Structures and dynamics investigation of phase selection in metallic alloy systems

by

Lin Yang

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Physics (Condensed Matter Physics)

Program of Study Committee:
Kai-Ming Ho, Major Professor
    Paul Durbin
    James W. Evans
    Alex Travesset-Casas
    James P. Vary

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this dissertation. The Graduate College will ensure this dissertation is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University
Ames, Iowa
2018

Copyright © Lin Yang, 2018. All rights reserved.
DEDICATION

I would like to thank my supervisors, my family, and my friends, for their instructive and loving guidance on all aspects during my Ph.D. study.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>x</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>xii</td>
</tr>
<tr>
<td>CHAPTER 1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>CHAPTER 2. LITERATURE REVIEW</td>
<td>7</td>
</tr>
<tr>
<td>2.1 Simulation Methods</td>
<td>7</td>
</tr>
<tr>
<td>2.1.1 Molecular dynamics simulation</td>
<td>8</td>
</tr>
<tr>
<td>2.1.2 Monte Carlo simulation</td>
<td>15</td>
</tr>
<tr>
<td>2.1.3 Tricks of computer simulation</td>
<td>16</td>
</tr>
<tr>
<td>2.2 Analysis Methods</td>
<td>21</td>
</tr>
<tr>
<td>2.2.1 Structural analysis</td>
<td>22</td>
</tr>
<tr>
<td>2.2.2 Dynamical analysis</td>
<td>25</td>
</tr>
<tr>
<td>2.2.3 Nucleation process modeling</td>
<td>26</td>
</tr>
<tr>
<td>2.2.4 Free energy computing</td>
<td>28</td>
</tr>
<tr>
<td>2.2.5 Experimental techniques</td>
<td>31</td>
</tr>
</tbody>
</table>
CHAPTER 3.  A PROCEDURE TO EVALUATE THE FREE ENERGY OF LIQUID AND SOLID PHASES OF AN ALLOY SYSTEM .......................... 33

3.1 Introduction .......................................................... 33

3.2 Computational Details ............................................. 36
   3.2.1 Pure fcc-Al and Al liquid ..................................... 37
   3.2.2 Al$_{1-x}$Sm$_x$ liquid ......................................... 40

3.3 Applications ....................................................... 42
   3.3.1 Driving force for nucleation of fcc-Al in supercooled Al$_{1-x}$Sm$_x$ liquid 42
   3.3.2 Melting curve (liquidus line) for fcc-Al and Al$_3$Sm .................. 44

3.4 Conclusion ......................................................... 47

CHAPTER 4. KINETIC EFFECT IN THE DETERMINATION OF THE NON-STOICHIOMETRIC SITE OCCUPATIONS .................. 48

4.1 Introduction ......................................................... 49

4.2 Experiments ....................................................... 50

4.3 Simulation Details ................................................ 55
   4.3.1 Monte Carlo simulation ....................................... 55
   4.3.2 Molecular dynamic simulation ............................... 59

4.4 Conclusion ........................................................ 62

CHAPTER 5. PERSISTENT-EMBRYO METHOD SIMULATION OF DEVITRIFICATION PROCESS OF $\epsilon$-Al$_{60}$Sm$_{11}$ PHASE .......... 64

5.1 Persistent-Embryo Method ....................................... 65

5.2 Simulation Details ................................................ 67
   5.2.1 Identifying solid/liquid-like atoms .......................... 67
   5.2.2 Embryo of $\epsilon$-Al$_{60}$Sm$_{11}$ .............................. 70

5.3 Results and Analysis ............................................. 71
LIST OF TABLES

| Table 6.1 | Accuracy Testing Overview | 81 |
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1</td>
<td>Debenedetti and Stillinger’s model.</td>
<td>3</td>
</tr>
<tr>
<td>Figure 1.2</td>
<td>Phase transition of Al$<em>{90}$Sm$</em>{10}$ metallic glass when heated.</td>
<td>5</td>
</tr>
<tr>
<td>Figure 2.1</td>
<td>Flowchart of Monte Carlo simulation.</td>
<td>16</td>
</tr>
<tr>
<td>Figure 2.2</td>
<td>The periodic boundary condition in two dimensions.</td>
<td>17</td>
</tr>
<tr>
<td>Figure 2.3</td>
<td>The feasibility of the neighborlist method.</td>
<td>19</td>
</tr>
<tr>
<td>Figure 2.4</td>
<td>The GPU Devotes More Transistors to Data Processing.</td>
<td>21</td>
</tr>
<tr>
<td>Figure 2.5</td>
<td>$g(r)$ of supercooled Al$<em>{90}$Sm$</em>{10}$ liquid.</td>
<td>23</td>
</tr>
<tr>
<td>Figure 2.6</td>
<td>The classical nucleation theory model.</td>
<td>28</td>
</tr>
<tr>
<td>Figure 2.7</td>
<td>The free energy perturbation method.</td>
<td>29</td>
</tr>
<tr>
<td>Figure 3.1</td>
<td>The flowchart of our algorithm.</td>
<td>35</td>
</tr>
<tr>
<td>Figure 3.2</td>
<td>The integrand of Equation 3.1 for fcc Al at $T = 800$ K.</td>
<td>37</td>
</tr>
<tr>
<td>Figure 3.3</td>
<td>The solid-liquid interface velocity for pure Al in the [100] direction as a function of temperature.</td>
<td>38</td>
</tr>
<tr>
<td>Figure 3.4</td>
<td>Absolute Gibbs free energy of fcc and liquid Al as a function of temperature under the ambient pressure.</td>
<td>39</td>
</tr>
<tr>
<td>Figure 3.5</td>
<td>The integrand of Equation 3.7 for $x = 0.25$ and $T = 1500$ K.</td>
<td>41</td>
</tr>
<tr>
<td>Figure 3.6</td>
<td>Driving force for nucleation of fcc-Al in Al$<em>{1-x}$Sm$</em>{x}$ liquid at 700 K as a function of the Sm composition.</td>
<td>43</td>
</tr>
</tbody>
</table>
Figure 3.7 The formation free energy of $\text{Al}_{1-x}\text{Sm}_x$ liquid, referenced to solid phases fcc-Al and $\text{Al}_3\text{Sm}$, at $T = 880$ K.

Figure 3.8 Melting curve for fcc-Al and $\text{Al}_3\text{Sm}$ from experiments and the current calculations for $\text{Al}_{1-x}\text{Sm}_x$ in the Al-rich region of the Al-Sm system.

