clusters

Intensive computational studies aimed at finding minimal energy configurations, show that structures with five fold symmetries, namely Icosahedral (ICO), Decahedral(Dh) and their variants, are prevalent in clusters with diameters smaller than 2 nm (∼250 atoms), while larger cluster sizes, have optimal structures with hexagonal symmetries, such as truncated Octahedra (TO) and variants [50]. These structures (see fig. 1.6) have been observed in samples of gold clusters, synthesized by different techniques and analyzed via electron transmision spectrocopy[12, 13]. The transition between Icosahedral, Decahedral and FCC structures has strong dependence on the number of atoms being considered, as for some of these structures a determinate number of atoms forms compact structures, i.e. The plots of total energies versus cluster sizes are not thoroughly uniform. This is especially marked on the Icosahedral structures, that form closed shell structures at magic numbers.

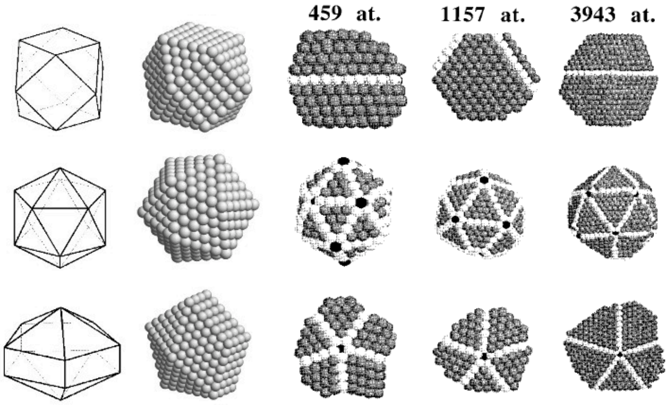


Figure 1.6: Typical structures materializing spontaneously during the solidification of supercooled gold clusters. The first two columns depict representations of 'perfect structures' (taken from ref.[51]). In the central pictures, Dark gray spheres represent gold atoms with an FCC local structure, light gray spheres are atoms in an HCP environment, and black atoms indicate sites with 5-fold (Dh) symmetry.(Taken from ref.[21])

Fig. (1.7 left), shows the relative energy of different structures with respect to the cuboctahedral arrangement, fig. (1.7 right), compares the energies of a series of icosahedra type structures. One must however consider that not only

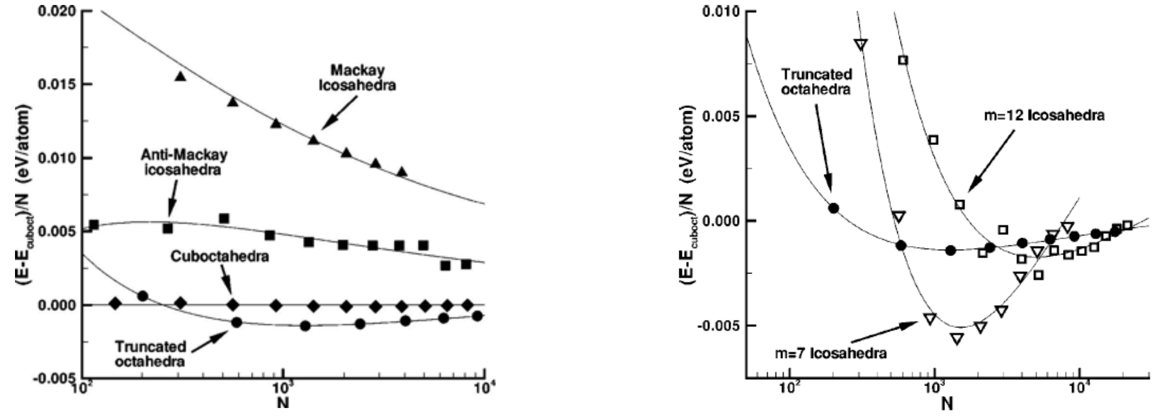


Figure 1.7: Left: Relative energies with respect to the cuboctahedral shapes. Right: Comparison of total energies per atom. Both figs. were taken from ref. [16].

thermodynamics but kinetic factors may determine the crystalline structure in which a cluster freezes.

Surface phenomena may play an important role in the freezing of nanoparticles. For example, the phase transitions observed in MD simulations on gold clusters up to sizes of about 3000 atoms for example, show that different cooling conditions lead to the spontaneous formation of various structrual patterns (ICO, Dh and TO), although Icosahedra were found to be most prevalent[21]. Yet, in another theoretical work, it is shown that the prevailing structure for clusters larger than 450 atoms, is a more energetically favorable TO structure, while other structures such as ICO and Dh did appear as well (see fig. 1.6) despite the fact that the FCC structure is the most stable for the particle sizes under consideration[52]. These simulations nevertheless, suggest that the formation of solid structures proceeds through characteristic stages. In particular the phase transition from liquid to Icosahedral crystals is thought to be started by a reordering of the atoms in the surface which propagates to the inner core of the cluster, pointing to the possibility of rationally directing the process of crystallization by actively changing the thermodynamic conditions at which such transformations take place (see fig.1.8).

In a recent atmospheric study, nucleation rates for the freezing of water droplets containing nitric acid, and suspended in the trophosphere [8] were compared against

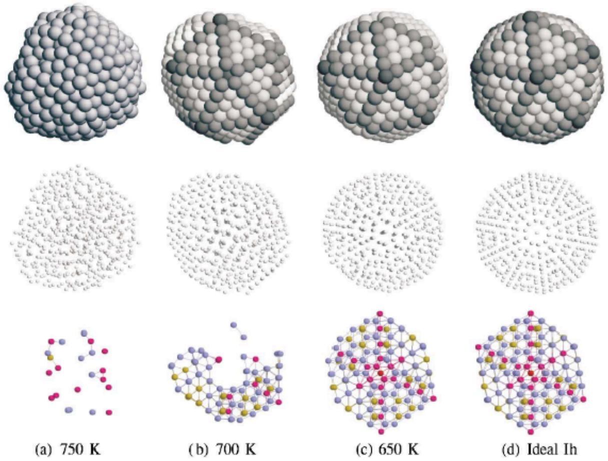


Figure 1.8: Cluster configurations at different cooling stages (a) in a liquid state (at 750°K), (b) just after freezing (at 700°K), (c) after complete rearrangement (at 650°K), and (d) an ideal icosahedron for comparison. In the upper row, only surface atoms are shown, while in the middle row, all the atoms are shown at a smaller size to display their inner arrangement. In the lower row, solidlike atoms with a well-defined local symmetry are shown in two-crosssectioned views by a ball-and-stick model. Blue, gold, and red balls represent atoms with hcp, fcc, and fivefold local symmetries, respectively. (taken from ref. [52])

the model provided by the classic theory of homogeneus crystallization, where nucleation starts at the interior of the droplet (see fig. 1.9.a). When the laboratory data was reinterpreted and the alternative surface nucleation was proposed, the correlation between experiment and the new model was dramatically improved. A phenomenological model, using surface thermodynamics consistent with the capillarity approximation[9], showed that the work required to form a critical nucleus for the crystal within a liquid droplet was greater than that for a crystal embryo formed at the surface (see fig. 1.9.b) and hence had a solid-vapor interface, compared to a critical nucleus forming in the centre of the droplet, when the inequality $\sigma^{vs} - \sigma^{vl} < \sigma^{ls}$, is satisfied, where σ^{vs}, σ^{vl} and σ^{ls} stand for vapour-solid, vapour-liquid and liquid-solid surface free energy densities respectively. This corresponds to the condition where the liquid partially wets its crystal.

Free energy computer simulations for a gold cluster comprising around 500 atoms, and using an atomistic model [23] show that a TO structure is the most stable for temperatures below the melting point. However, the energy barrier between liquid and an Icosahedron structures is much smaller ($\sim k_B T$) than that between the liquid and a TO structure ($\sim 7k_B T$), proving the ubiquitous prevalence in simulations and experiments, of the Icosahedral form , in spite of its energetic metastability at this cluster size. This work, however, does not show a detailed account of the mechanism of freezing for the structures for which the free energies were computed, and neither does ref.[52]. Moreover, another problem brought into question is the fact that these previous works have employed a global order parameter Q_6 instead of the definition of an *n-sized* embryo that we introduce in the present work, which is the central quantity in most nucleation theories.

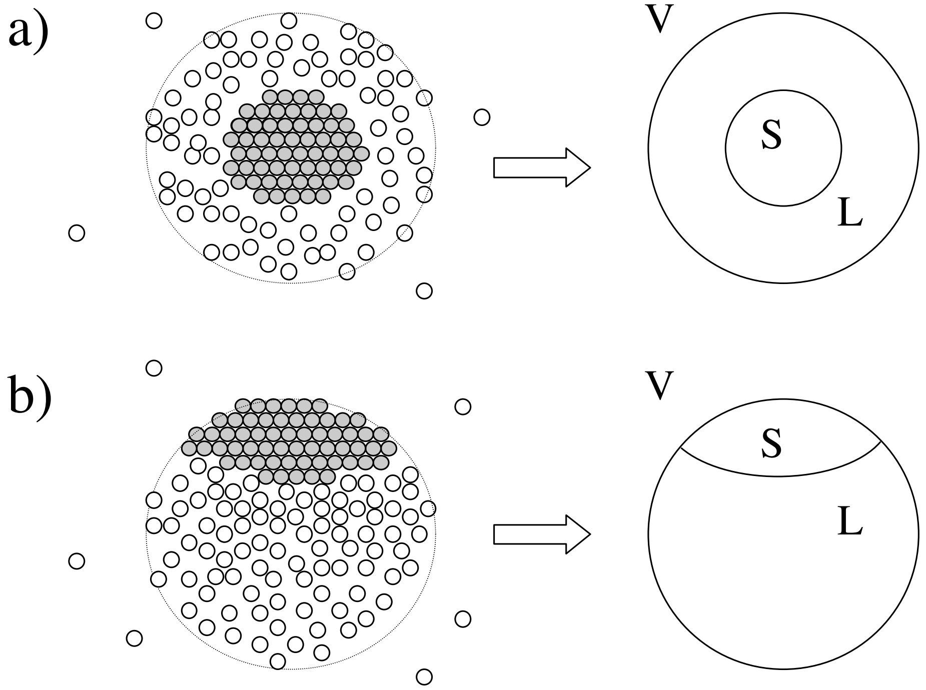


Figure 1.9: Possible nucleation schemes. a) transversal cut of a volume nucleation model embryo growing in a liquid droplet. b) transversal cut of a cluster where surface nucleation takes place. The symbols S, L, V, stand for each of the three phases, namely solid, liquid and vapor.

Chapter II.

Nucleation in Gold Clusters

Our first goal is to study nucleation at the molecular level, in the freezing of gold nanoclusters. To achieve this, we will focus on the calculating the free energy barrier to nucleation using the molecular simulation techniques outlined in the introduction (section 1.2). These techniques also allow use to examine the molecular properties of the n-sized embryos responsible for freezing which will help us address the question of whether freezing occurs in the surface of the cluster or in the core of the cluster, as suggested by classical nucleation models.

This chapter is organised as follows: We begin by defining our model and describing the molecular potential in Section (2.1). To identify the freezing temperatures for a number of clusters of different size, we calculate their caloric curves in Section (2.2). We describe the algorithms used to identify surface atoms in the cluster and identify the n-sized embryo in Sections (2.3) and (2.4) respectively. In Section (2.5) we present our free energy barrier calculations for a 456 atom gold cluster. Section (2.6) uses the classical nucleation model for core nucleation model to calculate the liquid-solid surface free energy density while in Section (2.7) we present evidence suggesting surface nucleation is in fact the preferred nucleation method. At the end of the chapter, we also suggest a phenomenological surface nucleation model.

2.1 Model Definition and Molecular Potential

At temperatures near the melting temperatures of the clusters, the vapour pressure of gold is very low and simulations rarely show any evidence of evaporation events where atoms escape from the cluster. This allows us to treat our system of an N sized gold cluster in a vacuum as a canonical ensemble with $\{N, V, T\}$ all held constant. In both the molecular dynamics and Monte Carlo simulations, we use a square simulation cell with periodic boundary conditions where the cell volumes are selected so that they are large enough to prevent periodic images from interacting, but small enough to prevent the liquid cluster from evaporating.

The use of pair potentials to describe atomic interactions has proved to be successful in the case of liquified noble gases, on which the Lenard-Jones interaction constitutes an important paradigm of simplicity[53]. However, when the complexity of the atomic interactions is increased, and the atoms are no longer weakly bounded, the nature of the chemical bond on every atom, and its particular chemical environment must be accounted for.

Our model makes use of the Effective Mass Approximation (EMA) methodology [48, 49]. The EMA potential is built upon the ideas provided by density functional theory (DFT) which uses the local electron density to deduce the energy of a system. This approach starts by assuming that the total-electron density in a metal is reasonably approximated by the linear superposition from individual atoms. The electronic density in the vicinity of each atom, can then be expressed as a sum of the density contributed by the atom in question plus the constant background density. The total energy of every particle is then given by

$$E_i = F_i(\rho_i) + \frac{1}{2} \sum_{j,(i \neq j)}^{N_{nb}(i)} V_{i,j}(r_{ij}). \tag{2.1}$$

Here the pair potential between atom i and a neighbor j is given by $V_{i,j}(r_{ij}) = \frac{1}{4\pi\epsilon_0} Z_i^a(r_{ij}) Z_j^a(r_{ij})/r_{ij}$, where $Z_i^a(r_{ij})$ stands for the screened nuclear charge of atom i of type a, and $r_{i,j}$ is the inter-atomic distance for the atom pair. The electron density ρ for atom i, is defined by

$$\rho_i = \sum_{j(\neq i)} \rho_j^a(r_{ij}). \tag{2.2}$$

The values for the parameters in the electron densities $\rho_j^a(r_{ij})$, the repulsive pair potential V_{ij} and the embeded atomic potential functional $F_i(\rho_i)$, are obtained by fitting experimental data to experimentally obtained information. (i.e. lattice parameters, elastic constants, cohesive energies, vacancy formation energy and so forth).

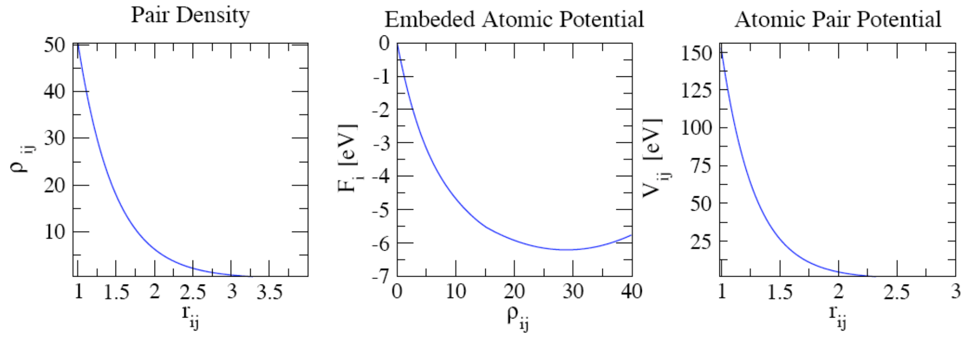


Figure 2.1: EMA potentials for gold: From left to right, pair density, Embeded Potential and Atomic pair potential V_{ij}. Distances are given in $\mathring{A}$ units.

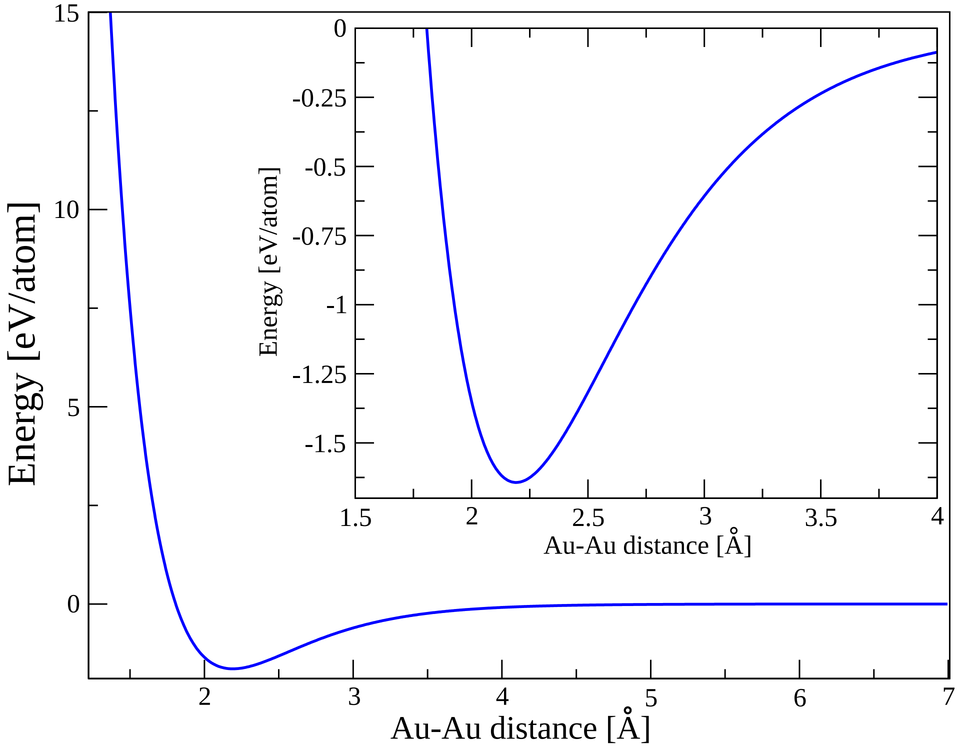


Figure 2.2: Atom dimer potential energy plot using the EMA interaction potential.

It is worthwhile to mention that in eqn. (2.1), the sum runs first over the neighbors of every atom i defined as those atoms within a cutoff radius r_c, which means that the chemical environment of every atom is described in detail, hence this method allows for an accurate description of surfaces and defects, a matter of fundamental importance in cluster simulations.

Finally, the energy of the entire structure is given by:

$$E_{tot} = \sum_{i}^{N} E_i. \tag{2.3}$$

Total Energy calculations for bulk crystalline structures were carrierd out to verify the accuracy of our Monte Carlo and Molecular Dynamics simulations. The elastic coefficient B was obtained by means of the universal equation of state for solids [54, 55]. Table (2.1) summarizes our results and proves the accuracy of our code. We note that the closeness of our results to those from the experiments are not surprising as the potential was especially designed to fit the available experimental information.

Quantity	Exp. [49]	Fit [49]	Current Work $\pm(error)$ *
Interatomic distance: d($\mathring{A}$)	4.07	4.07	$4.07 \pm (1 \times 10^{-2})$
Cohesive energy $(eV/atom)$	3.78	3.78	$3.78 \pm (1 \times 10^{-2})$
Bulk modulus $B(10^{12}dyn/cm^2)$	1.803	1.803	$1.803 \pm (1 \times 10^{-3})$
$E_{BCC} - E_{FCC}(eV)$	0.04	-	$0.038 \pm (3 \times 10^{-3})$
$E_{HCP} - E_{FCC}(eV)$	0.05	-	$0.048 \pm (2 \times 10^{-3})$
$\nu_T(X)$ (THz)	2.75	2.75	-
Thermal expansion coefficient at $773K(10^{-6}K^{-1})$	15.2	13.8	-

Table 2.1: Comparison between experimental and theoretical values for gold: Lattice parameter for FCC lattice, cohesive energy, Bulk modulus, BCC versus FCC energy, and HCP versus FCC energy. Ref. [49] contains a full account of both, experimental measurements and fits. * Current work errors compared to experimental fit data.

2.2 Caloric Curves

Once bulk energies and other parameters have been found to be satisfactory, we turn our attention to the study of the phase behavior of small clusters. In particular, we need to identify the freezing temperatures of the clusters, which we expect to be different from the bulk due to size effects. We compute the caloric curves for different cluster sizes to find a range of temperatures in which the caloric plots show a marked discontinuity that can be associated with a phase transition.

Initial cluster configurations were produced by trimming a FCC crystalline structure into spherical like shapes with different raddi. In this way we can produce clusters of arbitrary size ranging from a few hundred to several thousand atoms.

In order to produce reliable caloric plots using molecular dynamic simulations, we need to ensure that the integration step size for the Verlet algorithm (cf. subsection 1.2.2) produces energy conserving results. This is achieved by computing long time constant energy simulations for different trial step sizes and observing the differences in energies after a determined number of iteration steps.

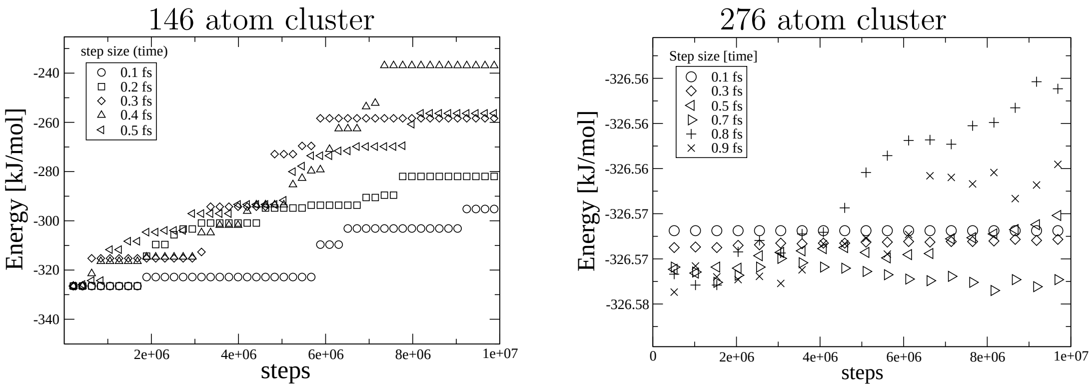


Figure 2.3: Determination of the optimal time step size.

Fig. (2.3) shows adiabatic calculations for two cluster sizes. The key is to select a time step that establish a balance between the error tolerance, and the time need for the computer simulation is to be carried out. Smaller time steps mean more calculations are needed for a simulation of a given length. For example, for a 146 atom cluster, by using a time step of $0.1 fs, (1 fs = 1 \times 10^{-15} seconds)$ we lose

$30 kJ/mol$ of energy on a simulation time of $10 ns$, $(1 ns = 1 \times 10^{-9} seconds)$, clearly an inappropriate time step. On the other hand, the loss is below $1 \times 10^{-3} kJ/mol$ for the same simulation time on a cluster with 276 atoms, resulting in a choice that could probably be too stringent due to the requirement of simulating producing long time simulations to ensure that the system has been thermally equilibrated.

Simulating a constant temperature equilibration for a structure comprising around 5000 atoms on a simulated time of 200 million simulation steps, last about 48 hours of computer time, [1]. For a time step of 1.5 *fs* we can run a simulation time of about $1 \times 10^{-4} s$ on real computer time of 12 hours. This underscores the importance of choosing the appropriate time step size and the need for efficient numerical codes.

In the present simulations for gold clusters, we choose time steps of $0.05 fs$ for the cluster sizes below 150 atoms, time steps of $0.25 fs$ for clusters sizes from 150 to 300 atoms, $0.5 fs$ for systems between 300 and 500 atoms, and $1 fs$ for systems larger than 500 atoms. The square simulation cell has a dimension of $1000 \times 1000 \times 1000 \mathring{A}^3$ for clusters smaller than 1000 atoms, and dimensions of $1500 \times 1500 \times 1500 \mathring{A}^3$ for larger structures.

Fig. (2.4) illustrates the drop in total energy as a function of number of time steps for four different temperatures. In this example we equilibrated the cluster to a temperature of $750°K$ for about 1×10^6 steps, and then reset the temperature drops to 300, 350, 400, 450 and 500 degrees Kelvin. It is apparent from this figure that the most dramatical changes in energy occur during the first two hundred nanoseconds. However, to obtain the caloric plots we need to go to a point of detailed equilibrium, i.e. the system evolves with a characteristic average temperature and a well defined standard deviation (fluctuations in energy).

To obtain our caloric plots, the clusters were initially heated to some temperature T_{max} as to produce completely disordered structures. To generate the caloric curve, we measure the energy at a temperature, and then decrease this by $\Delta T = 50°K$, until a minimum temperature T_{min} had been reached. At each temperature, we run the system for 10 *ns*, until the system is thermally equilibrated to the new temperature.

---

[1] Using a westgrid machine, see http://www.westgrid.ca

We then collect data for the last ten thousand time steps, and divide the accepted data into 100 sections, from where the standard deviations are calculated. The error bars are equal to two times the width of the average standard deviation and shown along the data in fig. (2.5).

Fig. (2.5) shows the caloric curves for cluster sizes with 90, 276, 456, 1220 and 3892 atoms. The temperatures T_{max} and T_{min} for each cluster are shown in the plot, where the approximate location of the transition zone can be seen as a discontinuity which broadens for atoms with fewer number of atoms. When the system has less than 100 atoms the transition zone can no longer be distinguished from the caloric plots, in consistency with the behavior observed in the freezing of small clusters reported elsewhere (cf. section 1.3).

We also observe hysteresis effects in the transition region, which depends on the rate and the direction with which temperature is increased. In fig. (2.6) we illustrate this effect for a 1220 atom cluster on which heating/cooling takes place at the rate of 50 $^{\circ}K/100fs$. When the rate is further reduced the transition point difference between a cooling and a heating process will be increased. Details of this phenomena are described in ref. [56].