Figure 4.1 The $\epsilon$-$\text{Al}_{0.60}\text{Sm}_{1.11}$ phase.

Figure 4.2 Raw and processed HAADF-STEM image.

Figure 4.3 Intensity along the $\langle 110 \rangle$ crystal lines.

Figure 4.4 The averaged correlation function along $\langle 110 \rangle$ lines of STEM samples (blue), and randomly generated samples (red).

Figure 4.5 Comparison of the formation free energy calculated using CE model ($E_{CE}$) and FS potential ($E_{FS}$) for 851 $\text{Al}_{120+y}\text{Sm}_{22-y}$ structures, at (a) $T = 500$ K and (b) $T = 800$ K.

Figure 4.6 MC simulation results.

Figure 4.7 Averaged correlation function along $\langle 110 \rangle$ lines of the MC samples, at 500 K (blue) and at 800 K (red).

Figure 4.8 Atomic configurations during molecular dynamic (MD) simulations.

Figure 4.9 Averaged correlation function along $\langle 110 \rangle$ lines of the MD simulations, with $r_c = \infty$ (blue), $r_c = 4.2\text{Å}$ (red) and $r_c = 2.6\text{Å}$ (black).

Figure 4.10 The estimated fitting parameters $A$ and $\xi$, together with their standard errors represented by a rectangular box for all the samples studied in this chapter.

Figure 5.1 The persistent-embryo method.

Figure 5.2 The histogram of $\delta_m$.

Figure 5.3 The embryo of $\epsilon$-$\text{Al}_{0.60}\text{Sm}_{1.11}$ phase.

Figure 5.4 The PEM simulation of the $\epsilon$-$\text{Al}_{0.60}\text{Sm}_{1.11}$ nucleation.
| Figure 6.1 | The flowchart of EAM/FS energy and force computing algorithm. | 79 |
| Figure 6.2 | Dependence of the potential energy on temperature by HOOMD-blue and LAMMPS during the cooling process of Cu_{64.5}Zr_{35.5} systems. | 82 |
| Figure 6.3 | Pair correlation functions $g(r)$ of Ni$_3$Al alloy at 1700 K by HOOMD-blue and LAMMPS are in agreement within statistical accuracy. | 83 |
| Figure 6.4 | Computing time required to perform a single MD time-step (TRPT) as a function of the number of atoms in the system. | 86 |
| Figure 6.5 | The GPU/CPU performance ratio as a function of the number of atoms in the system. | 87 |
| Figure 6.6 | Computing time required to perform a single MD time-step as a function of the number of atoms for EAM and LJ systems. | 88 |
ACKNOWLEDGEMENTS

I would like to take this opportunity to express my thanks to those who helped me with various aspects of conducting research and the writing of this thesis.

First and foremost, Dr. Kai-Ming Ho for his guidance, patience and support throughout my Ph.D. research and the writing of this thesis. His foresight and sagacity have often enlightened me on solving the obstacles I met in my graduate research. Encouraged by his words, my motivations for completing the research have frequently been strengthened.

I would like to thank my committee members for their efforts and contributions to this work: I would like to thank Dr. Paul Durbin, who was also the coordinating professor of my minor program Interdisciplinary Graduate Minor in Applied Scientific Computation, knowledge that I’ve learned in this program are very valuable in my research; I would like to thank Dr. James W. Evans, who was also my lecturer of Mathematical Methods for the Physical Sciences course, he taught me a lot methods, which are used in this thesis including Monte Carlo and molecular dynamics simulations; I would like to thank Dr. Alex Travesset-Casas, his comments in free energy computing, his suggestions in the persistent-embryo method (PEM), and his tutorials of the CUDA programming language make Chapter. 3, 5 and 6 can be finished; I would like to thank Dr. James P. Vary, who was also my lecturer of Quantum Field Theory I, my view has been broadened by being taught the applications of QFT in particle physics and condensed matter physics.

I would like to express my special thanks to Dr. Cai-Zhuang Wang, Scientist II of Ames Laboratory, who is my unofficial but de facto co-supervisor. In the group meetings and in the daily researches, his instructions and ideas always make my study clearer. Besides, his
life style, especially the regular schedule every day, also inspires me to be a diligent and wholesome researcher.

I would also like to express my sincere thanks to my collaborators, Dr. Fan-Qiang Meng, Dr. Lin Zhou and Dr. Matthew J. Kramer for explaining to me about STEM image processing, Dr. Huajing Song and Dr. Mikhail I. Mendelev for discussing with me about nucleation theory and classical potential, Dr. Tao Wang and Dr. Xiaobao Lv for sharing their understandings of Sm in Al-Sm glass in a diffusion and network point of view, and Dr. Shi-Huai Zhou for teaching me tips in free energy computing.

Also very importantly, I would like to thank my research group members Dr. Feng Zhang, Dr. Yang Sun, Dr. Xin Zhao, Dr. Yongxin Yao, Dr. Zhuo Ye, Dr. Manh Cuong Nguyen, Dr. Jun Liu, Mr. Boqun Song, and Mr. Shang Ren. We are an excellent team and share our knowledge selflessly: the free-energy computing approach in Chapter 3 is based on Dr. Feng Zhang’s idea, the PEM method used in Chapter 5 is invented in collaboration with Dr. Yang Sun and Dr. Feng Zhang, the usage of the cluster expansion package ATAT used in Chapter 4 is taught by Dr. Xin Zhao.

In addition, I would like to thank my friends in life Dr. Fan Zhao, Dr. Ying Mao, and Dr. Gang Chen, without whose help I would not have been able to complete this work.

This work was supported by the U.S. Department of Energy (DOE), Basic Energy Sciences, Materials Science and Engineering Division at the Ames Laboratory under contract number DE-AC02-07CH11358. Ames Laboratory is operated for the DOE by Iowa State University. The document number assigned to this thesis/dissertation is IS-T 3236.
ABSTRACT

Different phases of metallic alloys have a wide range of applications. However, the driving mechanisms of the phase selections can be complex. For example, the detailed pathways of the phase transitions in the devitrification process still lack a comprehensive interpretation. So, the understanding of the driving mechanisms of the phase selections is very important.

In this thesis, we focus on the study of the Al-Sm and other related metallic alloy systems by simulation and experiment. A procedure to evaluate the free energy has been developed within the framework of thermodynamic integration, coupled with extensive GPU-accelerated molecular dynamics (MD) simulations; The "spatially-correlated site occupancy" has been observed and measured in the $\epsilon$-Al$_{60}$Sm$_{11}$ phase. Contrary to the common belief that nonstoichiometry is often the outcome of the interplay of enthalpy of formation and configurational entropy at finite temperatures, our results from Monte Carlo (MC) and molecular dynamics (MD) simulations, imply that kinetic effects, especially the limited diffusivity of Sm is crucial for the appearance of the observed spatial correlations in the nonstoichiometric $\epsilon$ phase. Moreover, in order to overcome the time limitation in MD simulation of the nucleation process, a “persistent-embryo method” has been developed, which opens a new avenue to study solidification under realistic experimental conditions via atomistic computer simulation. Based on this thesis study, we have achieved deeper understanding of the driving mechanisms of the phase selections, and laid a foundation for further prediction and control of the fabrication of novel metallic alloy materials. This thesis consists of the following seven chapters:
Chapter 1 briefly introduces the history of the development of the metallic alloy materials, and their significant impact on human civilization. In particular, the research background of metallic glasses has been reviewed, and some unsolved questions have been raised.

Chapter 2 is the literature review. In the first section of it, simulation methods, including molecular dynamics (MD) simulation, Monte Carlo method (MC) simulation, and related technical issues in the computer simulation to mimic the real system have been reviewed. In the second section of it, analysis methods, including structural and dynamical analysis methods, classical nucleation theory, free energy computing algorithms, and experimental techniques have been reviewed.

Chapter 3 reports our work in a self-contained procedure to evaluate the free energy of liquid and solid phases of an alloy system. We start from the Einstein crystal as the reference system, using thermodynamic integration to solve the free energy of a single-element solid phase. Then we compute the free energy difference between the solid and liquid phases using Gibbs-Duhem integration. After that we construct an “alchemical” path connecting a pure liquid and a liquid alloy to calculate the mixing enthalpy and entropy. This procedure is of great importance because the evaluation of free energy is fundamental to achieving microscopic understandings of freezing and melting phenomena.

Chapter 4 elucidates the origin of the spatially-correlated site occupancy in the non-stoichiometric meta-stable $\epsilon$-Al$_{60}$Sm$_{11}$ phase. This STEM observed spatially-correlated site occupancy cannot be explained by the “average crystal” description from Rietveld analysis of diffraction data, or by the lowest free energy structure established in MC simulations. MD simulations of the growth of $\epsilon$-Al$_{60}$Sm$_{11}$ in undercooled liquid show that when the diffusion range of Sm is limited to $\sim 4\,\text{Å}$, the correlation function of the as-grown crystal structure agrees well with that of the STEM images. Conclusion thus has been made that the kinetic effects, especially the limited diffusivity of Sm atoms plays an important role in determining the non-stoichiometric site occupancy. In addition to the free energy point of view, this result
helps us to have a deeper understanding of phase selections from structural and dynamical points of views.

Chapter 5 describes the “persistent-embryo” method (PEM) nucleation simulation. The PEM is developed to facilitate crystal nucleation in MD simulations by preventing small crystal embryos from melting using external spring forces, so that the early state of rare nucleation events can be accessed. This method opens a new avenue to study solidification under realistic experimental conditions. The nucleation rates of pure Ni, and of B2 phase of glass former Cu-Zr alloy have been computed using PEM. We also apply PEM to the Al-Sm system to study the nucleation of $\epsilon$-Al$_{60}$Sm$_{11}$ phase in the undercooled Al-Sm liquid, complex and interesting behaviors, different from the Ni case, have been found.

Chapter 6 presents an implementation of EAM and FS inter-atomic potentials in HOOMD-blue, a GPU software designed to perform classical molecular dynamic simulations. The accuracy of the code has been verified in a variety of broad tests, the performance of the code is significantly faster than LAMMPS running on a typical CPU cluster. Furthermore, our hoomd.metal module follows HOOMD-blue code convention, which allows it to be coupled to the extensive python libraries. This package makes the MD simulations in the thesis and other related fields faster and more convenient.

Chapter 7 is a summary of my Ph.D. thesis study, and proposes a plan for future works.

**Keywords:** phase selection, free energy calculations, non-stoichiometric compound, molecular dynamics, Monte Carlo simulation, GPU computing
CHAPTER 1. INTRODUCTION

Materials are essential to human survival and development. Early human made use of natural materials such as stone, wood, bone, skins, fiber, clay, sand, limestone, and minerals [1]. From those materials, people gradually learned to work particular types of natural materials into novel tools and man-made materials. This process was so important in the history of civilization, that the pre-historical time was divided and named after the characteristic materials: Stone Age, Bronze Age, and Iron Age (proposed by Danish antiquarian C.J. Thomsen [2]). In these pre-historical ages, the control of fire was a milestone, which made it possible to break big rocks, extract metals from ores, etc. When it comes to the modern and contemporary era, thousands to millions of materials have been found or invented, fabricated, and applied in all aspects of human life, such as metal, ceramic, polymer and composite materials. The processing technologies have been developed as well: we are able to apply heat, mechanical forces, chemical reactions, etc. to change the microscopic or even atomic structures of the materials, in order to change their properties. From this brief historical review, we can summarize human’s application of materials to a positive two-step feedback loop: we fabricate a material, understand the relation between its properties and the processing method, as a positive feedback, this understanding helps us to design and fabricate new materials. Interestingly, the fabrication step often goes first, for example, the glasses had been fabricated nearly 4,000 years ago, however, “the deepest and most interesting unsolved problem in solid state theory is probably the theory of the nature of glass and the glass transition” [3] (Philip Anderson, 1995). In this thesis study, we follow the same loop, and focus on the understanding part: we investigate the structures and dynamics of phase selection in the metallic alloy systems.
It is interesting that the metallic alloy and the glass can be “married”, and have a daughter named “metallic glass”. As its name suggests, the metallic glass consists of metal atoms, but has a glassy structure, possessing short-range to medium-range order, but being absent of long-range order. Actually, both metal and glass are among the earliest and most widely used materials in human history: the bronze age can be dated back to c. 3300 BC, the first application of volcanic glass is found in the stone age relics, and there had been a rapid growth in glass making industry in Egypt and Western Asia during the late bronze age [4]. However, the history of the metallic glass is very short. The first metallic glass was $\text{Au}_{75}\text{Si}_{25}$, fabricated and characterized by W. Klement (Jr.), Willens and Duwez in 1960 [5], their method was to rapidly cool the alloy (to the order of $10^8 K/s$) to avoid crystallization, so the samples had to be thin enough in one dimension to achieve such extreme cooling rate; in fact, the first $\text{Au}_{75}\text{Si}_{25}$ glass is only $20 \mu m$ in thickness. In the 1970s, Pond and Chen et al. invented the twin roll strip casting method and the centrifugal rapid cooling method [6,7]. In 1975, Chen et al. fabricated the $1m m \sim 3m m$ Pd-Cu-Si glass [8], which has a lower critical cooling rate. In the early 1980s, Turnbull et al. increased the metallic glasses’ thickness to the order of $\sim 1cm$, using Pd-Ni-P and Pt-based metallic glass [9,10]. In 1990s, Inoue at Tohoku University and Johnson at Caltech made great contributions to the field of metallic glass with the discovery of glass forming alloys with much lower critical cooling rate (for example, the critical cooling rate of Zr-Ti-Cu-Ni-Be alloy is lower than $1K/s$ [11]). Bulk metallic glasses had been fabricated (for example, Zr-Al-Ni-Cu glass’s thickness can reach $\sim 3cm$ [12]), making the mass production of metallic glasses possible. During the following years, multiple kinds of metallic glasses have been fabricated, including the Mg-based [13–17], Zr-based [18–26], Pd-based [27–29], Cu-based [30–36], Fe-based [37–47], Co-based [48–51], Ni-based [50, 52–55], Ti-based [56, 57], Rare earth-based (RE) [58–60] metallic glasses. Although the origin of glass and the glass transition is still under investigation, in principle, people have realized that the metallic glass is a thermodynamic meta-stable phase.
which is absent of long-range orders. Turnbull proved that given sufficiently fast cooling rate, any metallic melt is able to form glass [58]; Debenedetti and Stillinger studied the glass transition and devitrification process in detail, proposed their supercooled liquids and the glass transition theory [61, 62].