The main goal of this section was to locate the freezing temperatures of a number

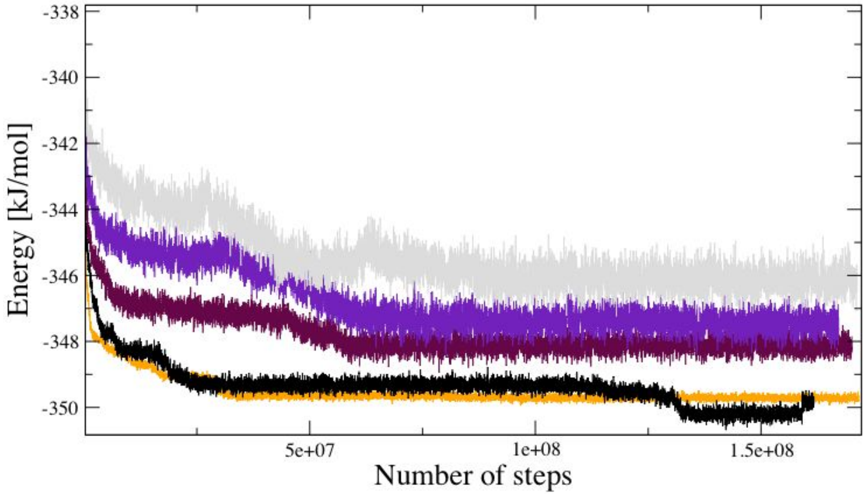


Figure 2.4: Total energy equilibration at constant temperature for a 456 atom cluster. Step size is $0.5fs$

of clusters since we will need to study the nucleation process below this temperature. Most of the work that follows focuses on nucleation in the 456 atom cluster. From fig. (2.5), we can locate the freezing temperature to be approximately $T = 750°K$.

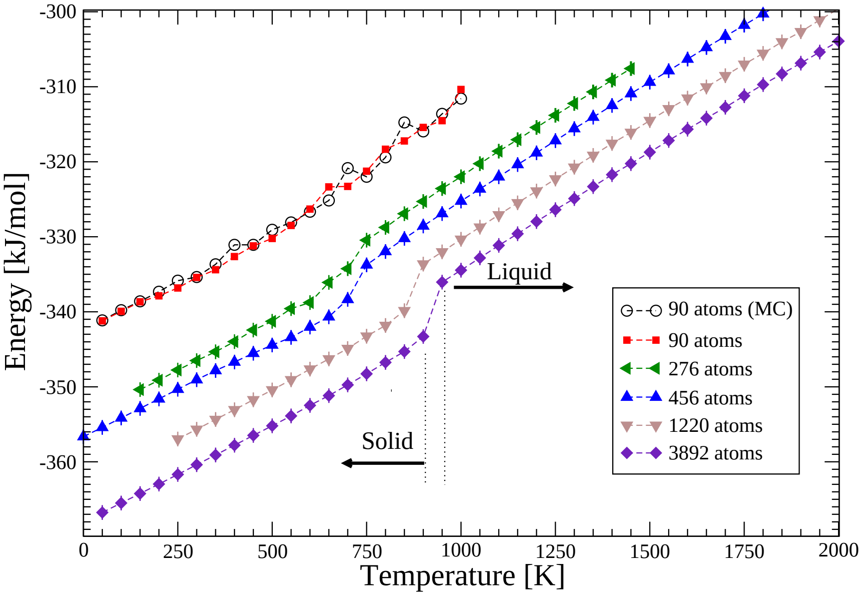


Figure 2.5: Caloric plots for various cluster sizes. Error bars are included in the plots.

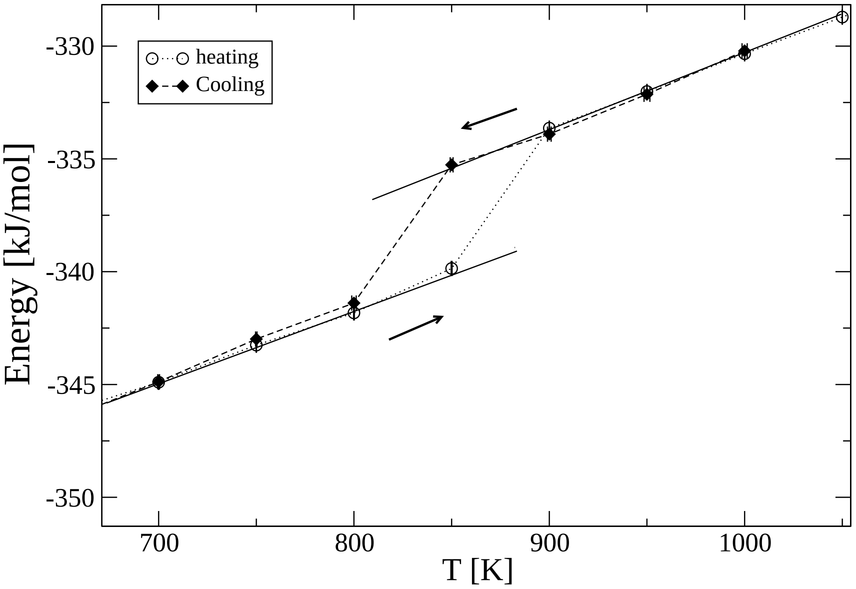


Figure 2.6: Hysteresis effects on cluster nucleation for a 1220 gold atom cluster.

2.3 Surface and Core atoms

The singular size dependent phase behavior attained by clusters on the caloric plots shown in the preceding section, suggests that surfaces play a mayor role in the phase transition behavior of clusters. An analysis of the energetics of clusters with different structures, shows that as the number of atoms in the core of the cluster increases, the energy per atom is greatly reduced.

Fig. (2.7 left), shows the dependence of the energy per atom on three different structural arrangements for clusters as a function of number of gold atoms. The high energy of the simple cubic structure with respect to the FCC and Icosahedra is part of the reasons why this structure is not found in nature for gold. On the other hand, the closeness in energies between the FCC and Icosahedra like structures suggests a competition between the two possible structural patterns. Fig. (2.7 right), shows that as the number of particles in a cluster is increased, the fraction of atoms in the surface is going to be reduced, with the inverse effect occurring for the number of atoms in the core of the cluster. In the limit, when the number of atoms is infinite, we reach the conditions of a periodic infinite structure, and therefore the fraction of atoms in surface is zero, with the number of atoms in the core equal to unity. Clearly, the number of neighbours surrounding the core atoms in greater than that at the surface, thus leading to a lower energy structure.

In this work we employed the cone algorithm [57], to calculate the number of atoms in the core and surface of the different structures considered. The apex of a cone, with azimutal angle $\theta = 60°$, and probe distance $r_c = 5.7\mathring{A}$ (see fig. 2.8), is placed at the centre of each atom. If, having tested all possible orientations, there exists at least one orientation in which the cone does not contain the centre of a neighbouring atom, then we declare such an atom a surface atom. To reduce the computing time employed in identifying the surface atoms in a cluster, we only test for the surface condition those atoms with less than fourteen neighbors. In the bulk, atoms usually have at most 12 neighbors. We expand this number to consider situations in which atoms are loosely bounded.

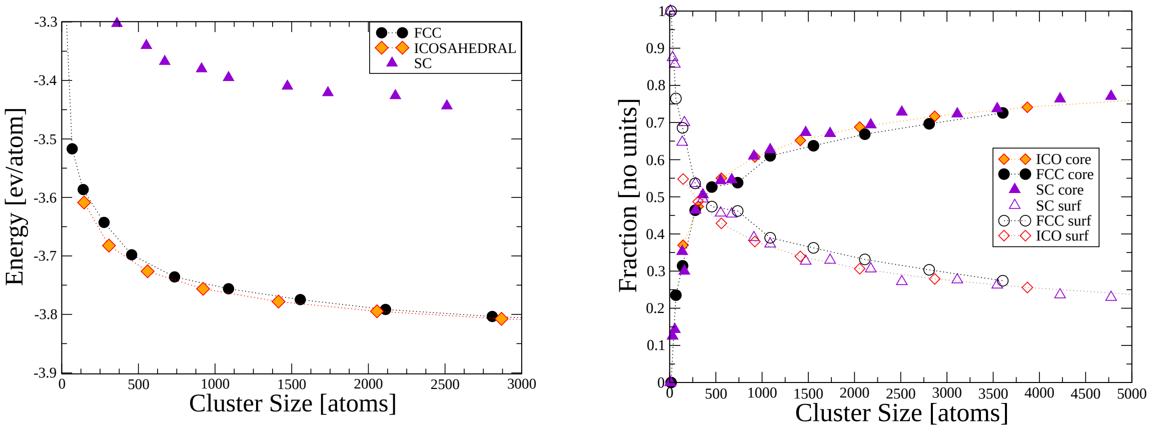


Figure 2.7: Left, energy as a function of cluster size (at $T = 0°$ K) for three different crystalline clusters, right surface and core fraction of atoms as a function of size.

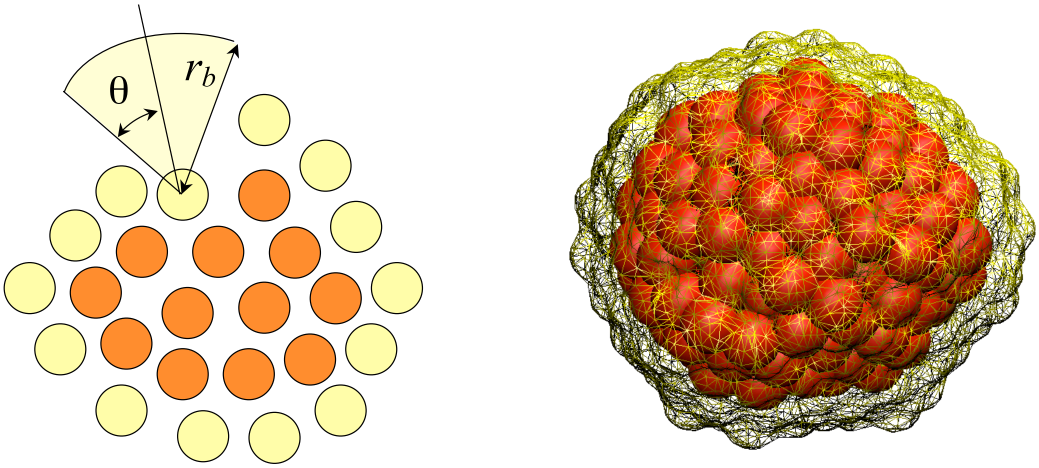


Figure 2.8: Left: A cone defined by an aperture angle θ and a probe distance r_c samples wether an atom in a cluster has a surface or a core-like environment. Right: Snapshot of a liquid 456 atom cluster. Core atoms are dark, and surface atoms light.

2.4 Solid Embryo Criteria

A molecular based approach to nucleation requires the identification of an *n-sized* embryo, constructed atom by atom from the liquid phase. We want the solid definition to be as broad as possible so that our constrained simulations do not preferentially select a particular crystal structure, e.g. FCC or icosahedron, over another. In addition to this, we require the order parameter to work when it encounters the surface of a cluster, where an atom has roughly half the number of neighbouring atoms as compared to an atom in a bulk-like environment.

The key to building an appropriate *n-sized* embryo algorithm lies in the idea that in a crystalline solid, the contiguous environment of two atoms i and j in close contact with each other is very similar, while the order in a liquid is uncorrelated. Also atoms at a grain boundary are characterized by their lack of correlation with respect to the neighboring atoms which are oriented with respect to the crystalline grains in which they belong.

2.4.1 Order Parameter

Following the ideas introduced by *Steinhardt et al.* [43], we propose the use of the dot product $c_{i,j}$ parameter to define the *n-sized* embryo, in the manner of *Frenkl et al.* [25]. Further, we implement a criteria for a threshold number of connections per neighbor to take into account the local environment of atoms in the surface of the cluster.

Dot product definition

To define the dot product parameter we start with a spherical harmonic $Y_{lm}(\theta, \phi)$ with the angles θ and ϕ defined by the radial vectors between atoms i and j $Y_{lm}(\theta(\hat{r}), \phi(\hat{r})) = Y_{lm}(\mathbf{r_{ij}})$. A sum is taken over a suitable number of neighbors and averaged over the total number of neighbors $N_{nb}(i)$ for every particle.

$$\hat{q}_{lm}(i) = \frac{1}{N_{nb}(i)} \sum_{j=1}^{N_{nb}(i)} Y_{lm}(\mathbf{r_{ij}}). \tag{2.4}$$

The next step is to define a complex vector of order $l = 6$ and size unity. We therefore divide this complex vector $\hat{q}_{6m}$ by its magnitude,

$$\tilde{q}_{6m} \equiv \frac{\hat{q}_{6m}(i)}{(\sum_{m=-6}^{6} |\hat{q}_{6m}(i)|^2)^{1/2}}.$$

Finally, the inner complex dot product is defined, this is the order parameter we use in our cluster algorithms and subsequent calculation of free energies.

$$c_{ij} = q_i \cdot q_j = \sum_{m=-6}^{6} \tilde{q}(i)_{6m} \cdot \tilde{q}(j)_{6m} \tag{2.5}$$

The dot product between atoms i,j has the characteristics of having a value close to one when two atoms have similar environment, otherwise its value is small. This idea is inspired in the crystalline order of solids which starts at the microscopic scale, and is extended over the network of surrounding atoms, until -in the case of a bulk crystal- a fundamental periodic unit is repeated throughout space.

Below we describe the procedure followed to choose the three parameters that define the dot product order parameter needed to construct the distribution of embryos within a cluster, these parameters are a neighbor distance radius r_b, a threshold dot product c_{min}, and last, the threshold number of connections per neighbor CxN_T.

Neighboring distance

To calculate the dot product c_{ij}, we need to identify the neighbours for each atom. Neighbor distances are usually defined as the local minima in between peaks of a distribution function. In fig. (2.9), we show the effect of probing the number of neighbors for various distances on a 3892 cluster, this cluster size was chosen because above 70% of its atoms are in a bulk-like environment and we wished to obtain a quick estimate of the proper neighboring distance. Plots were obtained by counting the number of atoms within a distance r_b for any given atom on a crystalline cluster and the best value was set to $r_b = 3.5\mathring{A}$ from probing different radii with $0.5\mathring{A}$ differences.

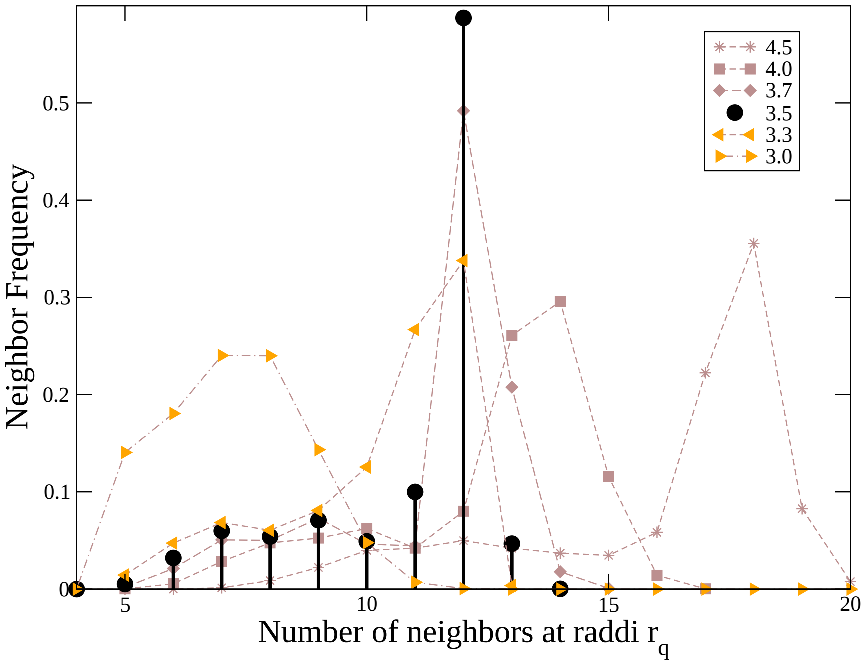


Figure 2.9: Left, histogram for the distance of the nearest neighbors as a function of a probe radius r_b. Filled circles show a peak value in 12 atoms, therefore we choose $r_b = 3.5\mathring{A}$ as the optimal first neighbor distance.

It was found that as the probe distance is increased, the maximum number of neighbors around some atoms starts to increase and as a counterpart, if the probe distance is reduced, the number of neighbor atoms is consequently reduced. The fraction of atoms in the surface of the cluster is characterized by the small tail of the histogram plot for sizes 4 to 10. The optimal probe radii was chosen in base to the number twelve, which corresponds to a close packed structure (FCC, HCP, ICO). Choosing a neighboring distance beyond the first neighbors would impose an undesirable long order correlation to our solid embryo algorithm. Incidentally, the fraction of atoms with 12 neighbors equals the fraction of atoms in bulk like environment, as computed by the cone algorithm method, which shows that the elected distance is appropriate.

In figs. (2.10) we analyze more rigorously the optimal neighboring distance. With aid of the cone algorithm, the histograms of the number of neighbors for atoms in the surface and the core of the clusters. In order to produce these plots, we carried out temperature equilibrations starting from a solid icosahedra-like structure [2]. Simulations were run for 1×10^9 iteration steps, with a $1fs$ step size. The initial temperature was $100°K$ and was increased by amounts of $25°K$ at the end of every simulation on an iterative process until the equilibration for a temperature of $1500°K$ was completed. The reference point in figs. (2.10) is the temperature $T = 725°K$ which indicates a cluster that is not completely melted. Histograms were generated by analyzing 150 configurations obtained by saving configurations every one thousand steps (i.e. saved every 10 ns on simulation time). Similar figures are obtained from the melting of an FCC like structure for the same number of atoms, and also for other different structures and sizes.

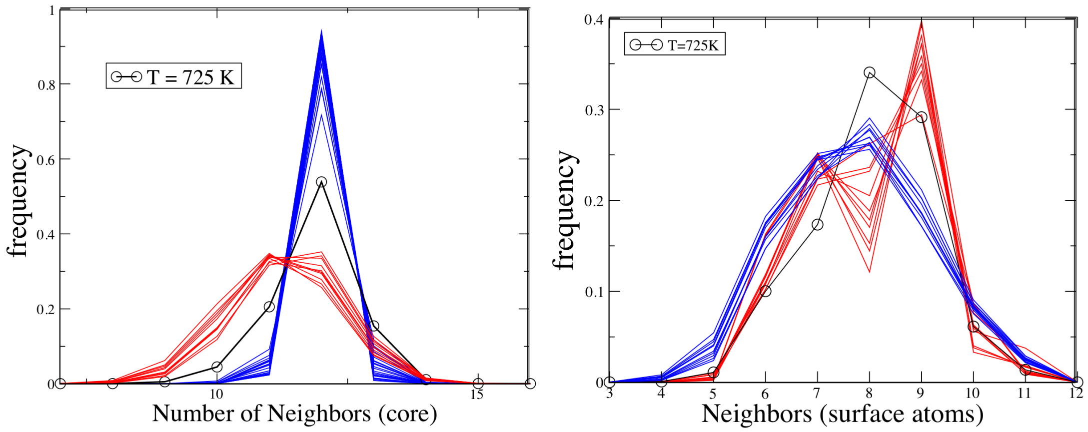


Figure 2.10: Neighboring distance for the 456 atom cluster at temperature intervals of $25°K$. A reference temperature of $T = 725°K$ in the figure is a temperature at which solid/liquid character of the cluster is undetermined. Dark (Blue on-line): Temperatures below melting point. Ligth (Red on-line): Temperatures above melting point. Both histograms have been normalized.

---

[2] A 456 atoms cluster can not be a perfect Icosahedra

Threshold dot product

We computed the dot product distributions from the last one hundred configurations for each one of the temperature equilibrations described in the preceding section, and obtained the histograms shown in figs. (2.11) and (2.12).

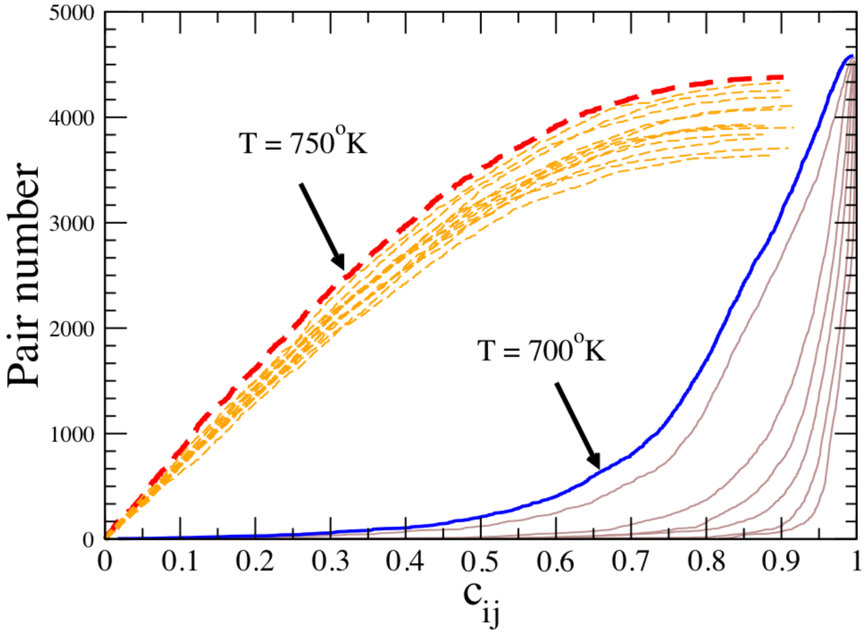


Figure 2.11: Cumulative dot product distribution for the melting of a small Icosahedra structure with 456 atoms. The plot shows the combined dot products of all the atoms in 100 cluster configurations for every temperature.

In fig. (2.11) we show the ordered dot pairs in increasing value, showing that larger values are more prevalent for clusters below the melting point. The temperatures $700°K$ and $750°K$ seem to be the limits for a liquid-like and solid-like behavior. Fig. (2.11) makes emphasis on the fact that for a range of temperatures a threshold value between 0.6 and 0.7 for the dot produce defines the transition from the solid to the liquid phase. Thus we use the threshold value of 0.65 as appropriate for our 456 atom gold cluster. We observe that for temperatures larger than the melting point, taken to be $750°K$ all the distributions look alike, i.e. the clusters are disordered. For temperatures bellow $700°K$ in the other hand we notice that the structure retains its crystalline like character with fewer liquid like atoms. The change of phase is evidenced in the breaking of the trend in the distributions from the solid to the liquid like histograms in either figure.

We further investigated the threshold value for the minimum dot product c_{MIN} for a cluster structure comprising 3892 atoms and with an initial FCC structure (see fig. 2.13). MD calculations were performed independently for many different configurations at constant temperature until the energies converged. In particular, we choose a time step of $1.5 \times 10^{-12}s = 1.5ps$, and performed 50 million equilibration steps. Each of these equilibration steps was repeated 4 times, totaling 200 million equilibration steps. The final results for the reported energies were averaged over one hundred MD iterations. The temperatures were chosen from $50°K$ to $1950°K$ in intervals of $50°K$.

Mainly it was found that for larger cluster sizes, the threshold value for the dot product is easier to define, as compared to small clusters, where fluctuations in energy are much larger. One would expect that in the thermodynamic limit (i.e. for a bulk like system), the histograms for the temperatures before and after the transition temperature interpenetrate at a minimum, a direct consequence of the

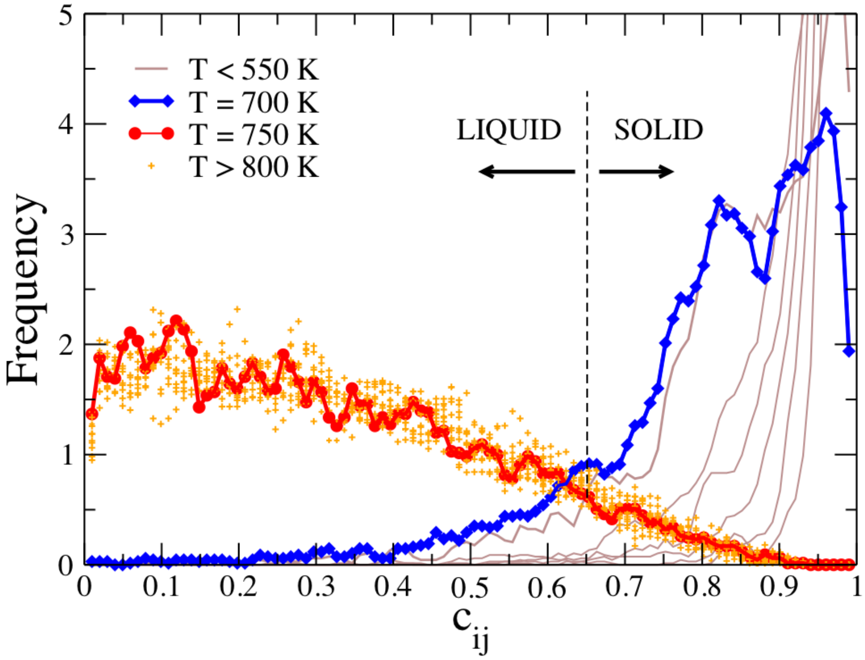


Figure 2.12: Normalized dot product histogram for the dot products shown in fig. (2.11), for the 456 atom cluster.

sharp phase transition for bulk materials, however, in a cluster we observe always atoms with solid or liquid like character.