![Figure 1.1 Debenedetti and Stillinger’s model.](image)

Figure 1.1 is retrieved from Ref. [62], (a) shows the volume and enthalpy with respect to the temperature of a system during a cooling process, $T_m$ is the melting temperature of the crystal, with ideal cooling rate $r_0 \rightarrow 0 \text{K/s}$; $T_{ga}$ and $T_{gb}$ is the glass transition temperature with cooling rate $r_a$ and $r_b$ respectively, where $r_0 \ll r_a < r_b$. (b) is the illustration of the free energy profile, ideal glass and crystal phases are labeled. As shown in Figure. 1.1, in the cooling process, when the temperature is below the melting point, the diffusivities of the atoms slow down, the system need more time to “find” and “move” to the lower energy state, if the cooling is fast enough, system can be “frozen” in the meta-stable, glassy state, this is called the glass forming process. Reversely, if we heat the glass gradually, allowing the atoms in the glass system diffuse faster, system will have enough time to search for and transfer to the low energy crystalline phase, this is called the devitrification process.

Devitrification of metallic glasses typically involves passing through low-temperature regimes where multiple competing structures are thermodynamically possible yet dynamical
processes are limited \[63\]. As a result, the metallic glass phase transitions usually consist multiple steps, among which the meta-stable phases may appear (the phase is “selected”) rather than the more stable phases, depending on the processing conditions. Hence, the metallic glass systems can serve as prototypes for investigation of the driving mechanisms in the formation of these diverse and complex meta-stable phases.

Most discussions in this thesis are related to the Al-RE (RE: rare earth) alloys, many of which are glass formers. Al-RE systems often have rich phase selection phenomena. We are interested in Al, because it is very useful in industrial applications, for its low density, high malleability, perfect corrosion resistance, and good thermal and electrical conductivity. The properties of Al can be improved further by adding RE into it to form Metal-RE alloys. For example, Y (Yttrium) enhances strength of alloys \[64\], Ce (Cerium) nitrate provides even better corrosion coatings \[65\], Gd (Gadolinium) increases durability of alloys \[66\], ... So the study of Al-RE alloys is important both in physics and applications. In this thesis, we focuses on the Al-Sm (Sm: Samarium) alloy.

Figure 1.2 shows a typical phase transition pathways of Al$_{90}$Sm$_{10}$ metallic glass when heated, experiment were conducted in Ames Laboratory by Lin Zhou, Fanqiang Meng, and M. J. Kramer.

- The above part shows the phase transition of the Al$_{90}$Sm$_{10}$ glass prepared by melt-spinning ribbon (MSR) method, when heated, the phase transition path is:
  \[\text{Al}_{90}\text{Sm}_{10} \text{ metallic glass} \rightarrow \text{big cubic phase } \text{Al}_{60}\text{Sm}_{11}(\epsilon) \rightarrow \text{Al}^{(\text{fcc})} + \text{Al}_5\text{Sm}(\pi) + \text{Al}_4\text{Sm}(\beta) \rightarrow \text{Al}^{(\text{fcc})} + \text{Al}_4\text{Sm}(\gamma) + \text{Al}_4\text{Sm}(\beta)\]

- The bottom part shows the phase transition of the Al$_{90}$Sm$_{10}$ glass prepared by sputtered thin film (STF) method, when heated, the phase transition path is:
  \[\text{Al}_{90}\text{Sm}_{10} \text{ metallic glass} \rightarrow \text{Al}^{(\text{fcc})} + \text{Al}_20\text{Sm}_{4} \rightarrow \text{Al}^{(\text{fcc})} + \text{Al}_5\text{Sm}(\pi) + \text{Al}_4\text{Sm}(\gamma) + \text{Al}_4\text{Sm}(\beta) \rightarrow \text{Al}^{(\text{fcc})} + \text{Al}_4\text{Sm}(\gamma) + \text{Al}_4\text{Sm}(\beta)\]
As shown in Figure. 1.2, the phase transition of the Al$_{90}$Sm$_{10}$ can be very complex: it can undergo a variety of paths and transfer to novel phases. For example, the Al$_{90}$Sm$_{10}$ glass and the Al$_{60}$Sm$_{11}$(ϵ) crystal are meta-stable phases, and the Al$_{60}$Sm$_{11}$(ϵ) crystal is a non-stoichiometric phase.

![Figure 1.2 Phase transition of Al$_{90}$Sm$_{10}$ metallic glass when heated.](image)

People have performed extensive researches and had considerable achievements in the realm of metallic glass, liquid and solid; however, the driving mechanisms of the phase selection, and the detailed pathways of the devitrification are still not clear. Motivated
by these questions, this thesis study focuses on the Al-Sm system, and investigates the structures and dynamics in its phase transition. Through these investigations, we have proposed a procedure for the free energy evaluation, which is mostly from a structural (energy) point of view; also, we have elucidated the kinetic effect in the devitrified phase, which is from both structural and dynamical points of views; furthermore, the growth pattern from non-crystalline to crystalline phase in realistic conditions has been modeled; in addition, molecular dynamics software package has been developed.
CHAPTER 2. LITERATURE REVIEW

In this chapter, the literature on the existing tools for investigating the structures and dynamics of phase selection in metallic alloy systems is reviewed. The molecular simulation methods are reviewed in Section 2.1, the analysis methods of the simulated results are reviewed in Section 2.2.

2.1 Simulation Methods

The molecular simulation actually has very a long history, even before the modern computer era, researchers were able to model liquids with macroscopic spheres under mechanical interaction (sphere-sphere collision) of large assemblies [67,68]. The invention of electric computer offers a powerful tool for the molecular simulation, although the electric computers started as a computational machine employed to perform code deciphering and bomb design, they were soon found useful for other non-military applications too, including the numerical simulation of particles [69], the first simulation of dense liquids was believed to be carried out on the MANIAC computer at Los Alamos [70], using the Metropolis Monte Carlo (MC) method invented by Metropolis, Rosenbluth, Rosenbluth, Teller and Teller [71,72]. Thanks to the fast development of the computer science, more and more methods of molecular simulation have been developed accordingly. Nowadays, molecular simulation has been widely used in the theoretical and computing study in physics, chemistry, materials, and biology, especially in the areas, where few if any exact theoretical results are known, and process is much hindered by the lack of unambiguous tests to assess the quality of approximate theories [72].
In this thesis study, molecular dynamics (MD) simulation and Monte Carlo (MC) method simulation are the fundamental tools in preparing simulated “samples” of the metallic alloy systems, and putting algorithms into practice. In terms of implementation, the GPU accelerated parallel computing is adopted.

2.1.1 Molecular dynamics simulation

The molecular dynamics (MD) simulation algorithm was first proposed in 1957 [73], which was for the simulations of hard spheres, then it was further adopted to model the condensed matter system including solid, liquid, gas, etc. [74–76].

2.1.1.1 Basic algorithm

The basic algorithm for MD is still as the same as that in the 1950s, it is about computing the structures of a system which contains $N$ particles (atoms or molecules). At time $t_0$, the structure of the system is determined by the position of each particle $\vec{r}_i$ ($i = 1, 2, \ldots, N$), to compute the “molecular dynamics” of the system, it involves updating the structure to the time at $t_0 + \Delta t$, where $\Delta t \to 0$. The basic theorem which governs the update process is the Newtonian equation:

$$\frac{d^2\vec{r}_i}{dt^2} = \frac{\vec{F}_i}{m_i} \quad (2.1)$$

where $\vec{F}_i$ is the summation of the forces acting on the $i$th particle by the other $(N - 1)$ particles, $m_i$ is the mass of the $i$th particle, and $\vec{r}_i$ is the position vector of the $i$th particle. By solving Equation. 2.1 at each time step, the trace of each particle in the system can be solved, these results can then be analyzed using statistical and mechanical methods to get the physical properties of this system or particular particles.

In the implementation of Equation. 2.1, limited by the discrete nature of the computers, the particles’ traces are of course not solved analytically, instead of which, the finite difference
method (FDM) based on Taylor series expansion is adopted. One of the most straightforward
FDM was introduced by Verlet [77]: given a system structure at time $t$, after a time interval
$\Delta t$, the positions and dynamic properties (velocities, accelerations, higher-order derivatives
of positions on time) at time $t + \Delta t$ can be expressed using Taylor series expansion as:

$$\vec{r}(t + \Delta t) = \vec{r}(t) + \Delta t \vec{v}(t) + \frac{1}{2!} \Delta t^2 \vec{a}(t) + \frac{1}{3!} \Delta t^3 \frac{d^3 \vec{r}(t)}{dt^3} + \frac{1}{4!} \Delta t^4 \frac{d^4 \vec{r}(t)}{dt^4} + \ldots$$

$$\vec{v}(t + \Delta t) = \vec{v}(t) + \Delta t \vec{a}(t) + \frac{1}{2!} \Delta t^2 \frac{d^3 \vec{r}(t)}{dt^3} + \frac{1}{3!} \Delta t^3 \frac{d^4 \vec{r}(t)}{dt^4} + \ldots$$

$$\vec{a}(t + \Delta t) = \vec{a}(t) + \Delta t \frac{d^3 \vec{r}(t)}{dt^3} + \frac{1}{2!} \Delta t^2 \frac{d^4 \vec{r}(t)}{dt^4} + \ldots$$

where $\vec{r}$ is the position of the particle, $\vec{v} = \frac{d\vec{r}}{dt}$ is the velocity of the particle, and $\vec{a} = \frac{d^2 \vec{r}}{dt^2}$ is the acceleration of the particle, note that the subscript is omitted since each of the particles follows the same form of equation. Then replace $\Delta t$ in Equation. 2.2 by $-\Delta t$:

$$\vec{r}(t - \Delta t) = \vec{r}(t) - \Delta t \vec{v}(t) + \frac{1}{2!} \Delta t^2 \vec{a}(t) - \frac{1}{3!} \Delta t^3 \frac{d^3 \vec{r}(t)}{dt^3} + \frac{1}{4!} \Delta t^4 \frac{d^4 \vec{r}(t)}{dt^4} + \ldots$$

submit Equation. 2.2 and 2.3, and omit the higher order terms (by assuming that they are small values), the particle position at $t + \Delta t$ can be expressed as a function of position at $t$ and $t - \Delta t$ and the acceleration at $t$:

$$\vec{v}(t + \Delta t) = 2\vec{r}(t) - \vec{r}(t - \Delta t) + \Delta t^2 \vec{a}(t)$$

such that, the Verlet’s algorithm can solve for the particle position at $t + \Delta t$ without computing the velocity, the velocity can be solved in a separate run if needed using particle’s position data:

$$\vec{v}(t) = \frac{[\vec{r}(t + \Delta t) - \vec{r}(t - \Delta t)]}{2\Delta t}$$