Moreover, although we can see size dependent effects when comparing the caloric

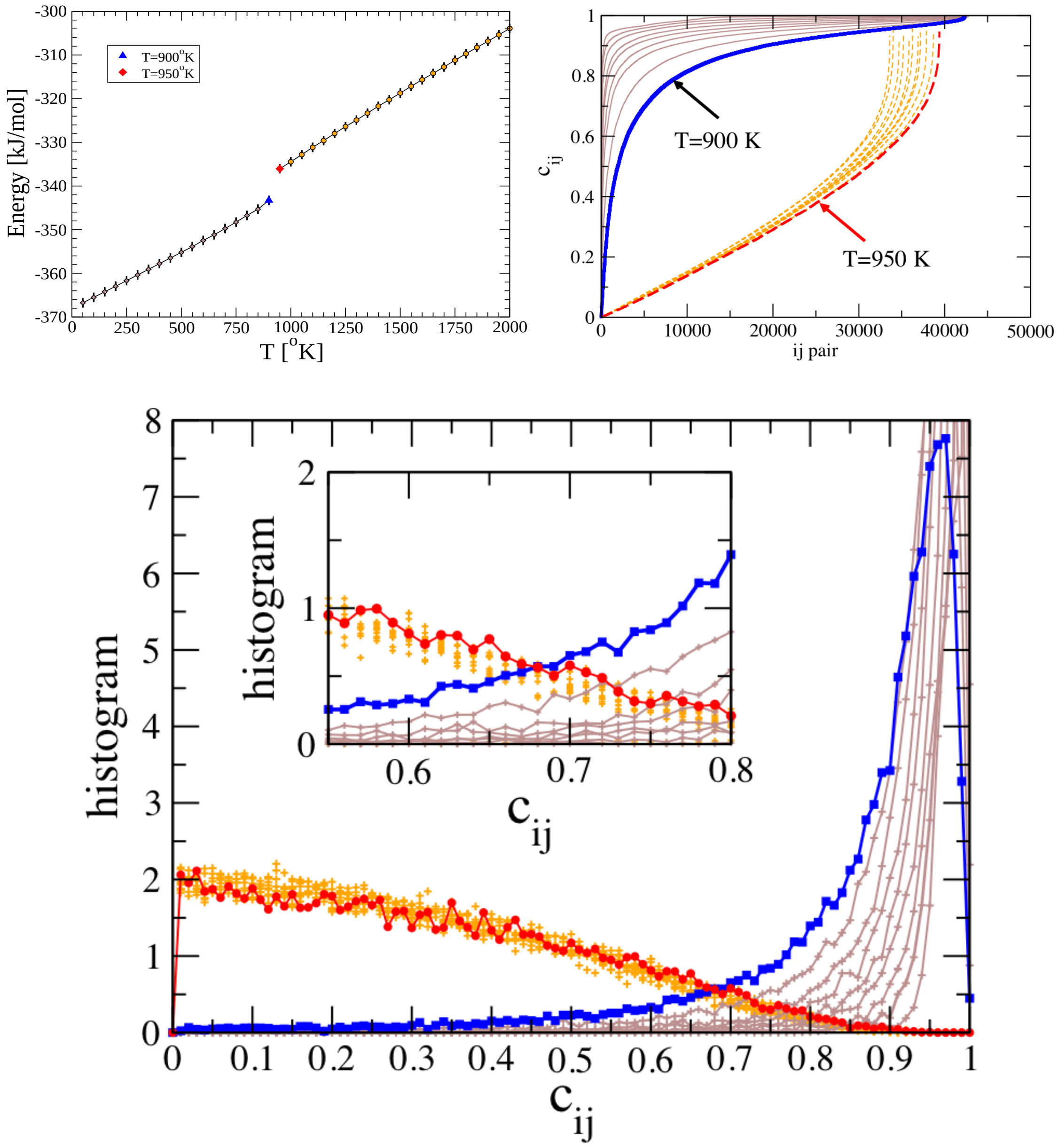


Figure 2.13: Calculation of the threshold dot product for the solid embryo criteria on a 3892 atom cluster. Top left, caloric plot. Top right, ordered dot product distribution. Bottom, dot product histogram for all temperatures in the caloric plot. The inset shows a crossing between the distributions of the temperatures where limits of stability are observed. i.e.$T = 900°K$ and $T = 950°K$. The horizontal axis is the dot product and the vertical axis has been normalized so that the area under the curves is equal to one.

plots of clusters with different atom sizes, the threshold dotproduct does not seem to vary greatly as a function of cluster size, suggesting this may be a constant for a wide range of clusters. In our simulations for the 456 atom gold cluster we choose a threshold dot product equal to 0.65.

Threshold number of connections per neighbor

The threshold number of connections per neighbor was introduced by a need to take into account case scenarios in which atoms in a cluster are located in the surface. Since these atoms will naturally have a smaller number of connections, these can not be characterized with the same criterion as atoms surrounded in all directions, which can form more connections. Furthermore, the spheroidal-like shape of clusters produces atoms with a very sparse number of neighbors in surface and bulk atoms (see figs. 2.10), with the distribution of number of neighbors varying more for highly disordered structures. We need therefore to account for cases in which atoms have a lesser number of connections for every neighbor they are interacting with.

Once again, we take the 150 configuration files for every temperature, from our 456 atom MD simulations, to generate plots for the number of connections per neighbor on surface and core-like atoms. This number is given by:

$$cxn(i) = \frac{\sum_{k=1}^{N_{nb}(i)} \Theta(c_T)}{N_{nb}(i)}, \tag{2.6}$$

where Θ is the heaviside function equal to one when the value of the dot product is larger than the minimum threshold C_{min}. Eqn. (2.6) simply expresses a count of the number of atoms which are connected to an atom i in function of the dot product pairs c_{ij}, this measurement is normalized when dividing by the number of neighbors of such an atom.

Fig. (2.14 left), shows the distribution of the number of connections per neighbor for atoms with core-like environment. Atoms were selected from the ensemble of 150 configurations by means of the cone algorithm, and the number of connections per neighbor arranged in increasing order and plotted for every temperature. For temperatures lower than $T = 725°K$, most of the atoms have large values for the $cxn(i)$ quantity, while for temperatures larger than this value, most atoms have small values. Fig. (2.14 right), shows the normalized histogram, obtained from the distribution in fig. (2.14 left), with a vertical line defining the threshold criteria for declaring an atom liquid or solid. Since the largest number of neighbors per atom has a maximum value of fourteen, increments in the number of connections are given by fractions of 1/14, as can be seen from the histogram generated by the clusters equilibrated at $T = 725°K$. We choose a threshold value $CxN_T = 0.5$, for values larger than this threshold, atoms are declared solid-like, irrespective of their local environment.

Figs. (2.15) are complementary to figs. (2.14), and describe in turn, the phase behavior reflected in terms of the number of connections per neighbor for atoms with a surface-like environment. A comparison of the distributions of (2.14 left), and (2.15 left), for temperatures slightly below the melting point $T = 725°K$, show that as temperature approaches the melting of the structure, disorder starts in the surface of the cluster.

Our inclusion of the number of connections per neighbor serves to introduce a measure independent of the environment of the atoms and serves to calculate the embryo distribution in a cluster as described in the embryo algorithm from section (2.4.2).

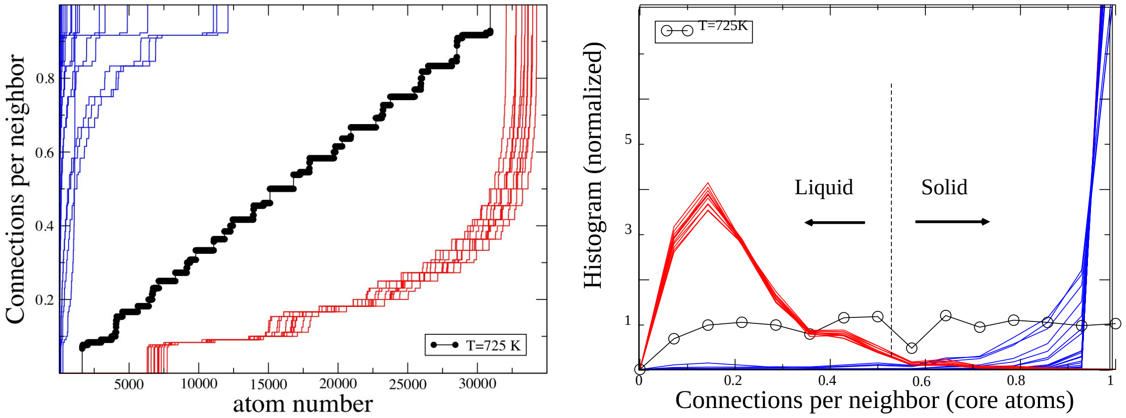


Figure 2.14: Left: Cumulative distribution of the number of connections per neighbor for atoms in a core-like environment for the 456 atom cluster. Data was obtained from 150 cluster configurations equilibrated at various temperatures. Right: Normalized histogram for the number of connections per neighbor for the figure in the left hand side.

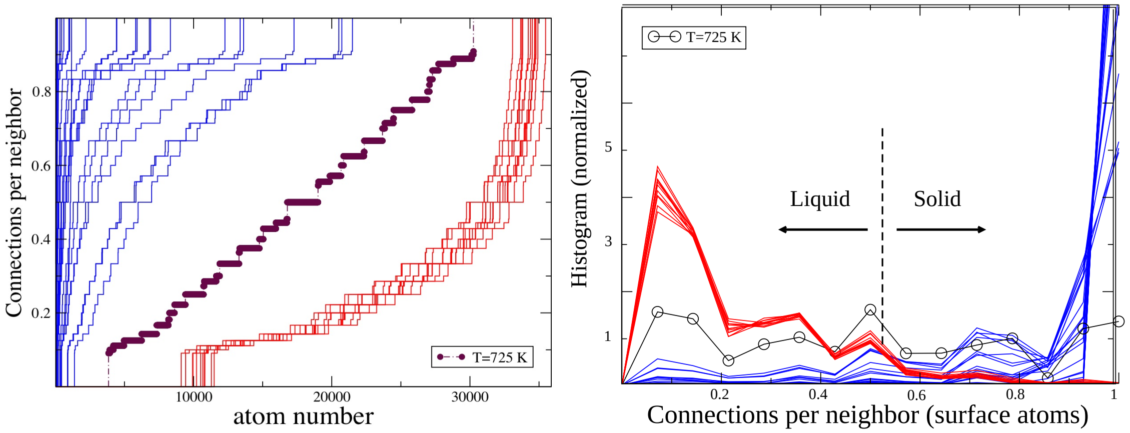


Figure 2.15: This figure is complementary to fig. (2.14). Left: Cumulative histograms for the number of connections per neighbor for atoms with surface-like environment. Right: Normalized histogram for the number of connections per neighbor for the figure in the left.

2.4.2 Summary: Solid embryo algorithm

The distribution of solid embryos is computed as as follows (see fig. 2.16):

(a) First we start by finding the distance $d(i, j)$ between an atom i and a candidate neighbor j, if this distance is less or equal than a minimum distance r_b then atoms i and j are said to share a bond.

(b) For all the bonded atoms of an “i” atom, a threshold dot product c_{MIN} discriminates those atoms which are connected from those which are non connected. i.e. for $c_{ij} \geq c_{MIN}$ two bonded atoms are connected.

(c) When a critical number of bonds for an atom are declared as connected, the atom is declared a solid, otherwise it remains classified as a liquid atom. This critical number of connections per neighbor is defined as $cxn(i) = N_{con}(i)/N_{Nb}(i)$ and compared to the threshold value CxN_T i.e. For $cxn(i) \geq CxN_T$ atom i is a solid.

(d) Two solid atoms that are connected are considered to be in the same embryo. In this way it is possible to extend the network of solid atoms.

The neighbor distance r_b, minimum dot product c_{MIN} and threshold number of connections per neighbor CxN_T used in the present work are 3.85 $\mathring{A}$, 0.65 and 0.5 respectively. The dependence of these parameters with cluster size is not significant.

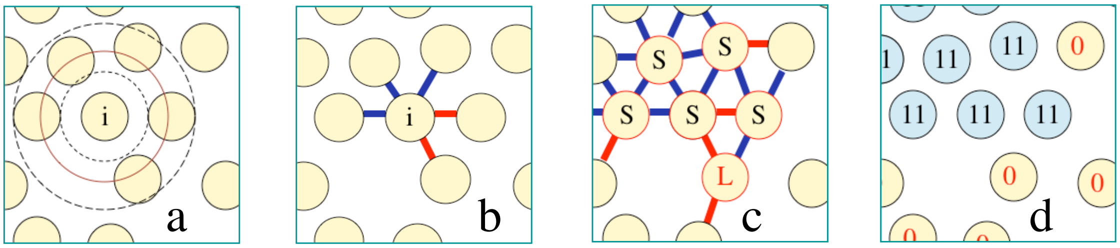


Figure 2.16: Building an order parameter to measure crystallinity. a) Neighbor Identification. b) Finding connections. c) Solid and Liquid definition. d) Computing embryo distribution.

2.5 Free Energy Barrier to Crystallization

In the present section, we investigate the freezing of gold nanoparticles by calculating the free energy barrier to nucleation using Monte Carlo simulation techniques. Nucleation phenomena is a process activated through the overcoming of a barrier to nucleation. However, since the probability of appearance of the critical embryo size is very small on a system under mild supercooling, an appropriate method to generate statistics has to be applied. We employ umbrella sampling plus the parallel tempering schemes described in Section (1.2). Our implementation of these methods is described in detail as follows:.

Let us denote by H_{EMA} a Hamiltonian defined by the sum of the kinetic plus potential energies of a system of particles, $H_{EMA} = E_{EMA} + E_K$. The EMA potential energy E_{EMA} is described in section (2.1), and the kinetic energy of a cluster containing N particles at a temperature T_i is obtained by means of the relation $E_k = N\frac{3}{2}k_B T_i$.

According to the umbrella sampling scheme (cf. section 1.2), we take an unconstrained hamiltonian H_{EMA} and add a potential to create a constrained Hamiltonian H_C,

$$H_\mathrm{C} = H_\mathrm{EMA} + \phi(n_\mathrm{max}), \tag{2.7}$$

the bias potential ϕ depends on two parameters and is function of the largest embryo in a distribution of solid embryos taken on a system containing N atoms:

$$\phi(n_\mathrm{max}; \kappa, n_0) = \frac{\kappa}{2}\left(n_\mathrm{max} - n_0\right)^2. \tag{2.8}$$

The parameter κ is termed of as the "umbrella constant" and n_0 is known as the "umbrella center". With this harmonic form for the bias potential, the umbrella center will force the sampling to be in the direction of the umbrella center while the size of the parameter κ limits the likelihood of sampling embryo distributions for values far away from the umbrella center. Umbrella centers were chosen with values located at intervals $\Delta n = 10$, i.e. $n_{0,\nu} = \{0, 10, 20, 30, 40, 50, 60, 70\}$ and the optimal value for the κ constant was fixed to a value of 0.001. With this choice of $n_{0,\nu}$ and κ parameters, swapping of configurations with common temperature and neighboring umbrella centers n_0, fluctuated between 60% and 70% of efficiency, which guaranteed efficient sampling without having to rely on the choice of a temperature or umbrella center dependency for the κ parameter. The choice of umbrella parameters is described on Appendix A-V.

In addition to the umbrella sampling method, we improved sampling by implementing the parallel tempering algorithm with eight simulated temperatures, $T_\mu = \{650, 660, 670, 680, 690, 710, 730, 750\}$ (in degrees Kelvin). The highest temperature in the simulation ($T = 750°K$) was chosen in such a way that the critical size of the system is well beyond the range covered by the sampling interval, (i.e. $n^* > 80$), this choice satisfies the requirement of the parallel tempering scheme that at least one temperature be chosen in such a way that for the sampling region the system can move freely between all the potential energy minima. This does not mean, however that the system will not nucleate for some critical size beyond the range that is it is being sampled.

The temperatures below $750°K$ were chosen in such a way that the swapping between configurations with the same umbrella center parameter $n_{0,\nu}$ and neighboring temperatures was roughly 60%. MD simulations were initially used as a guidance to select the range of temperatures at which the simulation of nucleation was more convenient to study.

The biased ensemble we are sampling is given by the product of all the subsystems with a number totaling the product of the total number of temperatures by the total number of umbrella centers. In our simulations we have 8 different umbrella centers, and 8 different temperatures, i.e. $8 \times 8 = 64$ nodes:

$$\begin{aligned} Q_C &= \prod_{\mu=1}^{8}\prod_{\nu=1}^{8} Q_{N,V,T_\mu,H_c(\mu,\nu)} \qquad (2.9) \\ &= \prod_{\mu=1}^{8}\prod_{\nu=1}^{8} \frac{1}{\Lambda_\mu^{3N} N!} \int \cdots \int dr^N exp[-\beta_\mu H_C(\mu,\nu\}]. \end{aligned}$$

The averaged embryo distribution $< N(n) >$ in the unconstrained space is obtained from the embryo distribution N_n in the constrained system via the relation:

$$\langle N(n) \rangle = \frac{\langle N_n \exp\left[\phi(n_{\max})/k_{\mathrm{B}}T\right] \rangle_C}{\langle \exp\left[\phi(n_{\max})/k_{\mathrm{B}}T\right] \rangle_C} \tag{2.10}$$

$$\cong \frac{\sum_{k=1}^{X_{MC}} \sum_{n=0}^{n_{max}} N_n \exp\left[\phi(n_{\max})/k_{\mathrm{B}}T\right]}{\sum_{k=1}^{X_{MC}} \sum_{n=0}^{n_{max}} \exp\left[\phi(n_{\max})/k_{\mathrm{B}}T\right]}, \tag{2.11}$$

where $\langle . \rangle_C$ in eqn. (2.10) denotes an average in the constrained ensemble. The averaged embryo distribution in the simulation is measured via the sum in eqn. (2.11), where N_n represents the number of embryos of size n in the constrained MC sampling iteration k, and ϕ_k is the biased potential for the *k-esim* sampling event, with the largest embryo having a size n_{max} and an associated biased potential ϕ_k, which multiplies the distribution of embryos on the entire configuration. The simulation is run for a total of number of X_{MC} Monte Carlo trajectories.

A MC step is defined as the process of selecting particles at random on a system and attempting to displace them by a random displacement ΔR a total number of N times in a system comprising N atoms. At the end of 10 MC steps, we test the final atomic configuration for the largest embryo size and only then apply the constrained potential. The process of performing 10 MC energy equilibrations, followed by a sampling under the constrained potential is termed of as one trajectory. The largest embryo distribution is only tested at the end of the trajectory instead of every MC step due to the computational effort required in the evaluation of the distribution of embryos in a cluster configuration.

We performed swapping attempts in umbrella centers every 10 trajectories, and swapping attempts every 10 trajectories, in temperatures for configurations with neighboring temperatures and same umbrella centers. The simulation was run for a total of about to 456,000 trajectories for every node, and repeated for four different initial conditions to ensure reliability of results. Further, we dropped out the initial $20,000$ MC steps for every node, from the ensemble statistics that produced the final free energy diagrams since the system has to be initially equilibrated. Three out of four of the simulation runs were started from equilibrated configurations, however, in spite of this, the initial $20,000$ trajectories were not taken into account to build free energy plots, this was made to ensure independency on each of the simulation runs. We ran, therefore, the system for a total of $(456,000 - 20,000)$ trayectories/node, which multiplied by 8 different umbrella centers and 4 independent simulations leaves a total of about 14 million trajectories for every temperature.

Once we have obtained the piecewise histograms with respect to every node, we extract the free energies via the relation,

$$\frac{\Delta G(n)}{k_B T} = -\ln\left(< N(n) >\right) + b. \tag{2.12}$$

The constant b, the temperature, and the average embryo distributions $< N(n) >$ are different for each one of the nodes in the simulation, i.e. $b = b(T_\mu, n_{0,\nu})$, $T = T_\mu$, and $< N(n) >=< N(n; T_\mu, n_{0,\nu}) >$.

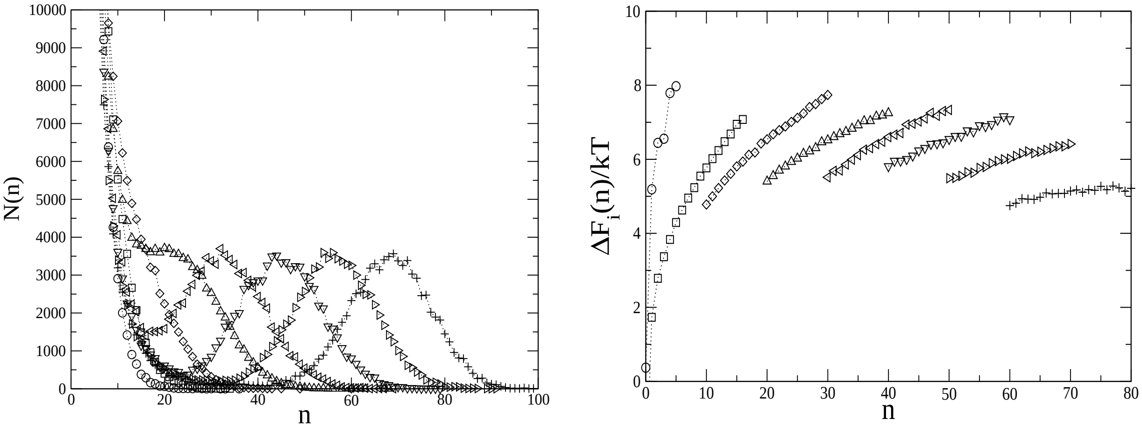


Figure 2.17: Left: Distribution of embryo sizes for every one of the umbrella centers at the simulated temperature of $T = 750°K$. This histogram is taken in the constrained ensemble. Right: Piecewise free energy of crystallization obtained from every simulated bin in the constrained embryo distribution.

In fig. (2.17 left), we show the constrained embryo distribution for all the statistics in the nodes with temperature $T = 750°K$ for the 456 atom cluster, the resulting piecewise free energies from these statistics once relation (2.12) has been applied are shown on fig. (2.17 right).

The optimal alignment of the common temperature free energy segments is achieved by obtaining the coefficients of the polynomial [25]:

$$\Gamma(n) = \sum_{k=1}^{k_{max}} a_k n^k + c, \tag{2.13}$$

according to:

$$\Omega = \sum_{n=0}^{max(n)} \left\{ \sum_{i=1}^{n_w} w_i(n) \left[\frac{\Delta G_i(n)}{k_B T_i} - \Gamma(n) \right]^2 \right\}. \tag{2.14}$$

Eqn. (2.14) is the squared error of the difference between the "measurements" obtained by the simulation (see eqn. 2.12) and the "expected" values given by the function $\Gamma(n)$. This difference is further multiplied by a weight equal to the inverse of the square of the standard deviation for the measurements in every simulation window. i.e. $w_i(n) = 1/\sigma^2_{\Delta G_i(n)}$. The standard deviation σ was estimated from the four independent simulations which were cut down in halves to produce a total of eight sets from which σ could be computed. The estimate of this error is below 0.6 $k_B T$ for statistical measurements around the umbrella center , and grows up to 1 $k_B T$ if the extremes of the umbrella sections are considered. In any case the extremes are not taken into account beyond a difference $|n_{max} - n_0|$, this is due to the fact that not enough sampling can be made in the constrained space for cases where the embryos are far from the umbrella center. This is taken care of, by defining the standard error as infinite in such events, i.e. the inverse of the square of the variance is zero. The sum in eqn. (2.14) is made for all the possible embryo sizes, from $n = 0$ to $n = max(n) = 90$, and for $i = 1$ to n_w, the number of umbrella windows, equal to 8.

Fig. (2.18) shows the free energy barriers to crystallization taking the liquid cluster as a reference. It is shown that as the temperature is lowered, the Helmholtz free energy barrier and the critical embryo size are reduced, with the solid phase increasing its relative stability. It is important to note that for small embryo sizes, the features of the free energy curves for all temperatures can not be explained via classic nucleation theory.

Table (2.2) contains a summary of the free energy barriers in fig. (2.18), with respect to every one of the eight temperatures in the simulation.

$T[^\circ K]$	$n^*[atoms]$	$\Delta G(n^*)/kT[]$
650	23	10.0
660	30	10.2
670	37	10.4
680	42	10.5
690	47	10.7
710	58	11.6
730	70	13.0

Table 2.2: Free energy barriers with respect to the critical embryo size n^* as a function of temperatures. These results were obtained via MC calculations for a 456 atom cluster.

The plot for fig. (2.19 left), shows a seemingly linear-like relationship of the

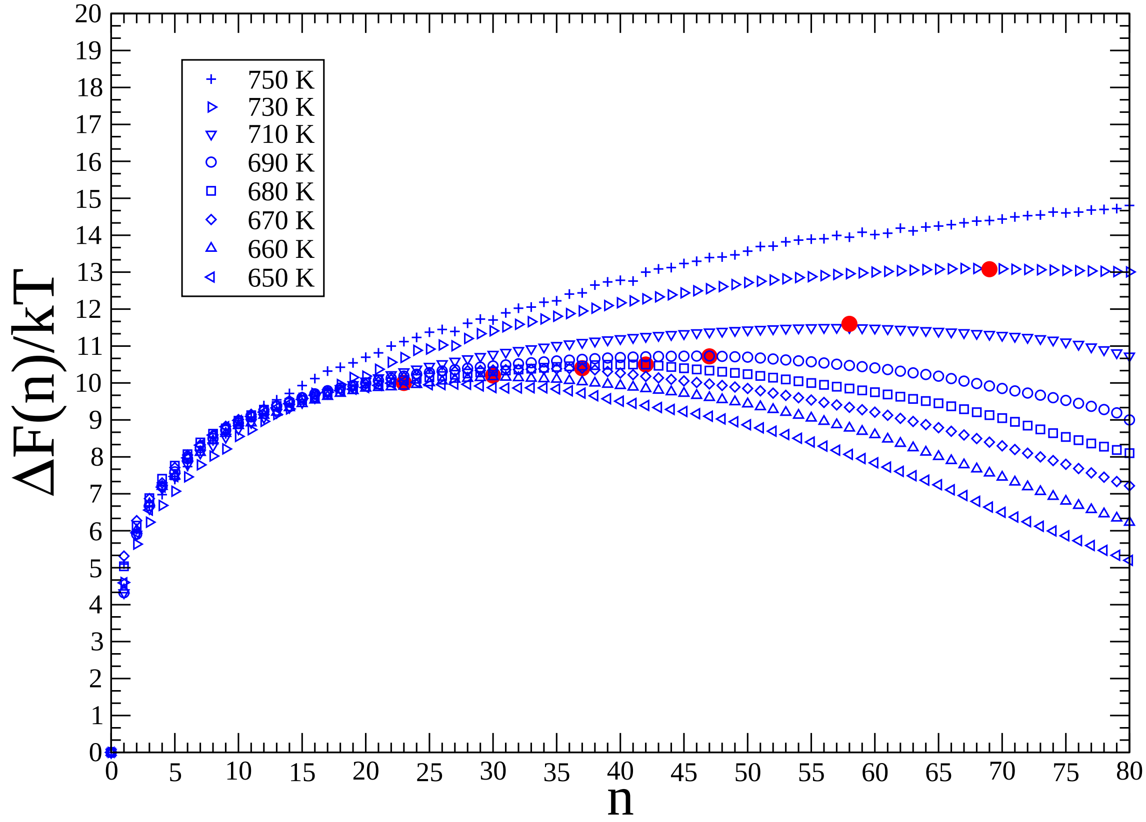


Figure 2.18: Free Energy barriers for various temperatures. 456 atom cluster. The filled circles indicate the maximum in the free energy for each temperature.