Although the Verlet’s algorithm accelerates the computation speed, it can incur inaccuracy, one source of error is from the huge accuracy difference between the small value $\Delta t^2 \vec{a}(t)$ and the large value $2\vec{r}(t)$ and $\vec{r}(t - \Delta t)$. Improved algorithms encompass the Swope’s algo-
rithm [78], in which the velocity is included in the Verlet’s algorithm:

\[
\begin{align*}
\mathbf{r}(t + \Delta t) &= \mathbf{r}(t) + \Delta t \mathbf{v}(t) + \frac{1}{2} \Delta t^2 \mathbf{a}(t) \\
\mathbf{v}(t + \Delta t) &= \mathbf{v}(t) + \frac{1}{2} \Delta t \mathbf{a}(t) + \frac{1}{2} \Delta t \mathbf{a}(t + \Delta t)
\end{align*}
\] (2.6)

and the Beeman’s algorithm [79]:

\[
\begin{align*}
\mathbf{r}(t + \Delta t) &= \mathbf{r}(t) + \Delta t \mathbf{v}(t) + \frac{2}{3} \Delta t^2 \mathbf{a}(t) - \frac{1}{6} \Delta t^2 \mathbf{a}(t - \Delta t) \\
\mathbf{v}(t + \Delta t) &= \mathbf{v}(t) + \frac{7}{6} \Delta t \mathbf{a}(t) - \frac{1}{6} \Delta t \mathbf{a}(t - \Delta t)
\end{align*}
\] (2.7)

However, the accuracy of all the above algorithms is still limited by the errors introduced by omitting the higher Taylor series expansion terms. Gear proposed a correction algorithm to correct the values computed based on Taylor series expansion [80], the strategy is: first, the particle’s properties, including position \(\mathbf{r}^T\), velocity \(\mathbf{v}^T\), acceleration \(\mathbf{a}^T\), etc., at time \(t + \Delta t\) are computed using Taylor series expansion 2.2, where the superscript \(T\) indicates the results are computed by the algorithms based on Taylor series expansion; once these values at time \(t + \Delta t\) are solved, the inter-atomic potential can be used to solve the particle’s acceleration \(\mathbf{a}^P\) at \(t + \Delta\), where the superscript \(P\) indicates the results are from inter-atomic potential or are “predicted” values; finally, the difference of \(\mathbf{a}^P\) and \(\mathbf{a}^T\)

\[
\Delta \mathbf{a}^G(t + \Delta t) = \mathbf{a}^P(t + \Delta t) - \mathbf{a}^T(t + \Delta t)
\] (2.8)

is solved, where the superscript \(G\) indicates it is of Gear’s correction, which can be used to correct the particle’s properties at \(t + \Delta t\):

\[
\begin{align*}
\mathbf{r}^G(t + \Delta t) &= \mathbf{r}^T(t + \Delta t) + \frac{1}{6} \Delta \mathbf{a}^G(t + \Delta t) \frac{\Delta t^2}{2} \\
\mathbf{v}^G(t + \Delta t) &= \mathbf{v}^T(t + \Delta t) + \frac{5}{6} \Delta \mathbf{a}^G(t + \Delta t) \frac{\Delta t}{2} \\
\mathbf{a}^G(t + \Delta t) &= \mathbf{a}^T(t + \Delta t) + \Delta \mathbf{a}^G(t + \Delta t) \\
\frac{d^3 \mathbf{r}^G}{dt^3}(t + \Delta t) &= \frac{d^3 \mathbf{r}^T}{dt^3}(t + \Delta t) + \frac{1}{3} \Delta \mathbf{a}^G(t + \Delta t) \frac{3}{\Delta t}
\end{align*}
\] (2.9)
The advantages of Gear’s correction are its high accuracy, and for each time $t$, only one set of correlation values at $t + \Delta t$ is required to be stored, so it is very capable to be implemented in the modern computers. Such that, Verlet’s algorithm together with Gear’s correlation is widely used in the MD simulations.

2.1.1.2 inter-atomic potentials

Since MD simulation is governed by the classical Newton’s second law Equation. 2.1, it is premised on the modeling of the force $\vec{F}_i$ act on the $i$th particle. Depending on the systems and the motivations that the researchers are interested in, different force models have been proposed accordingly, instead of giving out the forces directly, most models describe the inter-atomic potentials $U$, and one just need to take the negative gradient of $U$ to compute the force $\vec{F}_i = -\nabla_{\vec{r}_i} U$, note that the term “inter-atomic” should not be literally interpreted, the inter-atomic potentials are potential functions of any particles (atoms and/or molecules), for example: the hard sphere potential which describes particle-particle interaction with the elastic collision mechanics [73]; the pair potentials including the Lennard-Jones potential [81], the Morse potential [82], and the Yukawa potential (also called the screened Coulomb potential) [83], which describe the potential energy as a function of the distance between two interacting particles; and the many body potentials including the embedded atom method (EAM) potential [84], the Finnis-Sinclair (FS) potential [85], and the ab-initio methods [86, 87], which consider many body terms in the potential functions’ forms. The inter-atomic potentials have found fruitful applications in describing forces between particles in solids, liquids, gases, glasses, membranes, emulsions, suspensions and their interfaces. In the metallic alloy systems simulation, the ab-initio method and the EAM/FS potential are widely used.

The ab-initio method is the most accurate model to describe the inter-atomic interactions, because just as its name implies (ab-initio means “from first principles” or “from
the beginning”), it is a method based on quantum mechanics, attempting to solve the electronic Schrödinger’s equation. The most popular and fundamental class of the ab-initio method is the density functional theory (DFT), proposed by Hohenber, Kohn and Sham in the 1960s [88–90]. Hohenber and Kohn proved that

- **Hohenberg-Kohn Theorem 1**: The external potential \( V_{\text{ext}}(\vec{r}) \) is a unique functional of the electron density \( \rho(\vec{r}) \). H-K theorem 1 implies that the total energy \( E \), the kinetic energy \( T \), and the electron-electron interaction potential \( V_{ee} \) can all be expressed as functional forms of electron density \( \rho(\vec{r}) \): \( E[\rho] \), \( T[\rho] \), and \( V_{ee}[\rho] \), respectively.