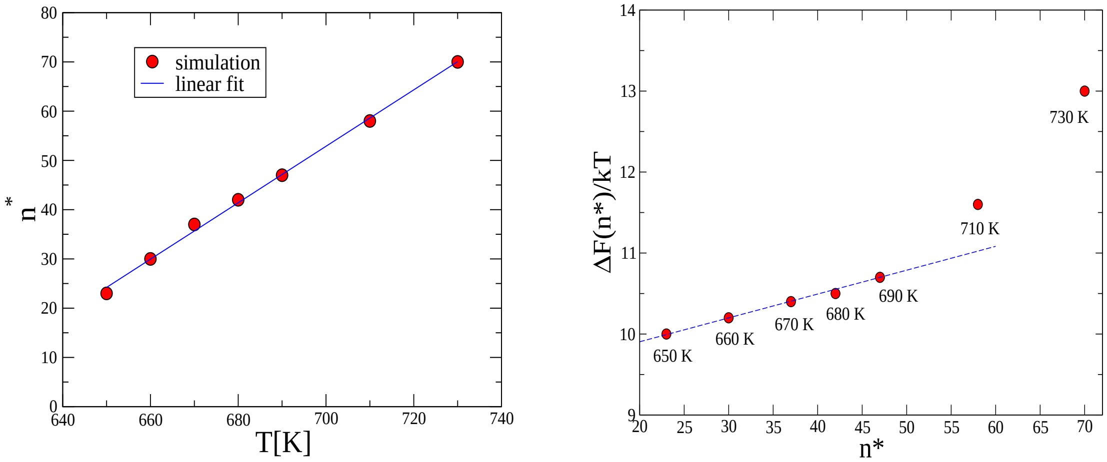


Figure 2.19: Barrier height as function of n_{max}. 456 atom cluster.

critical embryo size n^* with respect to temperature for all the temperatures in the simulation. Fig. (2.19 right), in the other hand, shows that a linear relationship of the free energy with respect to the critical embryo is only observed for high supercooling, i.e. at low temperatures. The size of the free energy barrier after temperature $T = 690°K$ follows a polynomial-like growth.

In any case if we are to follow the linear like trend in the prediction of the free energy with respect to the critical size at high supercoolings (i.e. at temperature lower than $690°K$), we find that the free energy barrier to nucleation for a critical size equal to zero will have an extrapolated free energy difference larger than $9.5kT_B$.

2.6 Core Dependent Nucleation.

Phenomenological models for nucleation provide a useful method for the calculation of surface free energy densities. In this section, to calculate the solid-liquid surface free energy density, we assume that nucleation obeys the CNT ansatz, i.e. nucleation is a process started at the core of the cluster.

Within the classical nucleation theory approximation and assuming a spherical cluster, ΔG^* is given by the expression:

$$\Delta G^* = \frac{16\pi}{3} \frac{\sigma_{sl}^3}{\Delta\mu^2}. \tag{2.15}$$

Here σ_{sl} is the interfacial free energy between the solid and liquid, and $\Delta\mu$ is the difference in chemical potentials between the cluster on the liquid state and the cluster with a solid embryo of critical size n^*.

The difference in chemical potentials is in turn given by

$$\Delta\mu = \Delta G_v + w', \tag{2.16}$$

where ΔG_v represents the free energy of freezing per unit volume, and w' the work per unit volume against the Laplace pressure, given by the expression:

$$w' = P_L(\rho_l - \rho_s)/\rho_l, \tag{2.17}$$

where the $\rho's$ are the densities of the liquid and solid phase in the clusters and $P_L = 2\sigma_l/r_0$ is the Laplace pressure [3] inside the cluster with surface tension σ_1 and radii r_0. Our goal now is to obtain an estimate for the chemical potential $\Delta\mu$ by means of thermodynamic approximations, and to use eqn. (2.15) with σ_{sl} as a fitting parameter. Our fit will be made against the data for the now known values for the free energy barriers ΔG^*.

The free energy of freezing ΔG_v may be derived from the Gibbs-Helmoltz rela-

---

[3] The Lapplace pressure describes the difference in pressure between the inside and the outside of a droplet due to the curvature of the interface

tion, which reads:

$$(\frac{\partial \Delta G}{\partial T})_p = -\frac{\Delta H}{T^2}. \tag{2.18}$$

The change in free energy from solid to liquid state can be obtained by integration of the change in enthalpies between the supercooled cluster with a solid embryo n^* and the cluster on the liquid state, hence we have $\Delta G = G_l(T) - G_s(T)$ and $\Delta H = H_l(T) - H_s(T)$.

A relation between enthalpy and temperature may be fit to a polynomial obtained from simulations. This is usually approximated well enough with a first or second order polynomial.

Entalphy is defined as H = U +P V . For an incompressible droplet and zero pressure the difference in enthalpies will be equal to the difference in internal energies. For the liquid branch in the caloric plot of the 456 atom cluster (see fig. 2.20), we obtained a second order polynomial fit for the internal energy U_l:

$$U_l(T) = a_0 + a_1 T + a_2 T^2 = -360.16 + 0.036878T - 0.20037 \times 10^{-5} T^2, \tag{2.19}$$

whereas for the solid branch, the internal energy as a function of temperature is given by:

$$U_s(T) = b_0 + b_1 T + b_2 T^2 = -356.66 + 0.025692T - 0.21146 \times 10^{-5} T^2, \tag{2.20}$$

The heat capacity at constant pressure is given by the relation $C_p = \frac{\partial H}{\partial T}$. From the derivatives of the internal energy functions (2.19), (2.20) we obtain

$$Cp_l(T) = b_1 + 2b_2 T = 0.036878 - 0.40074 \times 10^{-5} T, \tag{2.21}$$

$$Cp_s(T) = a_1 + 2a_2 T = 0.025692 - 0.42292 \times 10^{-5} T. \tag{2.22}$$

Now, the difference in entalphies is obtained with respect to a reference state T_m.

$$\begin{aligned} \Delta H &= H_l(T) - H_s(T) \\ &= \left[\int_{T_m}^{T} Cp_l(T')dT' + H_l(T_m)\right] - \left[\int_{T_m}^{T} Cp_s(T')dT' + H_s(T_m)\right] \end{aligned} \tag{2.23}$$

The difference in heat capacities under the integrals in eqn. (2.23) is positive for any temperature and is given in units of $kJmol^{-1}K^{-1}$. This difference may in turn be expressed as another polynomial,

$$\Delta C_p = Cp_l(T) - Cp_s(T) = c_1 + c_2 T. \tag{2.24}$$

Further, we define the reference state T_m as the melting point of the system and hence we may define the entalphy of fusion as the change in entalphies between the solid and the liquid phase at the melting point: $\Delta H_{fus} = H_l(T_m) - H_s(T_m)$. We may thus recast eqn. (2.23) as:

$$\Delta H = H_l(T) - H_s(T) = \int_{T_m}^{T} \Delta C_p(T')dT' + \Delta H_{fus}, \tag{2.25}$$

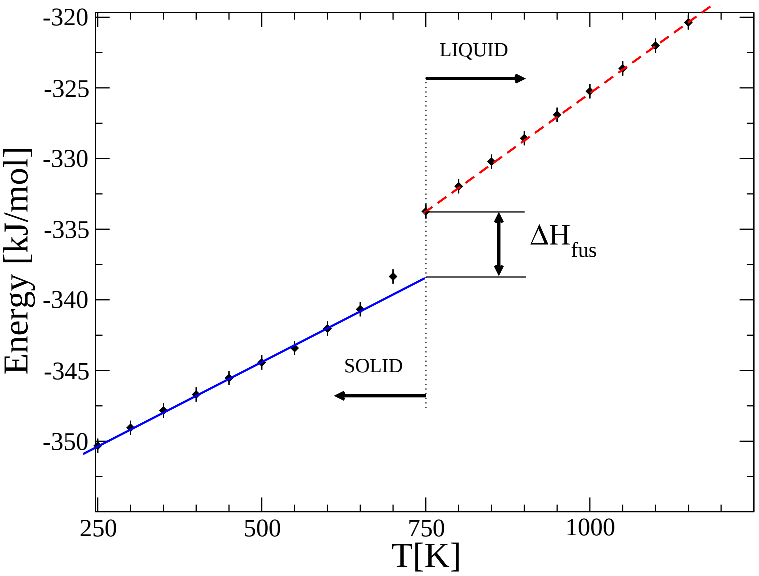


Figure 2.20: Entalphy of fusion ΔH_{fus}, 456 atom cluster

One would expect to have a sharp phase transition only in the thermodynamic limit. In the case of finite size systems the phase transition occurs in a transition zone and hence the melting point can not be uniquely determined [58].

Now, by substituting expression (2.24) into expression (2.25) we get

$$\Delta H = H_l(T) - H_s(T) = c_1(T - T_m) + \frac{1}{2}c_2(T - T_m)^2 + \Delta H_{fus}. \tag{2.26}$$

Integration of the expression (2.26) according to the Gibbs-Helmoltz relation (2.18).

$$\Delta G = -T\int_{T_m}^{T} \frac{c_1(T' - T_m) + \frac{1}{2}c_2(T' - T_m)^2 + \Delta H_{fus}}{T'^2} dt. \tag{2.27}$$

Substitution by a reduced temperature $T^* = T/T_m$, yields:

$$\begin{aligned}\Delta G = c_1 T_m\{1 - T^* + T^*\ln(T^*)\}\\ -c_2T_m^2\{1 - T^{*2} + 2T^*\ln(T^*)\}\\ -\Delta H_{fus}[1 - T^*]\end{aligned} \tag{2.28}$$

Quantity	Symbol	EMA potential	exp.
liquid	$\rho_l[kg/m^3]$	17 280	17 280
solid	$\rho_s[kg/m^3]$	19 000	18 400
surface tension			
liquid	$\sigma_l[J/m^2]$	0.74	1.13
	$-d\sigma/dT[mJ/m^2K]$		0.14
solid	$\sigma_s[J/m^2]$	0.90	1.40
	$-d\sigma/dT[mJ/m^2K]$		0.14
solid-liquid	$\sigma_{sl}[J/m^2]$	0.13-0.15	0.27

Table 2.3: EMA potential predictions for bulk gold vs. Experimental values. Obtained from ref. [21]

We set the value of the melting temperature as $T_m = 750°K$ with constants $c_1 = 11.186 Jmol^{-1}K^{-1}$, $c_2 = 2.218\times10^{-4} Jmol^{-1}K^{-2}$ and $\Delta H_{fus} = 4951.881 Jmol^{-1}$. A plot of expression (2.27) is shown on fig. (2.21 left). This function is zero at $T \approx 151°K$ and $T = 750°K$ with a global minimum at $T \approx 414°K$. It should be stressed out that for temperatures below the global minimum, this approximation produces two temperatures for the same chemical potential, which is unrealistic, however, the approximation is valid for temperatures near the melting point of the system, i.e. $T \approx 750°K$.

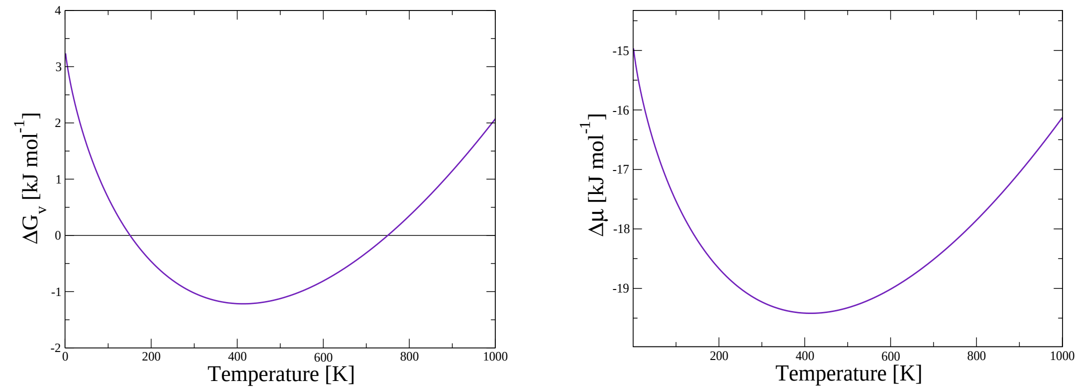


Figure 2.21: Left: Free energy of embryo freezing. Right: Chemical potential, 456 atom cluster.

To obtain the free energy of freezing per unit volume we divide by the volume of the cluster and multiply by the number of moles in the system (i.e. $\Delta G_v = \frac{n}{AV}\Delta G$), with n the number of atoms in the cluster and A Avogadro's number, and the volume of the system in the liquid phase with a value $V = 7.0 \times 10^{-27} m^3$.

Taking the volume V we approximate a radii for a spherically averaged cluster to be $r_0 = 11.8668 \times 10^{-10} m$ and the surface tension from the reference [21] in which the EMA potential used in the present work has been previously used to predict bulk quantities [21] (cf. table 2.4)

With the values in table (2.3) we obtain a value for the Laplace pressure $P_L = 11.8668 \times 10^{-10}(1.9786 \times 10^{10}[J/m^3])$ and the work per unit volume as defined by eqn. (2.17) is $w' = -0.1969 \times 10^{10} J/m^3(-18.2024509[kJ/mol])$. This work term is added to equation (2.28) to get a difference in chemical potentials:

$$\begin{aligned} \Delta\mu &= c_1 T_m\{1 - T^* + T^* \ln(T^*)\} \\ &\quad -c_2 T_m^2\{1 - T^{*2} + 2T^* \ln(T^*)\} - \Delta H_{fus}[1 - T^*] + w'. \end{aligned} \tag{2.29}$$

Table (2.4) shows the MC predicted values for the free energy barrier (third column) along with temperature and size of the critical embryo (first and second columns respectively).

$T[^\circ K]$	$n^*[atoms]$	$\Delta G(n^*)/kT[]$	$\sigma_{sl}[J/m^2]$ (fitted values)
650	23	10.0	0.1247
660	30	10.2	0.1237
670	37	10.4	0.1224
680	42	10.5	0.1205
690	47	10.7	0.1190
710	58	11.6	0.1166
730	70	13.0	0.1146

Table 2.4: Table with calculated solid/liquid surface free energy densities using the MD approach. The first three rows correspond to calculated MC simulations data.

Each one of the σ_{sl} values in the fourth column of table (2.4) was calculated by fitting the function

$$\Delta G^* = \frac{16\pi}{3}\frac{\sigma_{sl}^3}{\Delta\mu^2}, \tag{2.30}$$

to the Monte Carlo values listed on table (2.4) and by using σ_{sl} as a fitting parameter for each temperature set, and assuming the temperature dependent chemical from expression (2.29). The optimal value of the fits for each temperature is shown in fig. (2.22 left).

The values of the free energy density σ_{sl} as a function of temperature are listed in fig. ((2.22), right), with a linear dependent temperature relationship $\sigma_{sl}(T) = -0.0001315 \times T + 0.2102$. The correlation coefficient for this relationship is $r = 0.9958828$.

A plot comparing the MC values with those of the CNT like model with temperature dependent surface density σ_{sl} is shown on fig. (2.23).

The quadratic fit shown in fig. (2.23) has a temperature dependent functional form $\Delta G^*/k_BT = 203.04 - 0.59275 \times T + 0.00045526 \times T^2$, and predicts a minimum nucleation barrier for $T = 651^\circ K$ with the size of the barrier equal to $10.1 \pm 0.5 k_BT$, and the CNT like model from eqn. (2.30) in dimensionless units is given by the expression:

$$\frac{\Delta G^*}{k_B T} = 1.037130 \times 10^7 \frac{\sigma_{sl}^3}{T \Delta\mu^2}, \tag{2.31}$$

where σ_{sl} and $\Delta\mu$ have the temperature dependence given above. This model predicts a minimum for the nucleation barrier in the order of $9.799 k_B T$, for a temperature equal to $615° K$.

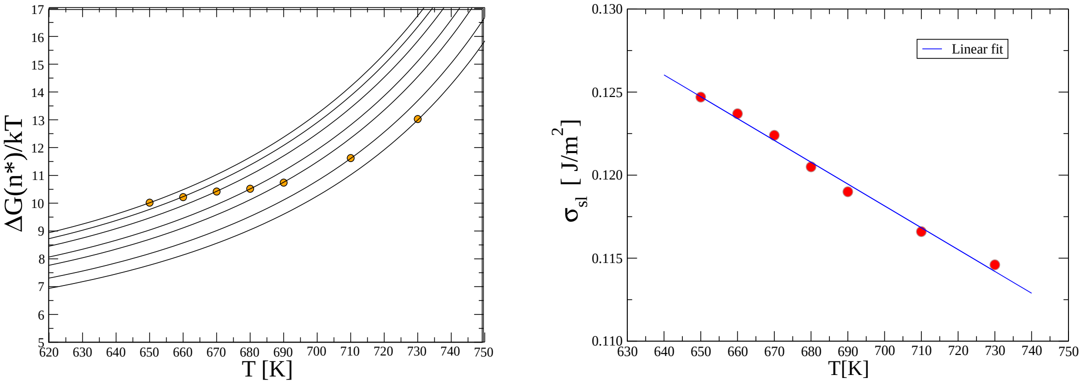


Figure 2.22: Left: calculated temperature dependent solid/liquid surface tension. Right: fits to the $\Delta F(n^*)/kT$ data. 456 atom cluster.

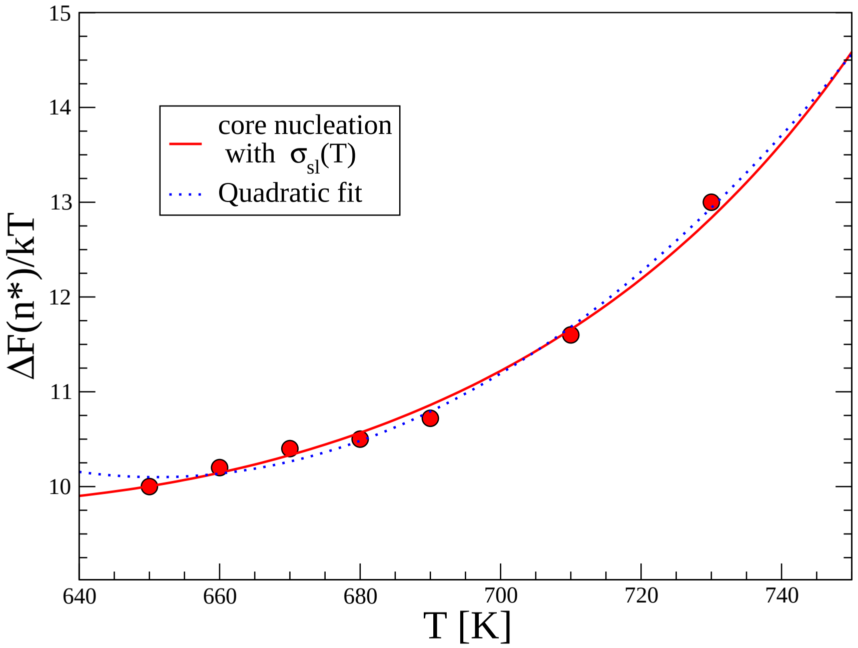


Figure 2.23: CNT fits to the $\Delta F(n^*)/kT$ data. The inset shows the temperature dependence of the free energy barrier for small temperatures under the temperature dependent surface tension approximation. 456 atom cluster.

2.7 Surface Nucleation

The usual model for classical nucleation in a cluster assumes that the embryo of the solid phase grows in the core of the nanoparticle. In terms of surface wetting phenomena, this assumes that the solid is completely wet by its liquid. We can test this assumption by examining configurations obtained from our MC simulations described in Section (2.5). Figs. (2.24) and (2.25) show one such configuration in different orientations as identified by the axis indicators. Fig. (2.24) shows just those atoms belonging to the largest embryo in the cluster while fig. (2.25) shows how the embryo is embedded in the cluster. Clearly, this embryo sits on the surface of the cluster suggesting that the solid phase may only be partially wet by its liquid. This contradicts the assumptions made in the previous section (see section 2.6) in which the CNT assumption of nucleation at the core of the cluster and a temperature dependent solid-liquid free energy tension σ_{sl} was proposed as a way to match a phenomenological model with our simulations.

In Section (2.7.1) we carry out a quantitative analysis of the wetting behaviour in the nucleation of clusters, and in Section (2.7.2) we propose an alternative phenomenological model that accounts for surface nucleation in a cluster, and with which we are able to explain the free energy features of our detailed molecular simulations.

2.7.1 Simulation Results

A simple measure of the degree of wetting can be obtained by counting the number of atoms in embryo of a given size that are on the surface of the cluster. If the embryo is completely wet there should be no embryo atoms (or very few) at the surface. We collect configurations every 1×10^3 trajectories from our MC simulations described in Section (2.5). This ensures that each configuration is independent from the previous one. As a result, we have a total of 3000 configurations at each temperature. We measure the number of surface atoms, $n_{max,surface}$, contained within the largest n_{max}-*sized* embryo in the cluster. This gives us between 20 and

50 configurations for each embryo size n_{max} at each temperature. Fig. (2.26) shows that $n_{max,surface}$ vs n_{max} can be fit to a straight line for all temperatures, for the data between n_{max} = 15 and 80. Table (2.6) shows that fit parameters and error estimates.

T[K]	$n_{nmax,surf} = m \times n_{max} + b$		Correlation coefficient.	slope std. error
	m	b		
750	0.47176	1.7639	0.9734	0.012898
730	0.43613	2.7985	0.9605	0.014885
710	0.46634	2.2349	0.9733	0.012533
690	0.50749	0.8902	0.9620	0.017092
680	0.50175	1.7298	0.9714	0.015187
670	0.43789	4.3382	0.9282	0.022110
660	0.47946	3.0887	0.9540	0.018543
650	0.47509	3.4498	0.9651	0.016623

Table 2.5: Linear fit for the number of atoms in the surface vs. total number of atoms for the n_{max} embryo.

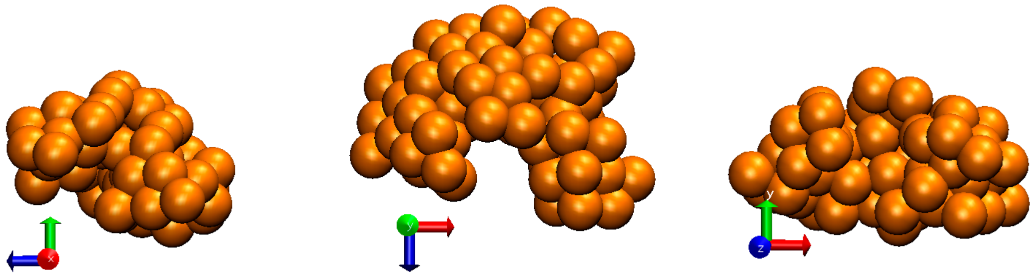


Figure 2.24: Snapshot of an n_{max} embryo. Left, topview from the X axis, center a topview from the Y axis, and right hand side, top view from the Z axis.

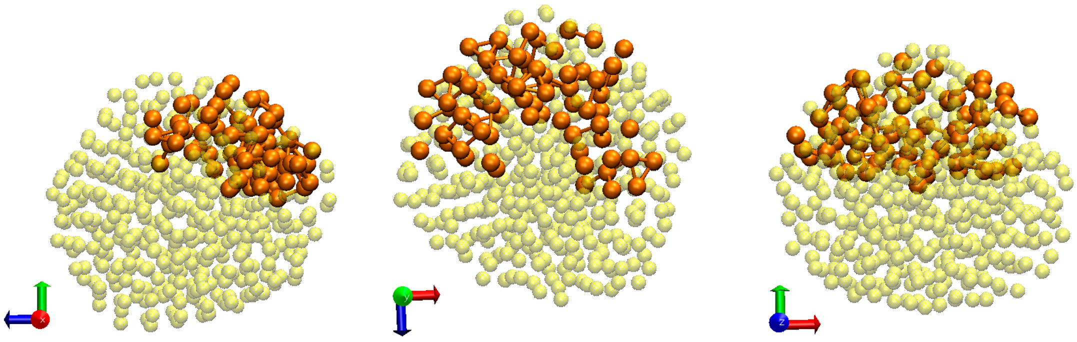


Figure 2.25: Snapshot of a cluster with the n_{max} embryo in the surface. Left, topview from the X axis, center a topview from the Y axis, and right hand side ,a top view from the Z axis.