- **Hohenberg-Kohn Theorem 2**: Given the external potential \( V_{\text{ext}}(\vec{r}) \), the density that minimizes the total energy is the exact ground state density. H-K theorem 2 means that the ground state energy (density) can be solved variationally by minimizing energy functional:

\[
E_{HK}[\rho] = \langle \Psi | T + V_{ee} + V_{\text{ext}} | \Psi \rangle = T[\rho] + V_{ee}[\rho] + \int d^3r V_{\text{ext}}(\vec{r})\rho(\vec{r}) = F[\rho] + \int d^3r V_{\text{ext}}(\vec{r})\rho(\vec{r}) \tag{2.10}
\]

where \( F[\rho] = T[\rho] + V_{ee}[\rho] \) is called the universal functional.

Introducing the ground state density \( \rho \), which is related to the wave function basis:

\[
\rho(\vec{r}) = \sum_{i=1}^{N} \psi_i^\dagger(\vec{r})\psi_i(\vec{r}) \tag{2.11}
\]

Kohn and Sham substitute Equation. 2.11 to Equation. 2.10, and separate the universal functional \( F[\rho] \) into terms of \( T_s[\rho] \), \( E_H[\rho] \), and \( E_{XC}[\rho] \), such that \( F[\rho] = T_s[\rho] + E_H[\rho] + E_{XC}[\rho] \), where

\[
T_s[\rho] = -\frac{\hbar^2}{2m} \sum_{i=1}^{N} \langle \psi_i(\vec{r}) | \nabla^2 | \psi_i(\vec{r}) \rangle
\]
\[
\frac{\hbar^2}{2m} \sum_{i=1}^{N} \int d^3 r |\nabla \psi_i(\vec{r})|^2
\]  

(2.12)

is the kinetic energy functional of a non-interacting electron gas of density \( \rho(\vec{r}) \).

\[
E_H[\rho] = \frac{1}{2} \int \int dr dr' \rho(\vec{r})\rho(\vec{r}') \frac{e^2}{|\vec{r} - \vec{r}'|}
\]

(2.13)

is the classical electrostatic (Hartree) energy of the electrons, and \( E_{XC}[\rho] \) is the exchange-correlation energy functional, which describes the unknown many body effect apart from the \( T_s[\rho] \) and \( E_H[\rho] \) terms. The Hohenberg-Kohn functional can be rewritten to the Kohn-Sham functional:

\[
E_{KS}(\rho) = T_s[\rho] + E_H[\rho] + E_{XC}[\rho] + \int d^3 r V_{ext}(\vec{r})\rho(\vec{r})
\]

(2.14)

the ground state energy of a many-electron system can be obtained by minimizing the energy functional 2.14, solve the variational problem, we have:

\[
\epsilon_i \psi_i(\vec{r}) = \left[ -\frac{\hbar^2}{2m} \nabla^2 + V_{ext}(\vec{r}) + \frac{1}{2} \int \int dr dr' \rho(\vec{r})\rho(\vec{r}') \frac{e^2}{|\vec{r} - \vec{r}'|} + \frac{\delta E_{XC}[\rho]}{\delta \rho(\vec{r})} \right] \psi_i(\vec{r})
\]

(2.15)

Equation. 2.15 is called the Kohn-Sham equation, it can be solved self-consistently, by assuming an initial ansatz, and compute the new density, and iterate again and again until convergence is achieved. Note that although the Kohn-Sham equation is self-consistent, but the form of the exchange-correlation functional \( E_{XC}[\rho] \) is unknown, in practice, approximations for the exchange term \( E_{XC}[\rho] \) must be made, the most commonly used approximations include Local Density Approximation (LDA), and General Gradient Approximate (GGA) [91–94], the details of these approaches are not very relevant to this thesis, so the further discussions are omitted here, generally speaking, LDA assumes that \( E_{XC}[\rho] \) depends solely upon the value of the electronic density at each point in space, whereas GGA considers the gradient of the electron density as well.

Ab-initio method can be implemented within MD, using software packages including VASP [95], Abinit [96], and CPMD [97], which is very accurate, but it often incurs a high
computational cost. For example, it is almost impossible to simulate large systems (> 1,000 atoms) and long time (> 1 ns), limited by the speed in solving the Kohn-Sham equation. However, many phenomena in metallic alloy systems require simulations with considerable spatial and temporal scales. Fortunately, it is possible to include many quantum effects into a simpler classical or semi-classical potentials, from which energy and forces are more efficiently calculated. In this thesis, we use the embedded atom method (EAM) [84] and the Finnis-Sinclair (FS) model [85] potentials, EAM is a classical empirical model, the total potential energy is expressed as a sum of a pair-wise term and an embedding term that parametrizes many-body effects, whereas the the FS is a semi-classical model, derived from the tight-binding approach based on the second moment approximation. While independently developed, these two potentials share significant similarity in their formulation.

The total potential energy of EAM [84] and FS [85] potentials can be expressed as:

$$E_{tot} = \sum_i F_{\tau_i}(P_i) + \frac{1}{2} \sum_i \sum_{j \neq i} \phi_{\tau_i\tau_j}(r_{ij})$$  \hspace{1cm} (2.16)

where $r_{ij}$ is the distance between atoms $i$ and $j$, $P_i = \sum_{j \neq i} \rho_{\tau_i\tau_j}(r_{ij})$ is the total electron density at the location of atom $i$, in which $\tau_i$ is the type of atom $i$, $\rho_{\tau_i\tau_j}(r_{ij})$ is the contribution to the electron density at the location of atom $i$ from atom $j$, the embedding function $F_{\tau_i}(P_i)$ represents the energy required to “embedded” atom $i$ into the electron cloud, and the pair function $\phi_{\tau_i\tau_j}(r_{ij})$ represents the pair-wise potential energy. For the “EAM” potential, the electron charge density $\rho_{\tau_i\tau_j}(r_{ij})$ depends only on atom $j$’s type $\tau_j$, and thus can be reduced to $\rho_{\tau_j}(r_{ij})$. While within the “FS” potential, $\rho_{\tau_i\tau_j}(r_{ij})$ generally depends on both $\tau_i$ and $\tau_j$.

EAM/FS potential can be implemented within MD, using software packages including LAMMPS [98], HOOMD-blue [99,100], and GROMACS [101], the simulation efficiency can be significantly improved compared with ab-initio based software packages. A common strategy is to use data from the ab-initio simulation, experiment and other resources to develop a reasonably consistent EAM/FS potential, then apply the EAM/FS potential in
the MD simulation, such that both the accuracy and efficiency of the simulation can be achieved.