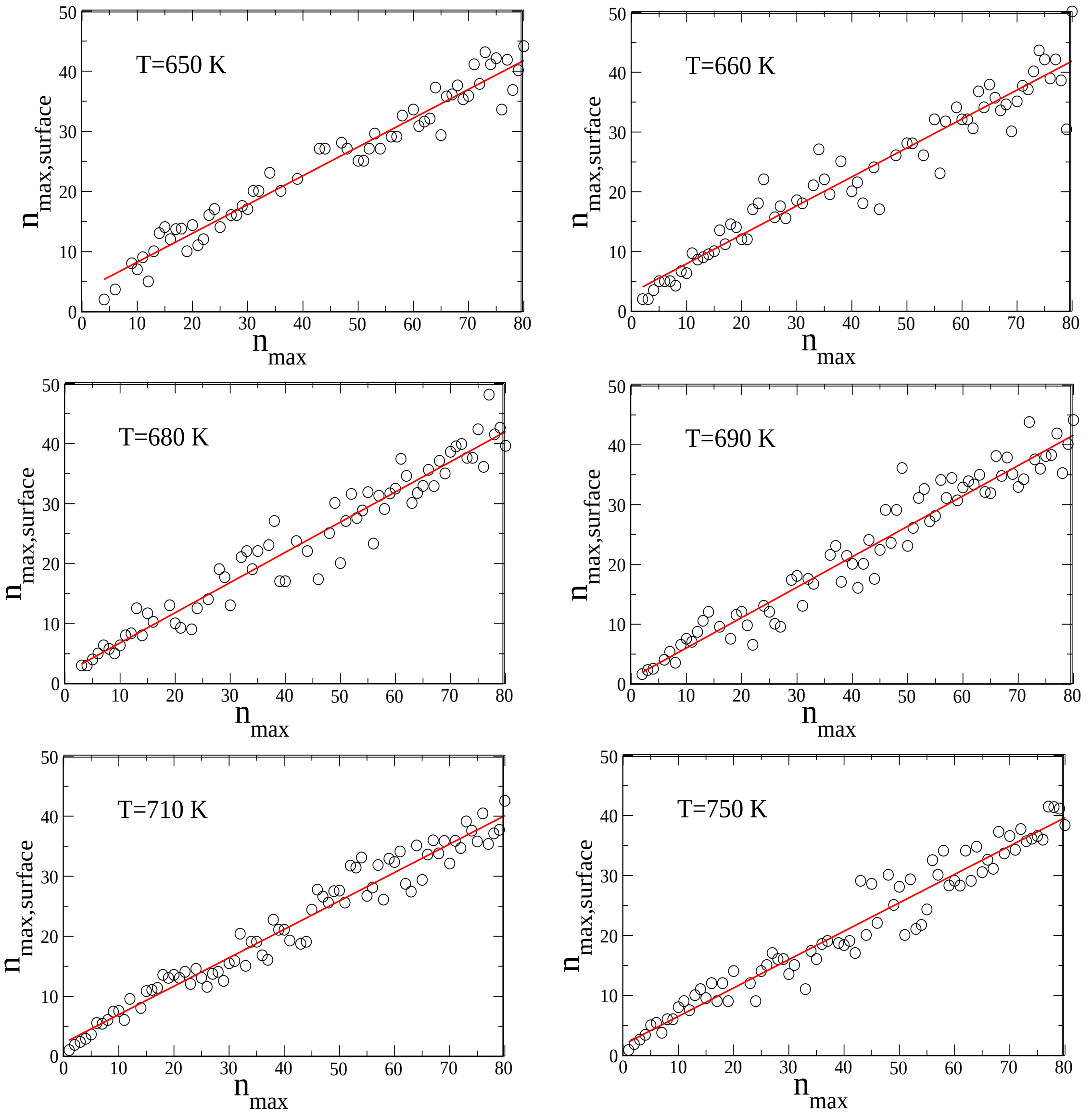


Figure 2.26: Fraction of atoms in the surface of the cluster $n_{max,surface}$ that belong to the largest embryo size n_{max} for various temperatures. The linear fits are taken over a distribution from about 100 sample points for every temperature to show the scatter of the data. 456 atom cluster.

Given the scatter of the data, the fraction of embryo atoms in the surface (obtained from the slope) appears to be independent of temperature (see figs. 2.26). To obtain better statistics, we averaged over all temperatures. The results, shown in fig. (2.27), suggests that smaller embryos have a higher fraction of atoms in the surface than larger embryos. Both large and small embryo regions can be fit by linear curves (see table 2.5). The data collected for the simulations at each temperature has been overlapped on fig. (2.27) (shown with plus symbols), the averaged data appears in filled circles, while the dashed line, both in the main fig. and the inset corresponds to the linear fit for small embryos. The data for the fits for both large and small embryos is shown in table (2.6).

Our results clearly indicate that a significant fraction of atom in the embryo, 47% for large embryos and 64% for small embryos, appear in the surface of the cluster. This supports the notion that in gold nanoparticles, nucleation occurs in the surface and that the solid is only partially wet by its liquid.

Region	$n_{nmax,surf} = m \times n_{max} + b$ m	b	Correlation coefficient.	slope std. error
$n < 20$	0.63519	0.66038	0.9939599	0.0058186
$n \geq 20$	0.46842	2.42550	0.9892535	0.0227692

Table 2.6: Linear fit for the number of atoms in the surface vs. total number of atoms for the n_{max} embryo.

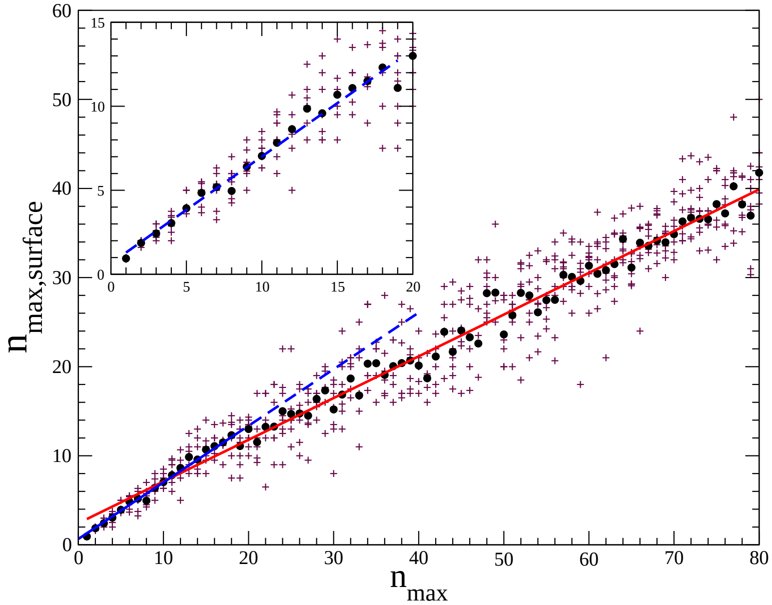


Figure 2.27: Number of atoms in the surface belonging to the n_{max} embryo vs n_{max} embryo size for the 456 atom cluster.

2.7.2 Lens Model

The condition of partial wetting places thermodynamic limits on the surface free energy densities σ_{ij} of the three phases such that $\sigma_{sv} - \sigma_{lv} < \sigma_{sl}$, where the subscripts s,l and v denote the solid, liquid and vapor phases respectively. For the EAM potential, $\sigma_{sv} = 0.90 J/m^2$ and $\sigma_{lv} = 0.74 J/m^2$ [59] which requires $\sigma_{sl} > 0.16 J/m^2$. *Bartell et al.* [21] found that a number of thermodynamic theories and empirical relations gave estimates of σ_{sl} within the range of 0.11 to $0.16 J/m^2$. In the same work, the authors used a CNT model that assumes complete wetting of the solid embryo by the liquid (core nucleation) to predict the solid-liquid surface tension based on fitting the rate of nucleation obtained from molecular dynamics simulations. For a cluster of N = 459, at $T = 700^\circ K$, they found $\sigma_{sl} = 0.084 J/m^2$ which is well below the wetting threshold.

To obtain an estimate of σ_{sl} under the conditions of partial wetting, we assume the solid embryo grows at a planar liquid-vapour interface in the shape of a lens (see fig. 2.28). The free energy required to form an embryo containing n atoms, from the liquid phase, can be expressed by:

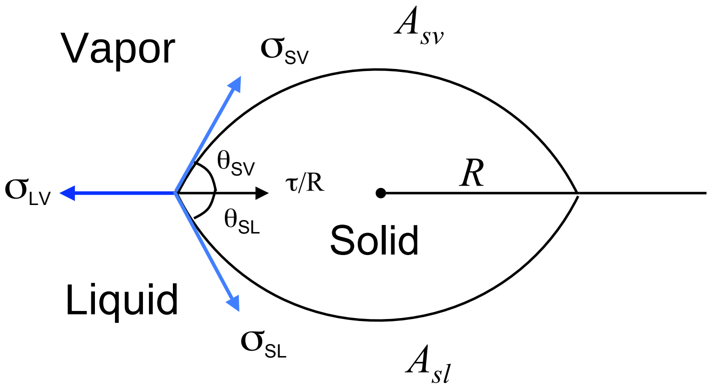


Figure 2.28: Schematic cross section of a solid lens nucleus forming at the liquid-vapour interface. A_{sl} and A_{sv} are the solid-liquid and solid-vapour interfacial surface areas respectively and R is the radius of the lens. The three arrows originating from the 3-phase contact are the force vectors of the surface tensions σ_{ij}.

$$\Delta G = n\Delta\mu + \sigma_{sl} A_{sl} + \sigma_{sv} A_{sv} - \sigma_{lv} A_{lv} \ , \tag{2.32}$$

where $\Delta\mu$ is the difference in chemical potential between the metastable liquid phase and the stable solid, σ_{ij} is the surface tension between phase i and phase j and A_{ij} is the respective surface area. The minus sign in front of the liquid-vapour surface free energy term accounts for the surface area lost due to the appearance of the embryo. Under the condition of mechanical equilibrium, the contact angles θ_{sl} and θ_{sv} can be obtained from Neumann's triangle construction[60] as

$$\cos\theta_{sl} = \frac{\sigma_{lv}^2 + \sigma_{sl}^2 - \sigma_{sv}^2}{2\sigma_{sl}\sigma_{lv}}, \qquad \cos\theta_{sv} = \frac{\sigma_{lv}^2 + \sigma_{sv}^2 - \sigma_{sl}^2}{2\sigma_{sv}\sigma_{lv}}. \tag{2.33}$$

The volume of the solid embryo containing n atoms is

$$nv = \frac{\pi}{3} R^3 A(\theta_{sl}, \theta_{sv}), \tag{2.34}$$

where R is the radius of the lens, v is the volume per molecule in the crystal and

$$A(\theta_{sv}, \theta_{sv}) = \frac{\sin\theta_{sl}(2+\cos\theta_{sl})}{(1+\cos\theta_{sl})^2} + \frac{\sin\theta_{sv}(2+\cos\theta_{sv})}{(1+\cos\theta_{sv})^2}\ . \tag{2.35}$$

The interfacial surface areas are given by

$$A_{sl} = \frac{2\pi R^2}{1+\cos\theta_{sl}}, \qquad A_{sv} = \frac{2\pi R^2}{1+\cos\theta_{sv}}\ , \tag{2.36}$$

where

$$R = n^{1/3}\left[\frac{3v}{\pi A(\theta_{sl}, \theta_{sv})}\right]^{1/3} \tag{2.37}$$

is obtained by solving Eqs. (2.34) and (2.35). The lens model was originally introduced to study droplet formation of the liquid vapor interface without the line tension [61]. A fit without a line tension contribution only provides a fit similar to that of CNT, and only if we significantly reduce the difference between σ_{sv} and σ_{lv}.

To fit the shape of the free energy curves calculated in our simulations we need to account for both the line tension τ and its curvature correction, τ_0. *Auer and Frenkel* [27], also found that these terms played an important role in the heterogeneous freezing of hard sphere colloids at a wall. By including the line tension term in the mechanical equilibrium analysis at the three phase contact line we obtain size dependent contact angles such that:

$$\cos\theta_{sl} = \frac{\sigma_{lv}^2 + \sigma_{sl}^2 - \sigma_{sv}^2 - 2\sigma_{lv}\tau/R + (\tau/R)^2}{2\sigma_{sl}(\sigma_{lv} - \tau/R)}, \tag{2.38}$$

and:

$$\cos\theta_{sv} = \frac{\sigma_{lv}^2 - \sigma_{sl}^2 + \sigma_{sv}^2 - 2\sigma_{lv}\tau/R + (\tau/R)^2}{2\sigma_{sv}(\sigma_{lv} - \tau/R)}. \tag{2.39}$$

The free energy may be cast as

$$\Delta G(n) = n\Delta\mu + R^2\left[\frac{2\pi\sigma_{sl}}{1+cos\theta_{sl}} + \frac{2\pi\sigma_{sv}}{1+cos\theta_{sv}} - \pi\sigma_{lv}\right] + 2\pi R\left[\tau + \frac{\tau_0}{R}\right]. \tag{2.40}$$

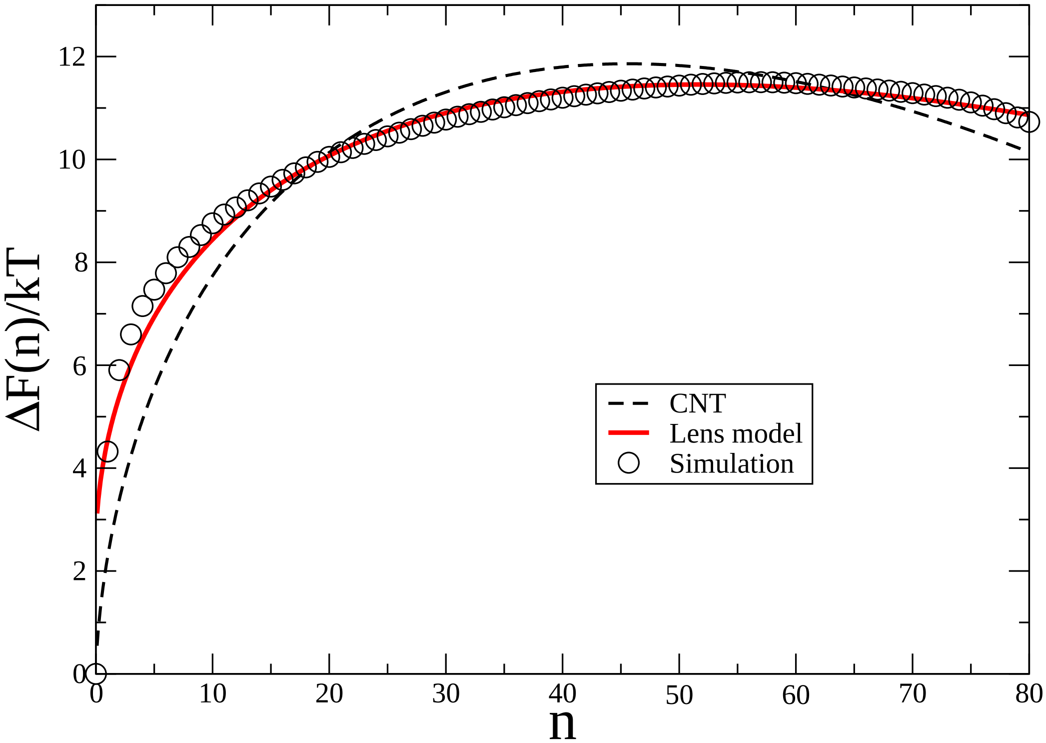


Figure 2.29: Phenomenological model data fits to the calculated free energies for the 456 atom cluster at $T = 710°K$. The lens model with the line tension included (solid line) and the CNT model (dashed line)

By fitting function (2.40) to our data at $710°K$, and using $\Delta\mu$, σ_{sl}, τ and τ_0

as fitting parameters, we obtained: $\sigma_{lv} = 0.18J/m^2$, $\tau = -1.17 \times 10^{-11}J/m$ and $\tau_0 = 3.92 \times 10^{-21}J$. Due to the non linear nature of the numerical fit, we cant guarantee that the values for these three parameters are those at the global minimum for the data fit and there is nothing in the equations that excluding the possibility of finding a positive line tension. Nevertheless, a negative value of this parameter would help to explain why smaller embryos have a greater fraction of atoms in the cluster surface. A negative line tension acts to expand a lens of a given volume or number of particles. This negative line tension thus, flattens the lens and pulls more particles into the surface of the solid embryo. The inset of fig. (2.27), relates the number of atoms in the surface of the largest solid embryo. At larger embryo sizes, the line tension becomes less dominant and the surface terms will act to make the lens more spherical-like reduces the number of surface atoms.

For matters of comparison, we also fit the CNT core nucleation model to our data at the same temperature of $710°K$ using the chemical potential μ and σ_{sl} as adjustable parameters and assuming a spherical geometry for the embryo. The resulting $\sigma_{sl} = 0.085J/m^2$ is the same as that obtained from direct measurements of the rate [21]. We see in Fig (2.29) that the CNT model clearly fails to predict the correct shape of the barrier but does obtain a close estimate of the barrier height. In contrast the lens model captures the essential features of the free energy plots.

2.8 Summary

We have used computer simulation techniques to calculate the nucleation free energy barrier to freezing in gold nanoparticles and show that the solid embryos form at the liquid-vapour interface. While we use a partial wetting model at a bulk interface to fit our data, adjusting the model for a realistic nanoparticle model that includes the curved liquid-vapour is only likely to change the quantitative values and the shape of the free energy curve for small embryos is still going to be dominated by the line tension. The present thesis work represents the first direct calculation of a free energy barrier for the nucleation process in a nanoparticle.

Chapter III.

Thermodynamic Stability of the Liquid Phase

In this chapter we address the question: *Is there a limit of stability to the liquid phase of a cluster?* This is a fundamental question that goes to the very heart of our understanding of what a liquid is, and has important implications for understanding freezing mechanisms and aspects of the glass transition in deeply supercooled liquids. The spinodal of a phase is usually described in terms of the stability criteria outlined in Section 1.1.2 and Appendix A-I, but in terms of nucleation the limit of stability is defined as the point at which the nucleation barrier goes to zero. In Chapter II, we saw that the nucleation barrier remained finite at all temperatures, which is consistent with the predictions of classical nucleation theory and suggests there is no spinodal. However, we also note that the simulation techniques used to obtain the free energy barrier for nucleation are really only appropriate for systems that are only mildly metastable where the appearance of an embryo is rare. To overcome this we introduce a free energy work function to describe the stability of the cluster in section (3.1). In subsection ((3.2)) we use molecular dynamics simulations to search for a kinetic signature of the limit of stability.

3.1 Free energy barrier from the largest embryo distribution.

We will use the size of the largest embryo in the cluster, n_{max} as our Landau type order parameter to define the state of the cluster. Clearly, if $n_{max} = 0$, the cluster must consist of only liquid-like atoms while at the other extreme, if $n_{max} = N$, then the cluster must be completely frozen. The advantage of using this order parameter is that the state of the cluster is uniquely defined by the largest embryo and hence we can rigourously define a free energy based on the probability $P(n_{max})$, of finding the cluster with the largest embryo n_{max}, given by the relation [26]

$$\Delta F(n_{max})/kT = -\ln P(n_{max}) \ . \tag{3.1}$$

In particular, we do not have to rely on arguments involving the rarity of the embryo in order to write eqn. (3.1) as was required in the case of N_n and eqn. (1.25). If we set $F(n_{max})/kT = 0$ at $n_{max} = 0$ then $\Delta F(n_{max})/kT$ is the work required to take the cluster to a state where the largest embryo is of size n_{max}.

An important distinction between $F(n_{max})$ and the work required to form an embryo of size n, $F(n)$, is that $F(n_{max})$ is the work required to change the state of the entire cluster, rather than a work associated with the formation of a particular sized embryo within the cluster.

We can in fact relate $P(n_{max})$ and N_n using the rarity of clusters as follows [26]: Let P_n be the probability that there is at least one cluster of size n in the system, and $P_n(i)$ be the probability that there are exactly i clusters of size n. Then,

$$\begin{aligned} P_n &= P_n(1) + P_n(2) + P_n(3) \ldots & (3.2) \\ N(n) &= P_n(1) + 2P_n(2) + 3P_n(3) \ldots . & (3.3) \end{aligned}$$

A rare cluster of size r is defined such that that $P_r(1)$ is small, and additionally, the appearance of a rare cluster is independent of prescense of other embryos in the clusters, i.e. $P_r(2) \approx P_r(1) \times P_r(1) \approx 0$. As a result, $P_n = N(n)$. [25, 41]. Also, two rare clusters of different sizes appearing at the same time, occurs with a vanishing probability $P_{r+m}(1) \times P_r(1) \approx 0$ [assuming $P_{r+m}(1) < P_r$ for $m > 0$, i.e. larger clusters are rarer], and so a rare cluster will also be the largest cluster in the system. Finally, we obtain $P_n(1) = P_n = N(n) = P(n_{\mathrm{max}})$, for $n \geq r$ (the equality holds up to a normalization constant that is irrelevant in determining the free energy).

To calculate $P(n_{max})$ we use the same simulation techniques, conditions and simulation run lengths used in Section (2.5) to calculate the nucleation free energy barriers. Using eqn. (1.35) with the biasing potential given by eqn. (2.8) we have

$$\langle P(n_{\mathrm{max}})\rangle = \frac{\langle P(n_{\mathrm{max}})\rangle_C \exp\left[\phi(n_{\mathrm{max}})/k_{\mathrm{B}}T\right]}{\langle \exp\left[\phi(n_{\mathrm{max}})/k_{\mathrm{B}}T\right]\rangle_C} \ . \tag{3.4}$$

In fig. (3.1 left) , we show the histograms in the constrained space, for a simulation equilibrated at $750°K$. Each fragment of the piecewise distribution belongs to a node with umbrella centers $n_{0,\nu} = \{0, 10, 20, 30, 40, 50, 60, 70\}$, part of a parallel tempering plus umbrella sampling scheme comprising eight different temperatures. $T_\mu[°K] = \{650, 660, 670, 680, 690, 710, 730, 750\}$. Fig. (3.1 right), is the result of applying eqn. (3.4) to the respective series of constrained embryo distributions.

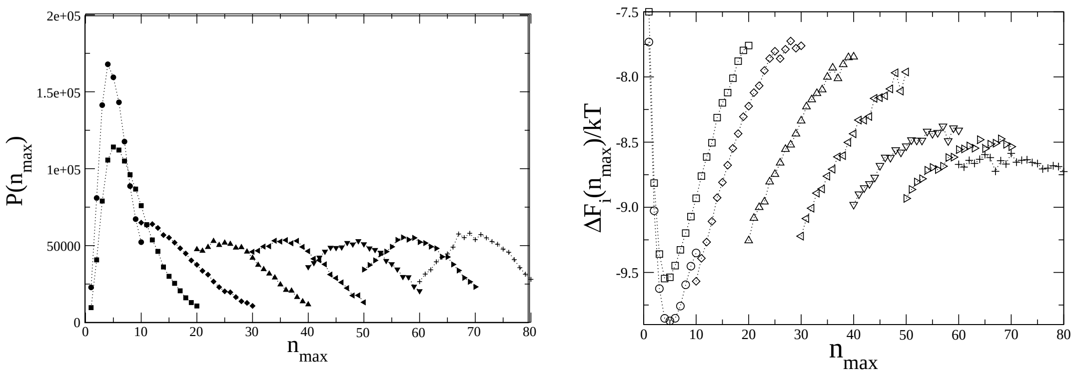


Figure 3.1: Left: Biased piecewise histogram with respect to the n_{max} and umbrella centers. Right: Piecewise construction of the free energy from the histograms in the left. The system is a 456 atom cluster at a temperature $T = 750°K$

The alignment of free energies in each sampling window was made according to eqn. (2.14), using a 20 order polynomial for the function $\Gamma(n) = \sum_{k=1}^{k_{max}} a_k n^k + c$, however excellent fits can be obtained by simply adjusting the "c" constants for every node. Figs. (3.3) compare the free energies calculated using eqn. (3.1) and eqn. (2.12). We can see that at high temperatures, where the system is only mildly supercooled, the two free energies are equivalent over a significant portion of the curves as we expected, and where the two free energies only differ for very small embryo sizes. However, at $T = 660°K$ the curves are very different.

In fig. (3.2) we show $\Delta F(n_{max})/kT$ over a range of temperatures. At high temperatures, we see an initial decrease in the free energy to a minimum at a small n_{max} indicating that the supercooled liquid cluster only contains a small solid-like embryo. Beyond the minimum, it requires work to increase the size of the largest solid embryo in the cluster which suggests that the cluster is in fact metastable. That is, small fluctuations on n_{max} around the minimum, are locally stable. However, larger scale fluctuations, that take the system beyond the critical embryo size cause the cluster to freeze since the cluster can lower its free energy by increasing the size of the largest embryo and moving towards the crystal state.

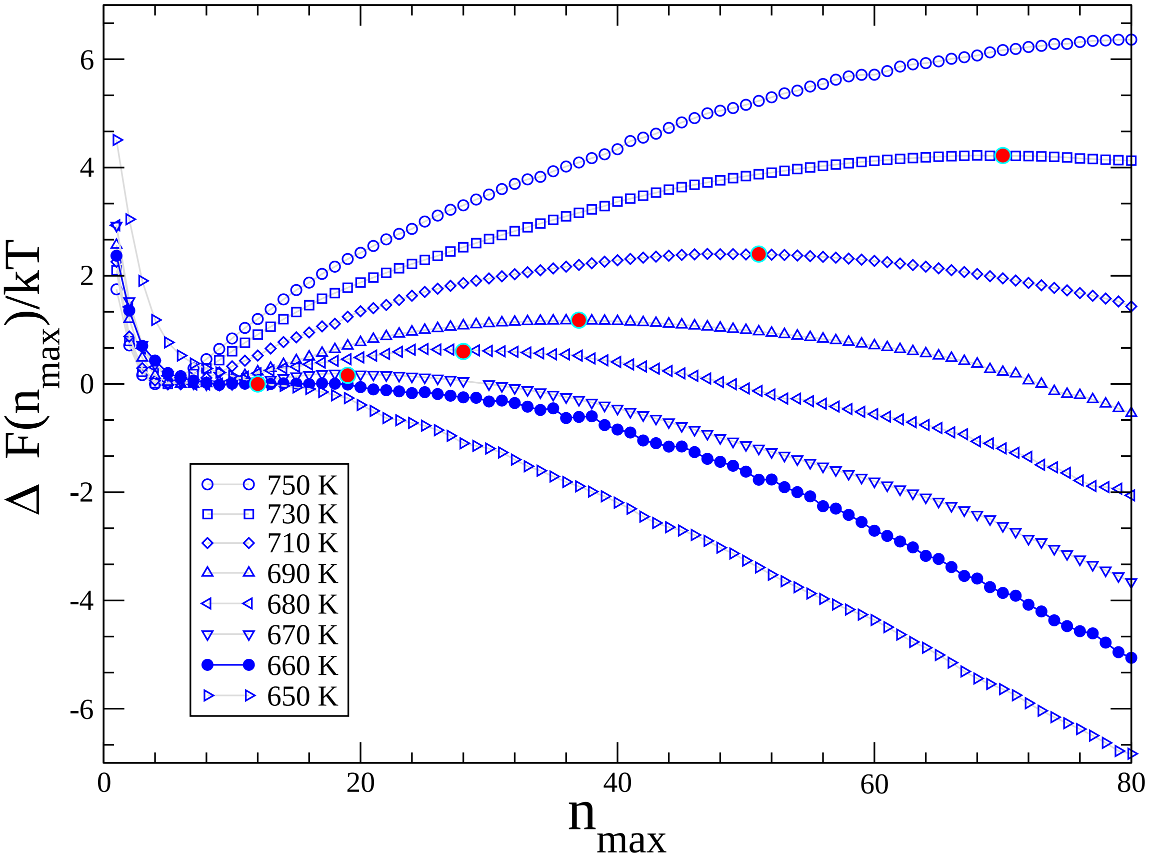


Figure 3.2: Free Energy barriers for various temperatures, for the 456 atom cluster. The spinodal temperature is shown with dark full circles, the light dark circles on each free energy curve signal the position of the critical largest embryo size n^*_{max}.

As the temperature is lowered the free energy barrier separating the metastable liquid state and the crystal becomes lower and the maximum moves towards smaller values of n_{max}. Eventually, we reach a temperature below which there is no barrier. At this temperature any fluctuation in the cluster that increases the size of the largest embryo decreases the free energy. This represents the limit of stability of the liquid phase of the cluster.

To clearly identify the spinodal point we define a free energy barrier $\Delta F(n^*_{max})$ as the difference in free energy between the maximum in the free energy occurring at n^*_{max} and the free energy at the minimum as described in fig. (3.4 top left). In table (3.1), we show that the barrier goes to zero at $T = 660°K$ while the critical size of the largest cluster at the spinodal remains finite as the spinodal temperature is approached (also see fig. 3.4 right). Fig. (3.3 right), compares $\Delta F(n_{max})/kT$ with $\Delta F(n)/kT$, obtained in Section (2.5) at the spinodal temperature $T = 660°K$. Clearly there appears to be a significant barrier to embryo formation as calculated from the equilibrium distribution at the point our present work suggests the liquid phase is unstable. We should stress that as previously described, these two free energies are fundamentally different, except in the regime of rare embryos.

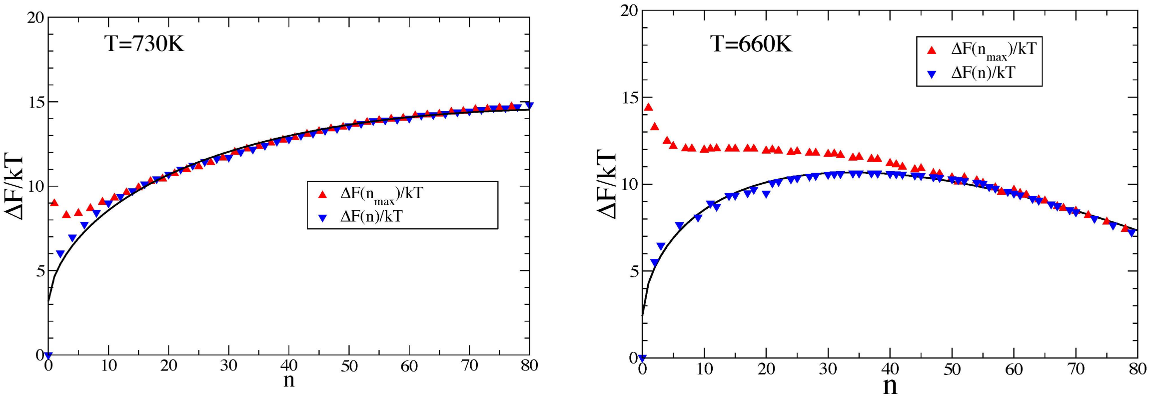


Figure 3.3: Comparison of free energy barriers for nucleation with respect to the n_{max} and embryo distribution n with respect to two different temperatures. The relative position of the $\Delta F(n_{max})/kT$ curves has been adjusted to maximize overlap with respect to the $\Delta F(n)/kT$ plots.

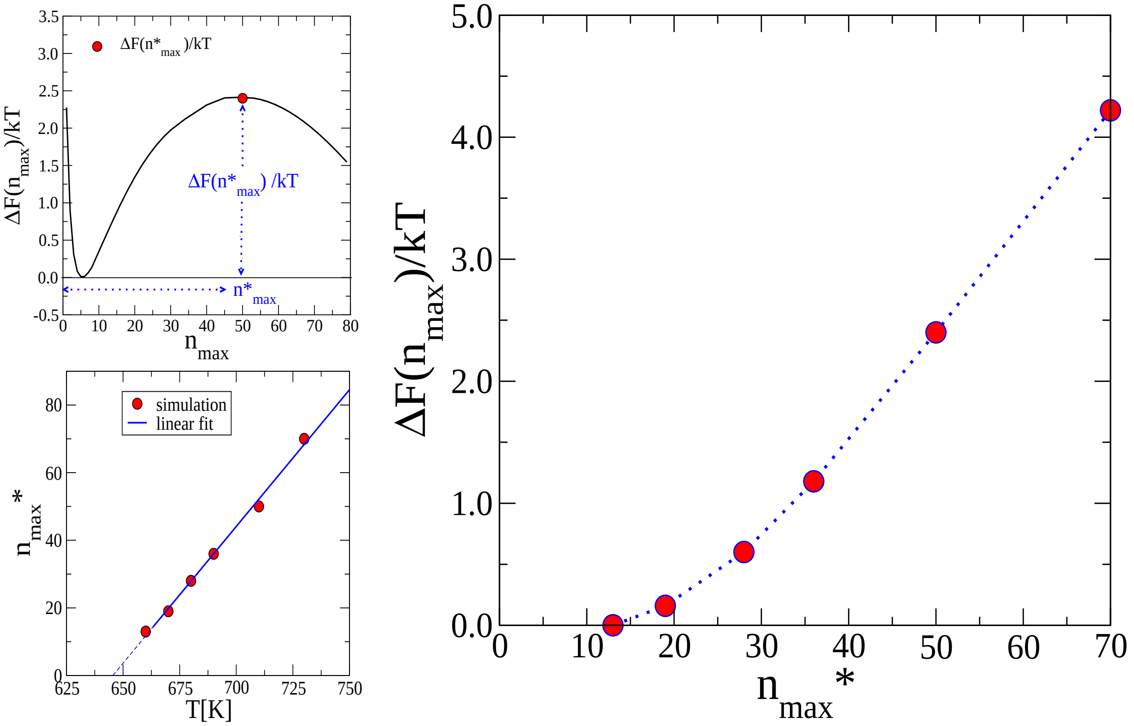


Figure 3.4: Top left. Barrier heights are computer with respect to the minimum in the curve, for small embryo sizes. Left, down. Linear like relationship for the embryo size as a function of temperature. Right. Barrier height as function of the largest embryo size, n_{max}. for the 456 atom cluster.

$T[^\circ K]$	$n^*_{max}[atoms]$	$\Delta F(n^*_{max})/kT[]$
650	-	-
660	12	0.0
670	19	0.16
680	28	0.60
690	36	1.18
710	50	2.40
730	70	4.22
750	-	-

Table 3.1: Free energy barriers to nucleation with respect to the critical largest embryo size n^*_{max}. The first to this data are shown in fig. (3.4).

3.2 Molecular dynamics calculation of nucleation rate.

The objective of the present section is to provide additional support to the idea that there is a limit of stability to the liquid phase of cluster. To this effect, we employed molecular dynamics simulations to calculate the nucleation rate in a fashion that is independent of the free energy barrier calculations carried out in Sections (2.5) and (3.1).

If a liquid cluster, equilibrated at a temperature above the freezing temperature is instantaneously quenched to a temperature below freezing, then the system will eventually nucleate to form the solid. We can calculate the nucleation rate by considering an ensemble of such runs. Assuming this process is described by a first-order rate law, the rate of nucleation J can be obtained from the relation [22, 63]

$$\ln\left[R_{\mathrm{x}}(t)\right] = -JV_c\left(t - t_0\right) \ , \tag{3.5}$$

where $R_{\mathrm{x}}(t)$ is the fraction of un-nucleated systems at time t, V_c is the volume of the system and t_0 is the lag time, i.e. the time required to achieve a steady state of precritical nuclei.

To make use of eqn. (3.5), we need to identify when a cluster has nucleated. While a number of different criteria have been used [26], to be consistent with our barrier calculation work we will follow the size of the largest embryo in the system as a function of time and say the crystal has nucleated once n_{max} has crossed some critical size, denoted n_c, for the last time in the simulation. We selected $n_c = 75$ as this size is larger than any of the critical embryo sizes calculated in this study.

We prepared a set of starting configurations by equilibrating a cluster to an initial temperature $T_i = 900°K$, which is well above the freezing temperature and save configurations every 5×10^5 time steps. The simulation time between one starting structure and another serves to ensure that the initial configurations are independent from each other. Our MD simulation uses a fifth order Verlet predictor corrector

MD algorithm with a time step of 0.25×10^{-12} seconds in the $\{N, V, T\}$ ensemble using velocity re-scaling every 10 time steps so constant temperature simulations could be carried out. The starting configurations were quenched to one of the final temperatures $T_f[^\circ K] = \{600, 630, 640, 660, 680, 700, 720, 740, 760\}$, and the largest embryo size was calculated at every picosecond of the simulation. A total of 3×10^2 quenches were carried out at each temperature. The simulations were terminated after a total time of 500ps.

Fig. (3.5) shows three sample time series where $T_f = 660^\circ K$. The runs A, B and C, from fig. (3.5) have nucleation times 52,190, and 343 ps respectively. In particular, run C highlights the need to select the last time the system crosses n_c as the nucleation time. The growth of the embryo is a stochastic process with atoms randomly adding to and leaving from the embryo. As a result, in run C, the embryo appears to grow beyond the critical size at an early time but it then dissolves later before finally crossing the barrier into the solid state. In general, it turns out that the nucleation rate is relatively insensitive to the exact selection of the nucleation criteria used to define the nucleation time, while the lag time obtained from the same calculation is very sensitive[26]. We are however, only interested in the rate to nucleation.

Fig. (3.6 left), shows the fraction of un-nucleated clusters as a function of time for all the temperatures studied. Even at the slowest rate, less than $5 - 7\%$ of the clusters remained liquid by the end of the simulation run. Fig. (3.6), on the right shows the linear fits to the plots of the negative of the logaritm of un-nucleated systems as a function of time. By using eqn. (3.5), we notice that the slope of these plots is equal to JV_c. This data is summarized on table (3.2) and shown in fig. (3.7). The width of the error bars in the figure is equal to two standard deviations.

To extract the rate of nucleation from table (3.2), we need to factor out the volume of the cluster. Estimates for this volume were obtained via the Voronoi tessellation method [64], which is essentially a space-filling partition of space among a given set of points (in this case the atom coordinates), and where convex polyhedral regions are assigned to each point to find which atoms are in the surface. The

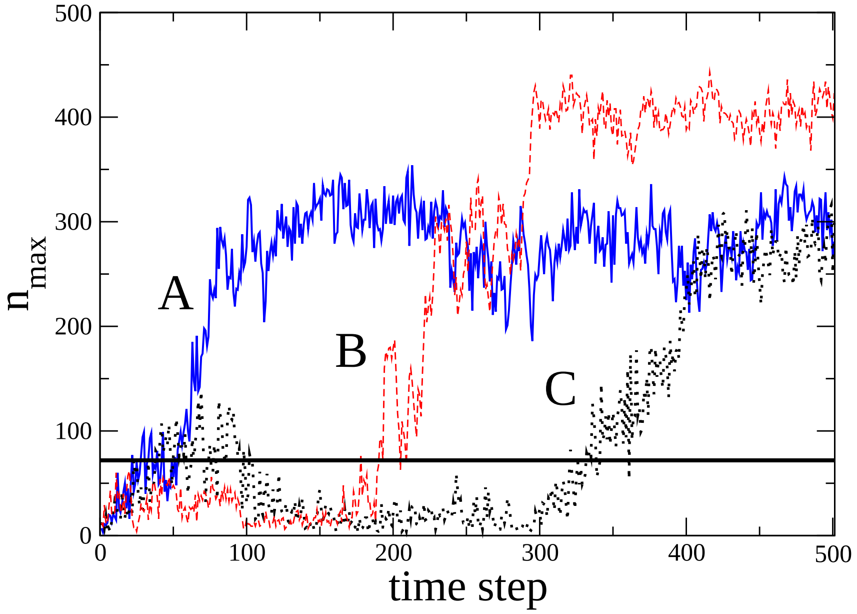


Figure 3.5: Evolution of the largest embryo size for three different configurations of the 456 atom cluster.

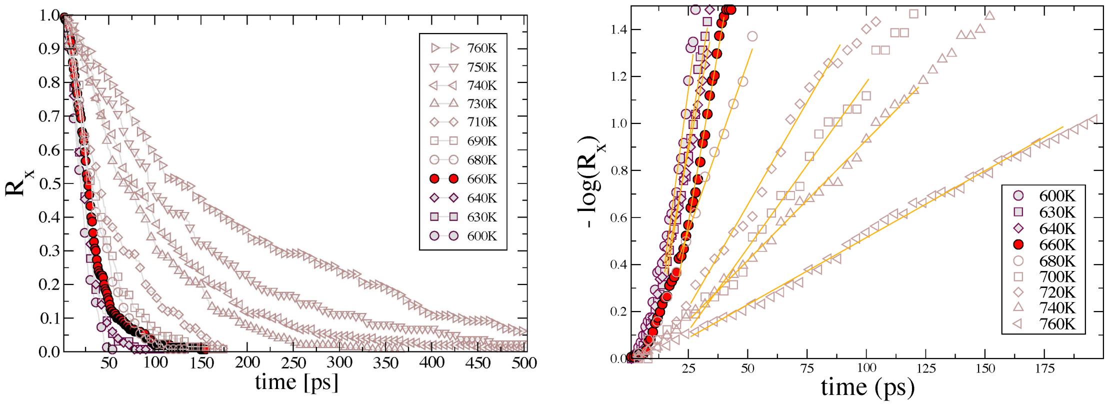


Figure 3.6: Left: Fraction of Un-nucleated systems. 456 atom cluster. Right: Taking the logarithm of the fraction of un-nucleated. The plot shows the linear fits from which the nucleation rate may be extracted by using eqn. (3.5).

total volume is therefore, the sum of the volumes of many individual polyhedra. The estimates for the volume are further refined by means of the "rolling sphere algorithm" [65] in which a probe sphere is rolled against a cluster consisting of hard spheres, resulting in a smooth surface from which the volume may be refined. A radius of 1.5$\mathring{A}$ was employed for both, the gold atoms and the probe sphere.

Temperature [$^\circ K$]	$JV_c[ps^{-1}]$	stdev
600	0.0710	0.002
610	0.0700	0.003
620	0.0660	0.002
630	0.0610	0.002
640	0.0577	0.002
650	0.0529	0.002
670	0.0395	0.002
680	0.0343	0.002
690	0.0291	0.002
700	0.0229	0.002
710	0.0180	0.002
720	0.0155	0.002
730	0.0122	0.003
740	0.0102	0.003
750	0.0078	0.004
760	0.0056	0.003

Table 3.2: Calculation of the rate of nucleation from the slope of the logaritm of un-nucleated systems vs time. The slopes are obtained from the slopes of the first order approximations to data illustrated in fig. (3.2), right hand side.

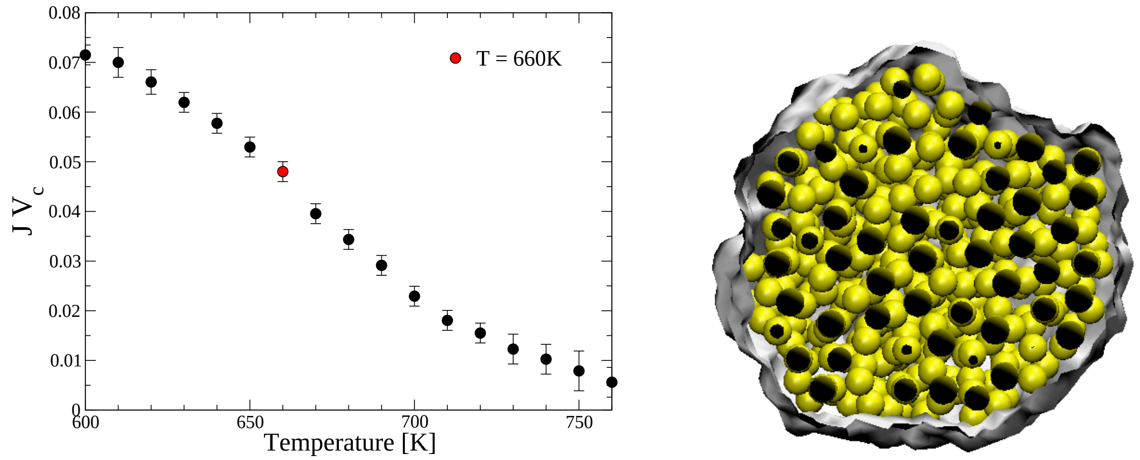


Figure 3.7: Left. Estimates of the dynamic rate from the slope of the logarithm of un-nucleated clusters. Right. Estimate of the volume of the cluster.

An average of over 100 configurations around the melting temperature of 750°K, resulted in a volume estimate of $7 \times 10^3 \pm 250 \mathring{A}^3$, for which the resulting averaged structure is spherical-like.

From the error deviations in the volume of the cluster V_c, and the quantity JV_c, from table (3.2), the largest error for the nucleation rate is less than $6 \times 10^{35} m^{-3} s^{-1}$. Finally, fig. (3.8) shows the rate of nucleation as a function of temperature, over a range of temperatures from $T = 760 - 600°K$.

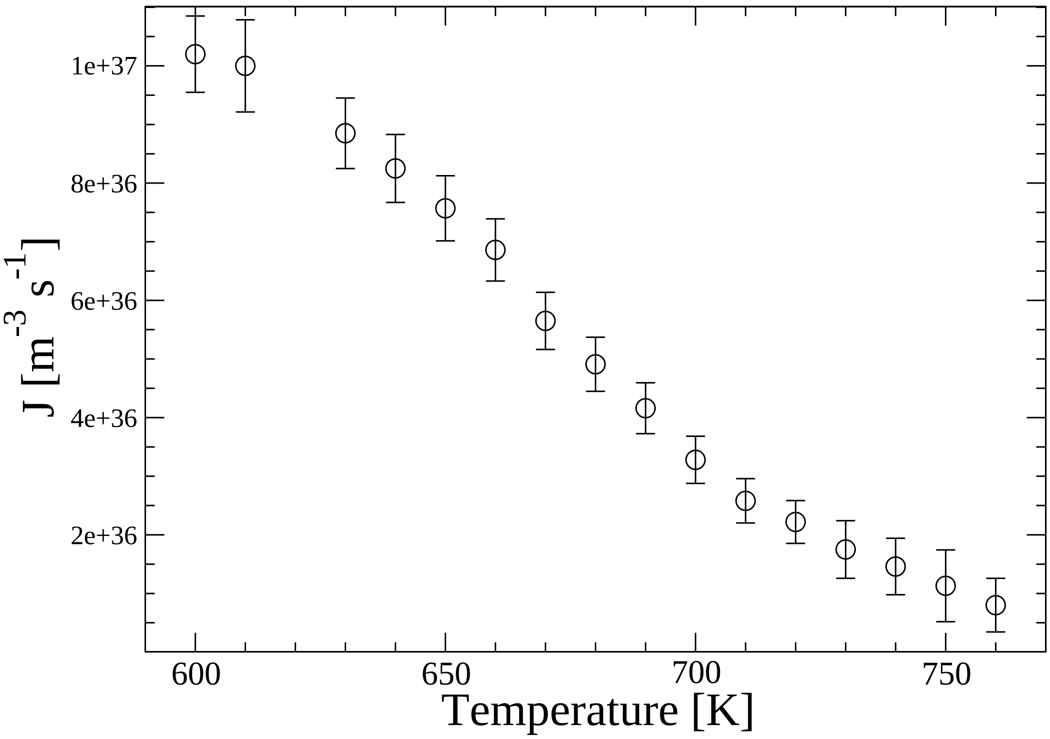


Figure 3.8: Estimates of the dynamic rate to nucleation as a function of temperatures for the 456 atom cluster. The volume of the cluster has been factored out from the data shown on table (3.2) and shown on fig. (3.7) left.

3.3 Summary

The rate of nucleation generally increases with decreasing temperature, as expected, since the nucleation barrier is getting lower as the system becomes more supercooled. This does assume that kinetic factors are not strongly temperature dependent. However, we see a marked change in the temperature dependence of J around $T = 700°K$ with the rate increasing more rapidly than expected. Such a trend would not be predicted on the basis of the free energy barriers calculated in Section (2.5). In fact, the temperature dependence of $\Delta F(n)/kT$ would suggest the rate of nucleation would become constant. i.e. the derivative of the rate with respect to temperature would become less negative with decreasing temperature. While the increase in the slope of the J vs T curve at $T = 700°K$ occurs at a temperature above our predicted spinodal point, it does coincide with the point where the barrier in $\Delta F(n_{max})/kT$ is in the order of kT, at which point the system can easily go over the barrier. This might explain the temperature dependence of the rate and support the idea that there is a limit of stability for the liquid at around $T = 660°K$.

An interesting feature of the spinodal predicted in our study is that the size of the critical embryo remains finite as the free energy barrier goes to zero. Mean field theories [35, 62] of the spinodal suggest that the critical embryo should diverge in size as the limit of stability is approached because of the ensuring very long range fluctuations. However, our results are consistent with recent simulations of *Pan et al* [44] who studied nucleation in a deep supercooled bulk Ising model.

CONCLUSIONS

Constrained Monte Carlo simulations, combined with parallel tempering, are used to calculate the free energy barriers to freezing in gold nano-clusters as a function of temperature.

First, we are able to prove that the difference between classic nucleation theory and our simulations is due to surface nucleation effects in which the solid embryo wets the surface of the cluster. This phenomena is incorporated in different phenomenological models which are built taking a series of refinements over classic nucleation theory and incorporating different surface and linear free energy terms.

Second, we propose a nucleation parameter with which we are able to identify a kinetic spinodal temperature where the nucleation barrier goes to zero and find that the critical cluster size remains finite at the limit of stability of the fluid phase. Molecular Dynamics simulations are used to examine the dynamics of freezing around kinetic spinodal temperatures, supporting the predictions of our Monte Carlo calculations.

Appendix I. Thermodynamic Stability

A phase transition is the transformation of a system from one thermodynamic phase into another. The distinguishing characteristic of a phase transition is the change in one or more physical properties with little change on the others, for example the abrupt change in heat capacity with respect to a small change in a thermodynamic variable such as the temperature signals a first order phase transition.

In order to define stability from a thermodynamic standpoint, let us first consider a system confined to a definite volume "V", with a constant number of particles "N" and no interchange of heat with the surroundings, i.e. the entropy "S" is constant. The conditions for stability of this system with respect to an initial state may be established by testing the changes with respect to all the possible variations under constant $\{S, V, N\}$ parameters:

$$[\Delta U]_{S,V,N} \geq 0. \tag{A-1}$$

If we expand L.H.S. of eqn. (A-1) as a variational of the changes in energies we obtain,

$$[\delta U + \frac{1}{2!}\delta^2 U + \frac{1}{3!}\delta^3 U + \ldots]_{S,V,N} \geq 0. \tag{A-2}$$

A vanishing linear term and a positive second order variation $\delta^2 U$ ensure that the minimum is stable for all variations subject to constant $\{S, V, N\}$:

$$\text{Stable equilibrium} \qquad \delta U|_{S,V,N} = 0, \qquad \delta^2 U|_{S,V,N} > 0. \tag{A-3}$$

while for both terms equal to zero we will have reached the limit of stability:

$$\text{Unstable equilibrium} \qquad \delta U|_{S,V,N} = 0, \qquad \delta^2 U|_{S,V,N} = 0. \tag{A-4}$$

The vanishing of the first order term δU constitutes therefore the equilibrium criterion, the positiveness of the second order variational $\delta^2 U$ constitutes the stability criterion. Thus to test the limit of stability, we require that the lowest order, non-vanishing energy variational of the energy functional be positive.

Stability of Pure Fluids

The internal energy for a fluid at constant temperature "T", pressure "P" and chemical potential "μ" is given by the expression:

$$dU = TdS - PdV + \mu dN, \tag{A-5}$$

since energy and entropy are not controlled experimental variables an alternative formulation of stability needs to be sought. By making changes of variables via Legendre transformations[1] we define the Gibbs free energy as

$$G = U - TS + PV. \tag{A-6}$$

Substitution of eqn. (A-5) into the eqn. (A-6) yields:

$$dG = -SdT + VdP + \mu dN. \tag{A-7}$$

Relation (A-5) is employed to define the conditions for stability in a manner analogous to eqs.(A-3) and (A-4). If we further impose the condition that we move along an isotherm, we are left with:

$$\Delta G = -V\Delta P + N\Delta\mu. \tag{A-8}$$

The new equilibrium conditions are formulated as $[\Delta G]_{T,P,N} \geq 0$ and the variational formulation similar to that in eqn. (A-2) can be used to test for stability criteria. When we have a multiphase system it is possible to write the expression for the internal energy in a more general way:

$$dU = TdS - PdV + \sum_{j=1}^{k} \mu_j dN_j, \tag{A-9}$$

Eqn. (A-9) describes changes in energy, entropy, volume and number of molecules

---

[1] These changes of variables are Legendre transformations, for example the Helmoltz free energy F is defined as $F = U - TS$, i.e. we change the variable entropy "S" by the variable temperature "T". i.e. The total derivative will now be a function of the changes in temperature, instead of the changes in temperature.

of a pure fluid along reversible quasistatic paths with k different fluids in thermodynamic equilibrium. We can further simplify eqn. (A-9) by writing:

$$dU = \sum_{j=1}^{k+2} Y_j dX_j, \tag{A-10}$$

with

$$Y_j = \left(\frac{\partial U}{\partial X_j}\right)_{X_1,X_2,\ldots,X_{j=1},X_{j+1},X_{j+2},\ldots,X_{k+2}} \tag{A-11}$$

The condition for stability $\delta^2 U|_{S,V,N} > 0$ can be written as:

$$\left(\frac{\partial Y_{k+1}}{\partial X_{k+1}}\right)_{Y_1,Y_2,\ldots,Y_k,X_{k+2}} > 0, \tag{A-12}$$

for a system in stable equilibrium (or metastable). For the limit of stability we have:

$$\left(\frac{\partial Y_{k+1}}{\partial X_{k+1}}\right)_{Y_1,Y_2,\ldots,Y_k,X_{k+2}} = 0, \tag{A-13}$$

These equations may only be used when the first variational is zero, i.e. $\delta U|_{S,V,N} = 0$. A detailed proof of eqs. (A-12) and (A-13) may be obtained from the literature. [33]

Gibbs Droplet Model

When we are dealing with interfaces where the number of particles in the system is fixed, we can no longer use eqn. (A-5) for the energy. Let us define the internal energy for this system : $U = TS - PV + \mu N + \sigma A$. Under constant T, P, σ and N we obtain:

$$dU = TdS - PdV + Nd\mu + \sigma dA, \tag{A-14}$$

where μ is the chemical potential of the system, σ is the energy cost of building an interface, and A is the surface of the interface. By taking the differential of the Gibbs free energy function $G = U - TS + PV$ and substitution of the differential of the internal energy in eqn. (A-14) we obtain:

$$dG = SdT + VdP + Nd\mu + \sigma dA, \tag{A-15}$$

furthermore, the change in free energies between an initial and a final state taken along the isotherm may be expressed by:

$$\Delta G = V\Delta P + N\Delta\mu + \sigma A. \tag{A-16}$$

We may write in the manner of eqs. (A-13) and (A-10), an expression for the limit of stability for the Gibbs condition $[\Delta G]|_{X_i} \geq 0$ as follows:

$$(\frac{\partial Y_2}{\partial X_2})_{Y_1,X_3,T} > 0, \tag{A-17}$$

for $X = \{\sigma, V, N\}$, and $\{A, P, \mu\}$.

$$(\mu, A, P) \Rightarrow \left(\frac{\partial \sigma}{\partial A}\right)_{N,P,T} > 0 \qquad (P, A, \mu) \Rightarrow \left(\frac{\partial \sigma}{\partial A}\right)_{V,\mu,T} > 0 \tag{A-18}$$

$$(\mu, P, A) \Rightarrow \left(\frac{\partial V}{\partial P}\right)_{N,A,T} > 0 \qquad (A, P, \mu) \Rightarrow \left(\frac{\partial V}{\partial P}\right)_{\sigma,\mu,T} > 0 \tag{A-19}$$

$$(P, \mu, A) \Rightarrow \left(\frac{\partial N}{\partial \mu}\right)_{V,A,T} > 0 \qquad (A, \mu, P) \Rightarrow \left(\frac{\partial N}{\partial \mu}\right)_{\sigma,P,T} > 0 \tag{A-20}$$

Eqn. (A-19 right), is especially useful for the experimentalist, since it may be compared to the isothermal compresibility κ_T. To get the condition on the limit of stability, with the spinodal temperature defined as the temperature at which:

$$\kappa_T = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T = 0. \tag{A-21}$$

For our detailed models, the variational expansion in the manner of eqn. (A-2) is preferred:

$$[\delta G + \frac{1}{2!}\delta^2 G + \frac{1}{3!}\delta^3 G + \ldots]_{X_i} \geq 0. \tag{A-22}$$

If we construct a polynomial expression dependent solely on n_{max}, the condition of equilibrium is reduced to testing the points where the first derivative respect to the order parameter is zero, and the condition of stability will be given by the second order derivative at the points where the equilibrium condition is met.

Appendix II. Equilibrium cluster distribution

Embryo size distribution vs Largest embryo probability.

The partition function for a system of N particles on a fixed volume and constant temperature T, is given by the expression

$$Q(N,V,T) = \frac{1}{\Lambda^{3N} N!} \int d\mathbf{r}^N exp[-\beta U(\mathbf{r}^n))], \tag{A-23}$$

where $U(\mathbf{r}^n)$ is the potential energy of a configuration of atoms with positions defined by $\mathbf{r}^n$ and $\Lambda = h/\sqrt{2\pi mkT}$ is the thermal Broglie wavelength. For now, we are to assume a criterion with which identification of embryos within a cluster of atoms is known, hence we define a special delta function as follows.

$$\delta_n(\mathbf{r}^n) = \begin{cases} 1 & \text{if all } n \text{ particles belong to the embryo,} \\ 0 & \text{otherwise.} \end{cases} \tag{A-24}$$

In addition to this, we define the function $\delta_r(\mathbf{r}^n)$, which ensures that all the other particles do not belong to the embryo.

$$\delta_r(\mathbf{r}^n) = \begin{cases} 1 & \text{if no other particle belongs to the embryo,} \\ 0 & \text{if any other particle belongs to the embryo.} \end{cases} \tag{A-25}$$

These two special delta functions are connected via the relation.

$$\delta_r(\mathbf{r}^n) = \prod_{i=n+1}^{N} [1 - \delta_{n+1}(\mathbf{r}^n, \mathbf{r}_i)] \tag{A-26}$$

We are now, in position to define a partition function for a system that contains at least one n-particle embryo

$$\begin{aligned} Q_n(N,V,T) = {} & \frac{1}{\Lambda^{3n} n!} \frac{1}{\Lambda^{3(N-n)}(N-n)!} \\ & \times \int d\mathbf{r}^n \int d\mathbf{r}^{N-n} \delta_n(\mathbf{r}^n)\delta_r(\mathbf{r}^N) \\ & \times exp[-\beta U(\mathbf{r}^n, \mathbf{r}^{N-n})], \end{aligned} \tag{A-27}$$

where is should be noted that there are as many as $N!/(n!(N-n)!)$ ways to select an n-particle embryo,

$$\begin{aligned} Q_n(N,V,T) = {} & \frac{1}{\Lambda^{3n} n!} \frac{1}{\Lambda^{3(N-n)}(N-n)!} \\ & \times \int d\mathbf{r}^{N-n} exp[-\beta U_{N-n}(\mathbf{r}^{N-n})] \\ & \times \int d\mathbf{r}^n \delta_n \delta_r exp[-\beta U_n(\mathbf{r}^n)] \\ & \times exp[-\beta U_{n,N-n}(\mathbf{r}^n, \mathbf{r}^{N-n})]. \end{aligned} \tag{A-28}$$

The effective potential for all the particles in the $n-sized$ embryo is defined as

$$U'_n = U_n - kT\ln[\delta_n], \tag{A-29}$$

and the interaction between the particles in the largest embryo size and the others

$$U'_{n,N-n} = U_{n,N-n} - kT\ln[\delta_r]. \tag{A-30}$$

Substitutions of eqs. (A-29), (A-30) into eqn. (A-28) leads to

$$\begin{aligned} Q_n(N,V,T) = {} & \frac{1}{\Lambda^{3n} n!} \frac{1}{\Lambda^{3(N-n)}(N-n)!} \\ & \times \int d\mathbf{r}^{N-n} exp[-\beta U_{N-n}(\mathbf{r}^{N-n})] \\ & \times \int d\mathbf{r}^n exp[-\beta U'_n] exp[-\beta U'_{n,N-n}]. \end{aligned} \tag{A-31}$$

Multiplication of the right-hand side of (A-31) by $Q(N-n,V,T)/Q(N-n,V,T)$

leads to:

$$
\begin{aligned}
Q_n(N,V,T) = & \frac{1}{n!\Lambda^{3n}} Q(N-n,V,T) \\
& \times \int d\mathbf{r}^n < exp[-\beta U'_{n,N-n}] > exp[-\beta U'_n],
\end{aligned} \tag{A-32}
$$

where we have defined the mean force potential:

$$
< exp[-\beta U'_{n,N-n}] >= \frac{\int d\mathbf{r}^{N-n} exp[-\beta U'_{n,N-n}] exp[-\beta U_{N-n}(\mathbf{r}^{N-n})]}{(N-n)!\Lambda^{3(N-n)} Q(N-n,V,T)}. \tag{A-33}
$$

The partition function for an n-particle embryo is defined as

$$
Q_n(N,V,T) = Q(N-n,V,T) q_n(V,T). \tag{A-34}
$$

The probability to find an embryo of size n in the N particle cluster is given by

$$
P_n = \frac{Q_n(N,V,T)}{Q(N,V,T)} = \frac{Q(N-n,V,T)}{Q(N,V,T)} q_n(V,T).
$$

Since the free energy of the system is given by $F = -kT\ln[Q]$, the above equation may be written as

$$
\begin{aligned}
P_n &= q_n(V,T) exp[+\beta\mu n], \\
q_n(V,T) &= \frac{1}{n!\Lambda^{3n}} \int d\mathbf{r}^n exp[-\beta U_{eff}].
\end{aligned}
$$

Where the effective potential U_{eff} is defined as $U_{eff} =< U_{n,N-n} > +U'_n$, thus we have:

$$
\begin{aligned}
q_n(V,T) &= \frac{1}{n!\Lambda^{3n}} \int d\mathbf{R}_{cm} \int d\mathbf{r}'^{n-1} exp[-\beta U_{eff}], \\
q_n &= \frac{V}{\Lambda_n^3} \times q_n^{internal},
\end{aligned}
$$

and the de Broglie wavelength for the embryo is defined by $\Lambda_n = h/\sqrt{2\pi nmkT}$ and:

$$
q_n^{internal} = \frac{n^{3/2}}{\Lambda_n^{3(n-1)}} \int d\mathbf{r}'^{n-1} exp[-\beta U_e].
$$

An intensive probability distribution will be defined as:

$$
\frac{P_n}{N} = \frac{1}{\rho\Lambda_n^3 n!} q_n^{internal} exp[-\beta\mu n],
$$

where ρ is the number density of the system. For rare embryos the probability can be written as

$$P_n = p_n(1) + p_n(2) + \cdots \approx p_n(1),$$

where $p_n(i)$ is the probability that there are exactly i embryos of size n. Furthermore, if we assume that the formation of different embryos is uncorrelated then it follows that $p_n(i) = [p_n(1)]^i$, then we can choose to neglect higher order terms as long as the probabilities associated are small, therefore the average number of embryos of size n in a cluster with N atoms equal to

$$N_n = 1p_n(1) + 2p_n(2) + 3p_n(3) + \dots$$

For rare clusters we have the following approximation,

$$\frac{P_n}{N} \approx \frac{N_n}{N} = \frac{1}{\rho\Lambda_n^3} q_n^{internal} exp[-\beta\mu n]. \tag{A-35}$$

At this point, it is important to stress that this quantity is classical, and hence should not depend on Planck's constant h, and in fact it does not, as the ideal gas part of the chemical potential $\mu = \mu^{ex} + kT\ln[\Lambda]$ cancels Plancks constant:

$$\frac{N_n}{N} = exp[-\Delta F(n)/k_BT]. \tag{A-36}$$

Eqn. (A-36) is used to obtain the free energy to crystallization $\Delta F(n)$ with respect to the liquid phase. The quantity N_n is obtained via ensemble averages, i.e.

$$\frac{< N_n >}{N} =< exp[-\Delta F(n)/k_BT] > . \tag{A-37}$$

Appendix III. Nucleation Program

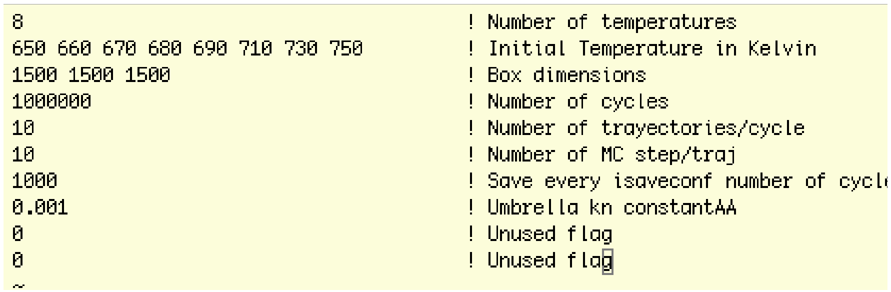
This section describes the computer code used to compute the free energy barriers for small clusters. The thermodynamic justification for such an algorithm has been described on section (1.2) and appendix A-II. The order parameter was introduced on section (2.4). The program consists on four main sections:

- Initialize.
- Distribute Workload.
- Check Point.
- Parallel Tempering.

We have implemented this code (see fig. A-2) in the fortran 90 programing language, and have used Message Passing Interface Libraries(MPI).

III.A Initialization

We start by reading the *options.in* input file which specifies the thermodynamic parameters for the simulation. The first two lines of the file define the size of the temperature array containing all the temperatures for the parallel tempering process. The third line defines the dimensions of the periodic boundaries at which the simulation is carried out (Angstrong units). Rows four, five and six are loop variables for number of cycles, the number of trajectories per cycle and the number of Monte Carlo steps per trajectory. Line number seven defines the number of configurations to be saved for post-processing. Line eight defines a value for the umbrella sampling parameter and the last two input lines are unused flags. A sample input file is shown on fig. (A-1). Once the data in the input file has been read, the program processes the pseudopotential file, setting cluster criteria parameters. It is straightforward



```
8                                       ! Number of temperatures
650 660 670 680 690 710 730 750         ! Initial Temperature in Kelvin
1500 1500 1500                          ! Box dimensions
1000000                                 ! Number of cycles
10                                      ! Number of trayectories/cycle
10                                      ! Number of MC step/traj
1000                                    ! Save every isaveconf number of cycl
0.001                                   ! Umbrella kn constantAA
0                                       ! Unused flag
0                                       ! Unused flag
~
```

Figure A-1: Typical *options.in* configuration file.

to change these parameters as well as the corresponding pseudopotential to recompile the code and carry out simulations for other unicomponent cluster systems. Furthermore, the code can be easily modified to study multicomponent systems.

Ideally, the number of processors required by a simulation would be chosen depending on the number of temperatures r and umbrella centers s that one requires, with the number of processors np given by $np = r \times s$. However, the computers at which we ran calculations require the specification of a restricted number of processing elements, therefore in practice one chooses a range of temperatures and a specific number of processors. From these, the number of umbrella centers is calculated for $s = np/r$.

The elements of the umbrella center vector are preset to start at $n_0 = 0$ with increments of $\Delta n = 10$, although again, it is straightforward to recompile the program to produce increments in any integer number.

III.B Workload Distribution

When the MPI directive is used to call the parallel program, every processor is identified with a unique ID number, a processor number i_{pe} ranging from 0 to $np - 1$. We associate every processor ID to an element of the $r \times s$ square grid with coordinates (n_0, T_0).

Once the umbrella center and temperature pair (n_0, T_0) for a processor has been defined, the processor seeks to load atomic configurations from a file containing the most recent configuration for the respective temperature and umbrella centers,

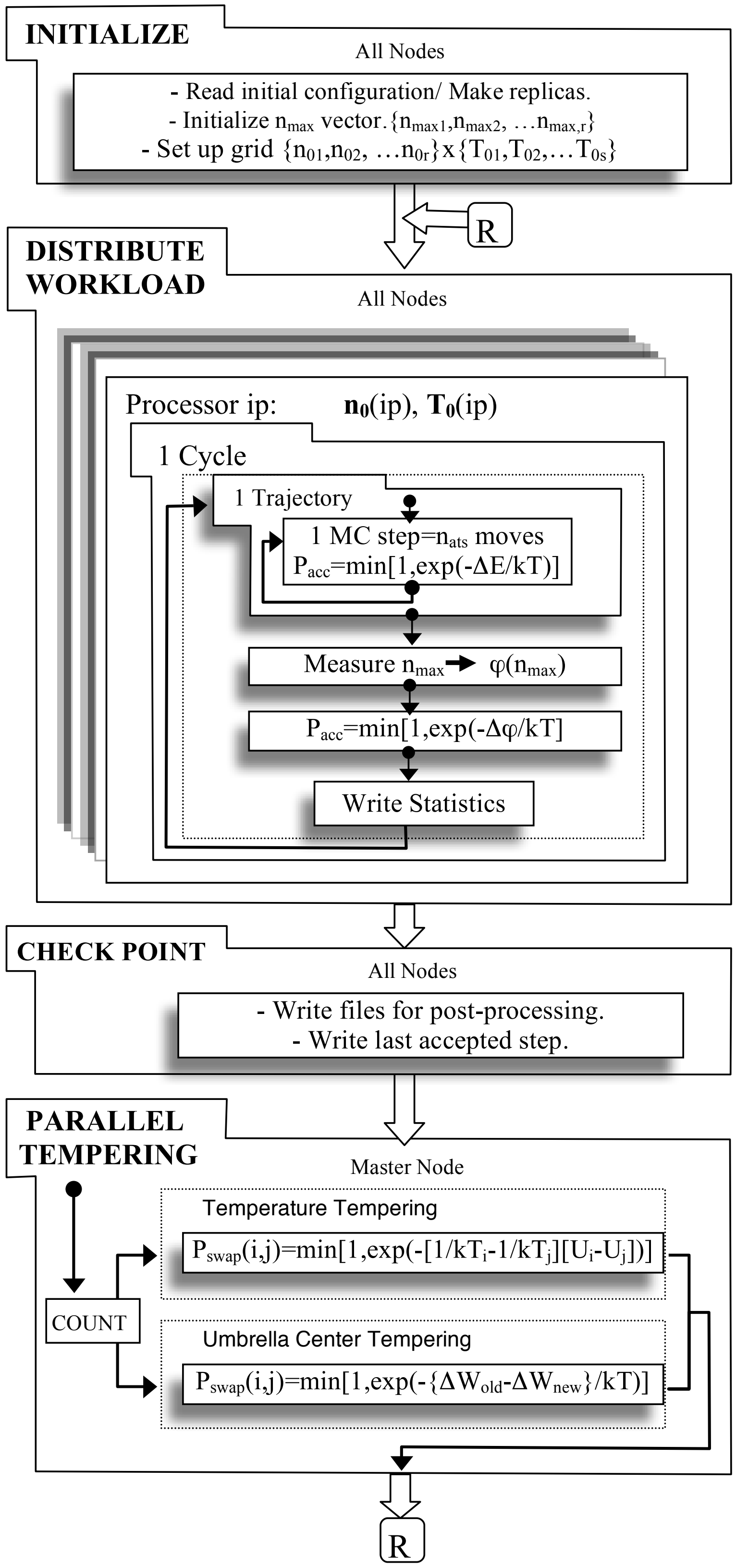


Figure A-2: Flow diagram for our Monte Carlo algorithm. Sampling techniques are discussed on the first chapter of this thesis work. Appendix (I) justifies the order parameter.

i.e. the simulation is restarted with the latest available coordinates. When the simulation runs for the first time, every processor loads a replica of the atomic coordinates located at the *atoms.xyz* file.

On the **xyz** format of the configuration files, the first line specifies the number of atoms and the second line is used to provide a short description , -usually text-describing the file contents. The rest of the file contains a four element listing that consists of atom type, and the X, Y, Z coordinates for each atom.

When the atomic coordinates for all the configurations for a cluster have been read, total energies are computed along with the largest embryo sizes and the corresponding embryo distribution. It is the task of every processor, to sample its own configuration according to the parameters specified in their individualized versions of a four dimensional vector which contains umbrella center $n_0(i_{pe})$, largest embryo size $n_{max}(i_{pe})$, temperature $T_0(i_{pe})$, and energy $E_0(i_{pe})$. This vector $V_{in}(i_{pe}) = (n_0, n_{max}, T_0, E_0)$, may have its second and fourth elements modified based on the Monte Carlo algorithm.

Trajectory Loop.

The Monte Carlo scheme followed by every processor has two main loops, namely the Monte Carlo steps loop and a Monte Carlo trajectory loop, with the former, nested inside the later.

Once the MC steps loop has been completed we evaluate the largest embryo size n_{max}. Ideally we would like to measure this order parameter every Boltzmann step on a single loop, we chose to make use of the fact that the largest embryo size fluctuates slowly as compared to the total energy, and hence we are able to avoid the need of making repeated n_{max} measurements.

The evaluation of the n_{max} order parameter in fact, requires computational time that scales as a polynomial ratio with respect to the number of atoms in the cluster. Unlike the total energy with which we can employ a range of computational algorithms to evaluate a difference in total energies made by a single Boltzmann move (see section 1.2) to make efficient calculations we require to process information from all the atoms to compute n_{max}.

The criteria for acceptance of a new cluster configuration is a function of the old and new largest embryo sizes and is given by:

$$P_{acc} = min[1, exp(-\Delta W/kT)], \tag{A-38}$$

with $\Delta W = \Phi(n_{max,old}) - \Phi(n_{max,new})$.

The umbrella potentials have harmonic forms, and are function of the umbrella parameter κ and the umbrella center $n_0(i_{pe})$.

$$\Phi(n_{max}) = \frac{1}{2}\kappa(n_0 - n_{max})^2. \tag{A-39}$$

When a trajectory has been accepted the embryo distribution and largest embryo size is updated, otherwise we reset the simulation to the values of the last accepted configuration, this includes retrieving old atomic coordinates as well as last embryo size and total cluster energy.

The combined effect of evaluating a chain of MC steps, followed by testing the embryo size and writing down the output to log files is termed of as one trajectory.

The number of trajectories that a processor performs is controlled by the fifth flag in the *options.in* file, "Number of trajectories per cycle". The whole chain of trajectories is termed of as once cycle.

When a cycle has been completed, the four dimension vector $V_{out}(i_{pe}) = (n_0(i_{pe}), n_{max}(i_{pe}), T_0(i_{pe}), E_0(i_{pe}))$ is turned over to the master node for purposes of tempering.

Monte Carlo Steps Loop.

In the Monte Carlo steps process, we perform a series of individual Monte Carlo moves (MC moves) with one MC move defined as the attempt to randomly select and move N atoms in an N-atom cluster. The acceptance of every atomic movement is ruled by Boltzmann sampling statistics where the probability of acceptance is given by:

$$P_{acc} = min[1, exp(-\Delta E/kT)], \tag{A-40}$$

with $\Delta E = E_{new} - E_{old}$ the difference in the configurational energies of the systems once the Boltzmann step has been made. The number of MC moves attempted in every MC steps loop is controlled by the flag specified in the sixth row of the *options.in* file. “Number of MC step/traj”[2]. At the end of every MC step, we adjust the magnitude of the random largest possible displacement ΔR in such a way that the acceptance rate is kept around 50%.

Log files.

The information of every trajectory is committed to two different files.

$Nmax.n_0.T_0$: The purpose of this file is to store information for the largest solid embryo at every accepted/rejected trajectory. The file contains information with accepted/rejected configurations, largest embryo size accepted, and cluster energy. From this file it is possible to evaluate the free energy from the ensemble average $< N_{nmax} >$.

$Nn.n_0.T_0$: This file contains detailed information on the last accepted embryo distribution. It contains embryo sizes, number of embryos, and largest embryo. From this file the free energy for the ensemble average $< N_n >$ may be computed. It is also possible to obtain $< N_{nmax} >$ as well, with the former file still required to provide an ensembe average on the energies. i.e. $< E_{nmax} >$.

III.C Check Point

The checkpoint section of the code was devised as a means to continue a simulation indefinitely. The idea behind this is to sample the system with the same thermodynamic conditions until some accuracy criteria in the resulting thermodynamic property being sampled has been reached. This usually requires much more time than the largest running time (usually 24:00 hrs) allowed by the high perfor-

---

[2] i.e. for $nmcsxt = 10$ we perform 10 MC steps

mance facilities we employed [3].

How many times a simulation ought to be executed, depends on the accuracy with which one desires to compute free energies. For any given Monte Carlo process, the variance of a sample is related to the number of sampling iterations by an inverse square root relationship, i.e. $\sigma(P(k)) = 1/\sqrt{(}k)$, with k the number of elements in the sample, P the measurement, and σ the standard deviation. In practice we performed calculations until the resulting standard deviations for free energies cold be set below 1 kT.

In addition to merely computational issues, it is not uncommon that the system may improperly terminate the execution of a simulation due to scheduled system maintenance. To this effect we save data in the binary format and update at the next check point. The files containing the latest accepted configurations have the generic name $ctmp.n_0.T_0$, with $n_0.T_0$ the umbrella center and temperature of the respective processor. Data is saved using binary format on files called $ctmp.n_0.T_0$.

In addition to the *ctmp* files, the program saves information into $config.n_0.T_0$ files with the intention of making post-processing of information available. The flag that controls how many cycles are to pass in between $config$ saving events is defined by the seventh line in the *options.in* file.

III.D Parallel Tempering

Once every node has completed a cycle, the master node receives np four dimensional vectors $V_{out}(i_{pe}) = (n_0(i_{pe}), n_{max}(i_{pe}), T_0(i_{pe}), E_0(i_{pe}))$ for parallel tempering, i.e. the interchange of conditions of simulation may be made based on temperature or based on umbrella centers. We chose to interchange thermodynamic parameters instead of atomic configurations for obvious reasons: Swapping configurations would cost us transferring at least $nats \times 3 \times np$ variables, notwithstanding the amount of time spent transferring data. Instead we just need to transfer $np \times 4$ parameters.

Our parallel program was preset in such a way that a counter rules that all pro-

---

[3] see http:www.westgrid.ca to whom we are grateful for the computing time that made this project possible

cessors alternate between trying 10 temperature tempering attempts followed by an equal number of umbrella center attempts. To this effect I follow a “double-ladder” approach that I describe below.

Umbrella Center Swaps

We start by selecting a pair of processors with ID’s $i = 0$, and $j = 1$, and obtain all the elements allocated on the respective vectors $V_{in}(i)$ and $V_{in}(j)$.

We then move the j flag upwards, i.e. by increments of 1, until we find that the temperatures of the vector pair are the same, and in such a case we attempt an umbrella center swap with probability:

$$P_{swap}(i,j) = min[1, exp(-[\Delta W_{new} - \Delta W_{old}]/kT)], \tag{A-41}$$

where

$$\begin{aligned} W_{old} &= -\frac{1}{2}\kappa[n_{max}(i) - n_0(i)]^2 - \frac{1}{2}\kappa[n_{max}(j) - n_0(j)]^2, \\ W_{new} &= -\frac{1}{2}\kappa[n_{max}(j) - n_0(i)]^2 - \frac{1}{2}\kappa[n_{max}(i) - n_0(j)]^2. \end{aligned}$$

If the acceptance criteria is met, we swap the umbrella centers of the vectors $V_{in}(i)$, $V_{in}(j)$. In such a case the value of the i flag has to be increased, with the flag j reset to $i + 1$, this is made to avoid repetition of swapping attempts.

We repeat the raising of the i flag until it reaches a value of $i = np - 2$ with the flag j in the interval from $i + 1$ to $np - 1$.

Temperature Swaps

Temperature swaps follow an analogous process to that described above. This time only we are sweeping V_{in} vectors until we find pairs with the same umbrella centers. Once we encounter an appropriate pair of vectors, we extract their parameters and evaluate the acceptance rule:

$$P_{swap}(i,j) = min[1, exp((1/kT_i - 1/kT_j)(E_i - E_j))], \tag{A-42}$$

where E_i is the energy of the cluster with ID number i, simulated at the temperature T_j whereas the cluster with ID number j is being simulated at a temperature T_j and possesses an energy E_j. The process is followed until we have tried all possible pairs for tempering. Finally we transfer control to all the processors by feeding to these the updated versions of the state vectors $V_{in}(i_{pe})$. The procedure outlined so far completes a cycle, which in practice is scheduled to run for twenty four hours, until the system manager terminates it.

The idea of running the process indefinitely is chosen instead of the alternative procedure of running a certain amount of cycles for a set amount of time. The purpose of this is to complete as many cycles as possible and avoid estimating computation time with the subsequent possibility of reducing available resources. It will be always possible to restart the program from the last available configurations until the statistics to guarantee a target degree of accuracy have been accumulated.

Appendix IV. Detailed Balance.

The general approach followed to demonstrate the validity of a MC algorithm is outlined below.

- A distribution function π has to be defined. This distribution depends on the thermodynamic constants of the system.
- The *detailed balance* condition is imposed, this condition states that the probability of a system to evolve from an initial state $q = o$ to a final state $q = n$ must be equal to that from the similar system evolving from state n to state o. In other words we require:

  $$\Phi(o \mapsto n) = \Phi(n \mapsto o), \tag{A-43}$$

  where Φ is the flow of configuration o to n given by the products of the probability $\pi(o)$ to be in the configuration o, the probability α of generating the configuration n, and $acc(o \mapsto n)$, the probability of accepting the move.

  $$\Phi(o \mapsto n) = \pi(o) \times \alpha(o \mapsto n) \times acc(o \mapsto n), \tag{A-44}$$

- Probabilities of generating a configuration are determined.
- The acceptance rule condition is evaluated until enough statistics are accumulated.

In particular, for the canonical ensemble $\{N, V, T\}$ the distribution function is given by

$$\pi(\mathbf{r}^N) = \frac{exp[-\beta U(\mathbf{r}^N)]}{\int exp[-\beta U(\mathbf{r}^N)] d\mathbf{r}^N}. \tag{A-45}$$

The probability of generating a particular configuration should be independent of the conformation of the system, hence we have with aid of the detailed equilibrium condition:

$$\alpha(o \mapsto n) = \alpha(n \mapsto o) = \alpha, \tag{A-46}$$

Substitution of eqn. (A-46) into the detailed balance condition, eqn. (A-43) and the further substitution of the sampling distribution (A-45), into expression (A-44), yields the acceptance rule condition for the $\{N, V, T\}$ ensemble, already introduced in section (1.2):

$$\frac{acc(o \mapsto n)}{acc(n \mapsto o)} = \exp\{-\beta[U(n) - U(o)]\}, \tag{A-47}$$

The detailed balance condition implies that enough MC energy sampling steps must be carried out once two configurations equilibrated to different thermodynamic conditions have been interchanged, this is to ensure that the sampling of embryo configurations is sampled under equilibrium conditions.

The umbrella sampling plus parallel tempering algorithm utilized in our calculations consists of 64 nodes, each one with its umbrella center and temperature, and an associated partition function given by (see section 2.5):

$$Q_C = \prod_{\mu=1}^{8}\prod_{\nu=1}^{8} Q_{N,V,T_\mu,H_c(\mu,\nu)}.$$

Let us denote the configuration of node i by $\mathbf{i} = \mathbf{r}_i^N$, and its associated constrained hamiltonian by $H(i) = H(i)_0 + \phi_i$, where $H(i)_0$ is the unconstrained Hamiltonian and ϕ_i is the bias potential, with an associated Boltzmann parameter $\beta_i = \frac{1}{k_B T_i}$.

The acceptance rule for a swap between ensembles i, and j, follows from the condition of detailed balance (A-43) and is given by the expression:

$$\begin{aligned}
&\pi(i,\beta_i)\pi(j,\beta_j) \times \alpha[(\mathbf{i},\beta_i),(\mathbf{j},\beta_j) \mapsto (\mathbf{j},\beta_i),(\mathbf{i},\beta_j)] \\
&\qquad \times acc[(\mathbf{i},\beta_i),(\mathbf{j},\beta_j) \mapsto (\mathbf{j},\beta_i),(\mathbf{i},\beta_j)] \\
&= \pi(i,\beta_j)\pi(j,\beta_i) \times \alpha[(\mathbf{i},\beta_j),(\mathbf{j},\beta_i) \mapsto (\mathbf{i},\beta_i),(\mathbf{j},\beta_j)] \\
&\qquad \times acc[(\mathbf{i},\beta_j),(\mathbf{j},\beta_i) \mapsto (\mathbf{i},\beta_i),(\mathbf{j},\beta_j)]
\end{aligned}$$

If we perform simulations in such a way that the probability of swapping umbrella centers and temperatures occurs with a sampling probability α, we obtain as acceptance rules:

$$\begin{aligned}
&\frac{acc[(\mathbf{i},\beta_i),(\mathbf{j},\beta_j)\mapsto(\mathbf{j},\beta_i),(\mathbf{i},\beta_j)]}{acc[(\mathbf{i},\beta_j),(\mathbf{j},\beta_i)\mapsto(\mathbf{i},\beta_i),(\mathbf{j},\beta_j)]}\\
&\qquad=\frac{[-\beta_i H(j)-\beta_j H(i)]}{[-\beta_i H(j)-\beta_j H(i)]}\\
&\qquad=exp\{(\beta_i-\beta_j)[H(i)-H)j)]\}.
\end{aligned} \tag{A-48}$$

Since the rate of exchange of umbrella centers and temperatures is not the same in the course of the simulation, we set different rules for swapping umbrella centers (umbrella sampling) than for exchanging temperatures (parallel tempering). It is straightforward to show that eqn. (A-41) and (A-42) correspond to these two case scenarios respectively.

Appendix V.
Estimation of parameters and efficient sampling.

In a typical MC simulation, we begin calculations with structures which do not comply with the detailed balance condition (A-43), and therefore the acceptance rule condition (A-47) is not valid. To avoid this problem we require the system to evolve for a reasonable number of time steps, until it settles to thermodynamic equilibrium. This requirement hinders the choice of appropriate umbrella sampling plus parallel tempering parameters.

Fig. (A-3) shows the effects of cutting down a number of initial trajectories in a simulation. Clearly, convergence in the calculated free energies to crystallization is reached only after around 2×10^4 trajectories have been cut down from the calculations of the free energy.

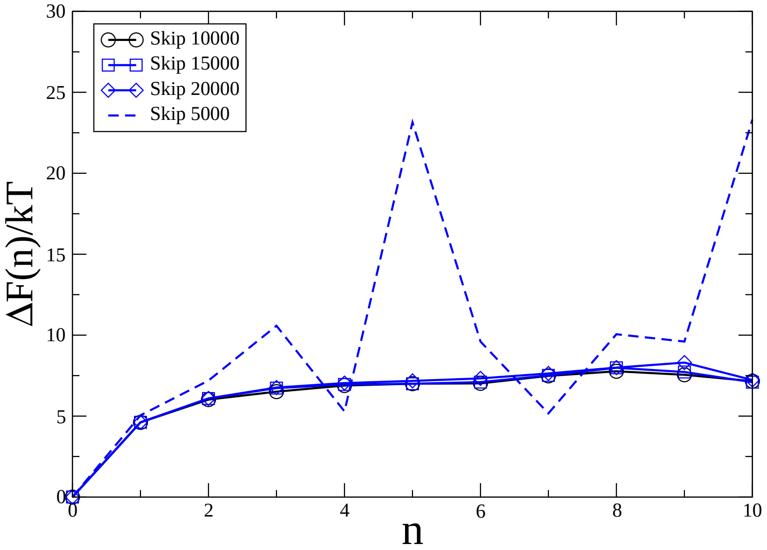


Figure A-3: This figure illustrates the need to skipping non equilibrated configurations to ensure convergence. The free energy appears to be convergent only after the initial 1×10^3 trajectories are ignored. With the further condition that sampling is being made uniformly over all n, the free energy will be more accurate for increasing sampling.

In order to produce an ensemble of configurations, each equilibrated to its respective umbrella center $n_{0,\mu}$ and temperature T_ν, we need to find appropriate umbrella sampling constants, as well as temperatures, and definitions of trajectory as to comply with the following conditions:

1. *Local equilibrium.* When the swapping of two adjacent configurations takes place, long enough energy equilibrations should be carried out, before the embryo distribution is computed. This will enable the system to evolve to the new thermodynamic conditions.

2. *Uniform swapping.* Swapping of configurations with different temperatures occurs at a comparable rate than swapping of configurations with different umbrella centers.

It is not unusual to have to test different combinations of temperatures and umbrella sampling constants, to produce long time energy simulations, only to discover that even when cutting down a large number of trajectories in the simulation the system does not converge properly. In such a case the simulation has to be repeated for the new conditions. A quick estimate for the appropriate umbrella center parameters may be obtained by assuming that the clusters of two neighboring systems i and j behave as an harmonic crystals, and hence eqn. (A-4) can be used to derive a targeted swapping acceptance probability. This is achieved by assuming some threshold value of acceptance for the quantity $\Delta W_{new} - \Delta W_{old}$ and computing the approximate umbrella parameter κ. In practice, when the umbrella constant is too large, the acceptance rate will be too high, and the system will evolve towards local energy minima. If the umbrella constant is too small, each one of the nodes will behave as an uncoupled Metropolis MC simulation.

Finally, the temperatures for the parallel tempering scheme, ought to be spaced in such a way, that the swapping acceptance rate for interchanging temperatures is close to the swapping acceptance rate for the interchange of umbrella centers.

In our simulation we achieved sampling rates for all nodes between 40% and in some instances about 75%, for an umbrella constant equal to 0.001. The efficiency of the algorithm is further justified, when one obtains smooth histograms for the n and n_{nmax} embryo distributions in the constrained space. In fig. (A-4 left), we illustrate the process of exchange of configurations (see Appendix A-III), for nodes with the same temperature and different umbrella centers, that correspond to successful exchange attempts under the umbrella sampling scheme. Fig. (A-4 right), illustrates the process of exchange of configurations for nodes with the same umbrella center $n_0 = 40$, but different temperatures. All the rest of the histograms, whether for the exchange at constant temperatures, or constant umbrella centers, are distributed in a similarly sparse fashion. The combination of the umbrella sampling and parallel tempering algorithm provides robustness to the sampling algorithm. The sampling rules of the algorithm were defined in appendix A-IV.

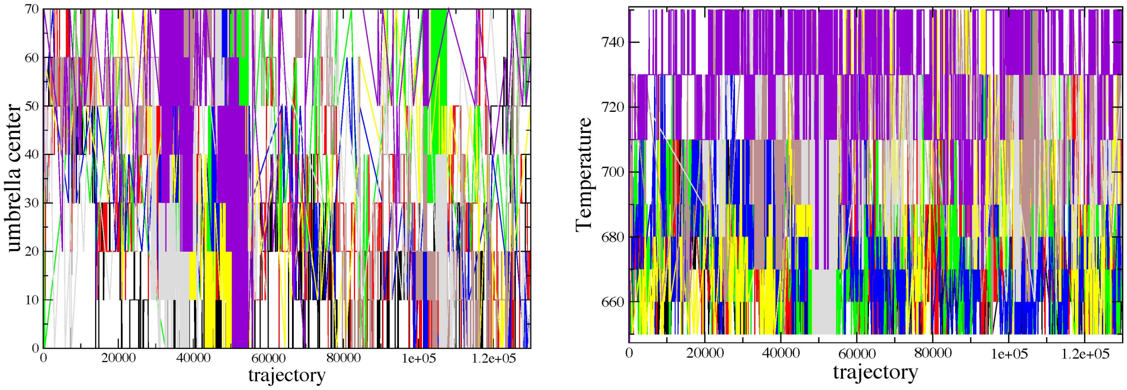


Figure A-4: Left: exchange of configurations with different umbrella centers, and the same temperature $T = 690^{\circ}K$, along 1.3×10^5 trajectories. Right: Exchange of configurations with different temperature for an umbrella center equal to $n_0 = 40$.

Fig. (A-5) is a closeup of the parallel tempering exchange of configurations from fig. (A-4 right), showing that exchange is regular at long and in the local scale. The data sets in the figure have been displaced vertically, just to show clearly the exchange. The umbrella sampling exchange behaves likewise.

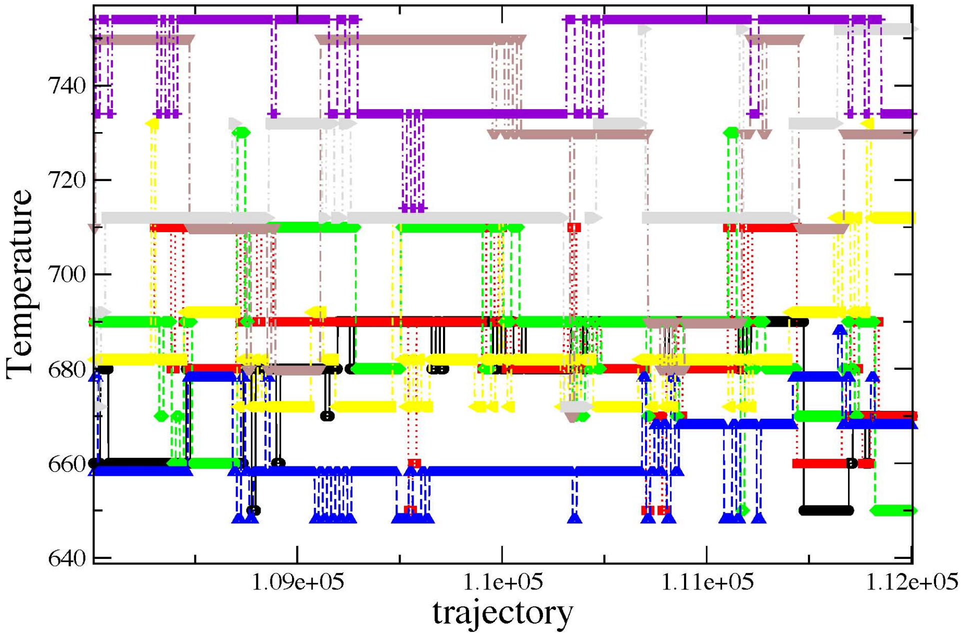


Figure A-5: Umbrella sampling: Switching configurations.

References

[1] A. Watson, *Science* **306**, 1309 (2004).

[2] G. D. Scholes and G. Rumbles, *Nature Materials* **5**, 683 (2006).

[3] J. V. Barth, G. Costantini and K. Kern, *Nature* **437**, 671 (2005).

[4] R. S. Berry, *Nature* **393**, 212 (1998).

[5] A. W. Woods and S.S. Cardoso, *Nature*, **385** , 518 (1997).

[6] *Climate Change 2001: The Scientific Basis*, Cambridge University Press,Cambridge U.K. (2001).

[7] A. Tabazadeh, Y. S. Djikaev and H. Reiss, *Proc. of the Nat. Acad. of Sci. of the USA* **99**, 15873 (2002).

[8] A. Tabazadeh, Y. S. Djikaev, P. Hamill and H. Reiss, *J. Phys. Chem. A* **106**, 10238 (2002).

[9] Y.S. Djikaev, A. Tabazadeh, P. Hamill and H. Reiss, *J. Phys. Chem. A*, **106**, 10247 (2002).

[10] P. G. Debenedetti, *Nature* **441**, 168 (2006).

[11] W. Steubing, *Ann. Phys.* **24**, 1 (1907); *Ann. Phys.* **26**, 329 (1908).

[12] J. G. Allpress, J.V. Sanders *Surf. Sci.* **7**, 1 (1967).

[13] P. A. Buffat, Marc Fleli, R. Spycher, P. Stadelmann and J.-P. Borel, *Faraday Discuss.* **92**, 173 (1991).

[14] K. Kimoto and I. Nishida, *I. Jpn. J. Appl. Phys.* **6**, 1047 (1967).

[15] S. Ino, *J. Phys. Soc. Jpn.* **21**, 346 (1966).

[16] Shaun C. Hendy and J. P. K. Doye, *Phys. Rev. B*, **66**, 235402 (2002).

[17] C. L. Cleveland, W. D. Luedtke, and Uzi Landman, *Phys. Rev. B* Vol. 60, Number 7, 5065 (1999).

[18] C. L. Cleveland, W. D. Luedtke, and Uzi Landman, *Phys. Rev. Lett.* 81, 2036 (1998).

[19] M. R. Goldsmith, C. C. George, G Zuber, R. Naaman, D. H. Waldeck, P. Wipf, D. N. Beratan, *Phys. Chem. Chem. Phys.*, **8**, 63 (2006).

[20] R.N. Barnett, C.L. Cleveland, H. Hakkinen, W.D. Luedtke, C. Yannouleas, and U. Landman *Eur. Phys. J. D* **9**, 95 (1999).

[21] Y. G. Chushak and L. S. Bartell *J. Phys. Chem. B* **105**, 11605 (2001).

[22] J. Huang and L. Bartell *J. Phys. Chem. A* **106**, 2404 (2002).

[23] H.-S. Nam, Nong M. Hwang, B. D. Yu, D.-Y. Kim, and J.-K. Yoon, *Phys. Rev. B* **71**, 233401 (2005).

[24] S. Auer and D. Frenkl, *Nature* **409**, 1020 (2001).

[25] S. Auer and D. Frenkl, *J. Chem. Phys.* **120**, 3015 (2004).

[26] I. Saika-Voivod, P.H. Poole and R. K. Bowles, *J. Chem. Phys.* **124**, 224709 (2006).

[27] S. Auer, D. Frenkel, *Phys. Rev. Lett.* **91**, 015703 (2003).

[28] *"Homogeneous nucleation Theory. The Pretransition Theory of Vapour Condensation"*. New York, Academic Press (1974). ISBN: 0120383616

[29] K. F. Kelton, A.L. Greer, and C.V. Thompson, *J. Chem. Physics.* **79**, 6561 (1983).

[30] Landau, L.D. and Lifshitz, E.M., Statistical Physics Part 1, Vol. 5 *also* Vol. 9 (Statistical Physics Part 2) Pergamon, 3rd Ed. (1994).

[31] Yu. B. Zeldovich *Acta Physicochim.,* URSS **18**, 1 (1943).

[32] P. G. Debenedetti, *Nature*, **441**, 168 (2006).

[33] P. G. Debenedetti *"Metastable Liquids", Chapter 2: "Thermodynamics".*
and *Appendix 2: "Thermodynamics of fluid interfaces".*
Princeton University Press: Princeton, NJ, 1996. ISBN 0 691 08595

[34] Wales D.J. *"Energy Landscapes", Chapter 8: "Clusters".*
Cambridge Molecular Science (2003). ISBN 0 521 81415 4

[35] M. Biskup and L. Chayes, *Commun. Math. Phys.* 238 (2003), no. 1-2, 53-93.

[36] T. S. Ray, *Journal of Statistical Physics*, **62**, 463, (1991).

[37] P. G. Debenedetti and H. Reiss, *J. Chem. Phys.* Vol. 108, Number 13 (1998).

[38] A. Laaksonen, R. McGraw, and H. Vehkamaki, *J. Chem. Phys.* Vol. 111, No. 5, 2019 (1999).

[39] G. Wulff, *Kristallogr. Mineral.* **34**, 449 (1901).

[40] P. R. Ten Wolde and D. J. Frenkel, *J. Chem. Phys.* **109**, 9919 (1998).

[41] H. Reiss and R. K. Bowles, *J. Chem. Phys.* **111**, 7501 (1999).

[42] C. Valeriani, E. Sanz, and D. Frenkel, *J. Chem. Phys.* **122**, 194501 (2005).

[43] P. J. Steinhardt, D. R. Nelson, M. Ronchetti, *Phys. Rev. Lett. B* **28**, 784 (1983).

[44] A. C. Pan, T. J. Rappl, D. Chandler and N. P. Balsara, *J. Phys. Chem. B* **110**, 3692 (2006).

[45] *"Reviews in Computational Chemistry", Vol. 9. Chapter 2: "Free Energy by Molecular Simulation", by T.P. Straatsma.* Kenny B. Lipkowitz and Donald B. Boyd, Editors. VCH Publishers, Inc. New York (1996). ISBN 1069-3599.

[46] G. M. Torrie and J. P. Valleau, *J. of Comp. Phys.*, **23** 187 (1977).

[47] W. C. Swope, H.C. Andersen, P.H. Berens, and K. Wilson, *J. Chem. Phys.* **76**, 637 (1982).

[48] M. S. Daw, S. M. Foiles, and M. Baskes, *Mat. Sci. Rep.* **9** ,251 (1993).

[49] F. Ercolessi, M. Parrinello, and E. Tosatti, *Phil. Mag. A* **58(1)** 213 (1988).

[50] C. L. Cleveland, U. Landman, T. G. Schaaff, M. N. Shafigullin, P. W. Stephens, and R. L. Whetten, *Phys. Rev. Lett.* **79**, 1873 (1997).

[51] D. Reinhard, B. D. Hall, P. Berthoud, S. Valkealahti, and R. Monot *Phys. Rev. Lett.* **78**, 8 (1997).

[52] H.-S. Nam, Nong M. Hwang, B. D. Yu, and J.-K. Yoon, *Phys. Rev. Lett.* **89**, 275502 (2002).

[53] J. E. Jones and A. E. Ingham, *Proc. R. Soc. A* **107**, 636 (1925).

[54] P. Vinet, J.H. Rose, J. Ferrante and J. R. Smith *J. Phys.: Condens. Matter* **1**, 1941 (1989).

[55] F. D. Murnaghan, *Am. J. Math.* **49**, 235 (1937).

[56] H. Deng and J. Huang, *J. Sol. State Chem*, **159**, 10 (2001).

[57] Y. Wang, S. Teitel and C. Dellago, *J. Chem. Phys.* **122** 214722 (2005).

[58] P. G. Debenedetti, *"Metastable Liquids"*, Chapter 3; Kinetics. Princeton University Press: Princeton, NJ, 1996. ISBN 0 691 08595

[59] L. J. Lewis, P. Jensen and J. L. Barrat, *Phys. Rev. B*, **56** 2248 (1997).

[60] H. M. Princen, in *Surface and Colloid Science*, edited by E. Matijevic (Wiley-Interscience, New York, 1969) Vol 2.

[61] A. A. Nepomnyashchy, A. A. Golovin, A. E. Tikhomirova and V. A. Volpert, *Phys. Rev. E*, **74** 021605 (2006).

[62] J. W. Cahan and J.E. Hilliard, *J.Chem. Phys.* **31**, 688 (1959).

[63] Y. G. Chushak, P. Santikary, and L. S. Bartell, *J. Phys. Chem A* **103**, 5636 (1999); G.W. Turner, Y. G. Chushak and L. S. Bartell, *J. Phys. Chem A* **108**, 1666 (2004).

[64] W. J. Brostow, P. Dussault and B. L. Fox, *J. Comput. Physics*, **29**, 81 (1978).

[65] P. Alard and S. J. Wodak, *J. Comp. Chem.* **12**, 918 (1991).