2.1.2 Monte Carlo simulation

Invented and named by Metropolis and Ulam [71], the Monte Carlo (MC) method has been used not only in the molecular simulation, but also in the realms of numerical integration, optimization, artificial intelligence, risk analysis, . . . . The extensive applications of MC method are based on its stochastic nature, in principle, MC method can be used to solve any problem having a probabilistic interpretation by modeling the probability distribution. It should be mentioned that the MC molecular simulation is just one application of the MC method in molecular dynamics problems, by modeling the proper probabilities in statistical mechanics. The most “classical” MC method is called the “importance sampling”, in each step of the simulation, two evolution rules must be applied: the ergodic rule and the detailed balance rule.

- **Ergodic rule.** An ergodic process in a system is such a process that the statistical properties of all the points in the system space, can be represented by a reasonably large selection of points within a given enough time. The design of any MC method must obey ergodic rule in order to deduce properties of the whole space from a selection of the space.

- **Detailed balance.** Let $P_0, P_1, \ldots, P_i$ be Markov process with stationary distribution $\pi_i$, the process satisfy detailed balance if

$$\pi_i p_{i \to j} = \pi_j p_{j \to i} \forall i, j$$  \hspace{1cm} (2.17)

where $p_{i \to j}$ is the probability of state $P_i$ to $P_j$, and $p_{j \to i}$ is vice versa. Most MC molecular simulation algorithms, including the Metropolis MC, obey the detailed balance
rule, in order to insure the sampling scheme leads to the equilibrium state. However, it need to mention that the detailed balance rule is a sufficient but not a necessary condition, some alternative schemes can also lead to equilibrium [102].

The Metropolis MC molecular simulation algorithm is constructed to model the Boltzmann distribution of a system, a flowchart of it is shown in Figure 2.1.

![Flowchart of Monte Carlo simulation.](image)

There are a variety of algorithms to modify the structure to $R'$. In molecular simulation, the Kawasaki algorithm [103] is often used, Kawasaki algorithm swaps the two particles instead of modifying one by one, which was started to be used in the Ising model in order to maintain the spin momentum of the system, in molecular simulation, the swap strategy maintains the concentration of the alloy. Cluster MC algorithm [104, 105] is another interesting algorithm, in which two clusters (regions) are swapped, which can be regarded as a generalized Kawasaki algorithm, but much faster. In this thesis, the ATAT software package [106] is used to perform the MC simulation.

2.1.3 Tricks of computer simulation

Some tricks must be played in the computer simulation, at least for two important reasons: one is to mimic the “real” system since the simulation is never a “real” experiment, the other
is to accelerate the computing speed. The former includes the periodic boundary condition, and the Nosé-Hoover thermostat, the latter includes the concept of the neighborlist, and the parallel computing.

### 2.1.3.1 periodic boundary condition

Comparing to most of the “real” experiments, the number of particles in a molecular simulation is often much smaller, particles which are close to the walls of the simulation box, or in experimental words, on the surface, would experience very different forces compared to the particles in the box (bulk particles) would do. If we are interested in simulating the properties of the bulk particles, periodic boundary condition must be applied [107]. The periodic boundary condition is illustrated in Figure 2.2.

![Figure 2.2](image)

Figure 2.2 The periodic boundary condition in two dimensions.
For simplicity, only two dimension is shown, particles leave and enter boxes through the four walls. The basic idea of the periodic boundary condition is: as a particle (particles with dark orange background) moves in the simulation box (shaded in dark blue), the periodic image of it (particles with light orange background) in the neighboring image boxes (shaded in light blue) also moves in the same way, such that when the particle leaves the simulation box (particle 1 with dark orange background), its image (particle 1 with light orange background) enters the simulation box from the neighboring image box through the opposite wall. Such that the environment (inter-atomic potentials) of the particles close to the wall is as the same as that of bulk particles, and the total number of particles keeps the same.

2.1.3.2 the Nosé-Hoover thermostat

The canonical ensemble is often used in the molecular simulation, which is also called the $NVT$ ensemble in simulation for it is with constant number of particles, volume, and temperature. Most real process can be modeled in the $NVT$ ensemble instead of the micro-canonical ensemble ($NVE$), because temperature is a measurable value from the experiment. In order to set and keep the temperature in the simulation, thermostat must be introduces, the Nosé-Hoover thermostat [108, 109] is widely used in molecular simulation. The basic idea is to introduce a fictitious variable $\zeta$ to the Newtonian equation. 2.1 which governs the dynamics of the particles:

$$\frac{d^2 \vec{r}_i}{dt^2} = \frac{\vec{F}_i}{m_i} - \zeta \frac{d\vec{r}_i}{dt}$$ (2.18)

effectively, the variable $\zeta$ represents a dumping force, which tunes the acceleration of the particles, $\zeta$ is related to the temperature by:

$$\frac{d\zeta(t)}{dt} = \frac{1}{Q} \left[ \sum_{i=1}^{N} m_i \frac{\vec{v}_i^2}{2} - \frac{3N + 1}{2} k_B T \right]$$ (2.19)

where $Q$ is the Nosé-Hoover mass which determines relaxation of the dynamics of the dumping, $k_B$ is the Boltzmann constant, and $T$ denotes the target temperature. It can be seen
that with this Nosé-Hoover approach, the system tends to the steady temperature $T$, in which $\sum_{i=1}^{N} m_{i} \frac{\vec{v}_{i}^{2}}{2} = \frac{3N+1}{2} k_{B} T$ (3-dimension, the +1 is because the additional degree of freedom $\zeta$). It is necessary to emphasize that Nosé-Hoover thermostat does not fix the system’s temperature, but allows some fluctuations, which is very similar to the experiment’s scenario.

2.1.3.3 neighborlist

![Diagram of neighborlist method]

Figure 2.3 The feasibility of the neighborlist method.

No matter which inter-atomic potential is used, the time taken to compute the forces on the $N$ particles in the system is $O(N^2)$ if we execute a double loop. To reduce the computational load, Verlet [76] suggested an approach by recording a list of the neighbor particles of a particular $i$th particle, which is updated periodically every a few simulation time steps, this method can reduce the time taken to compute the forces to $O(N)$. The basic idea of the Verlet neighborlist is to surround the potential cut-off radius $r_c$ by a “skin” with width $r_s$, particles whose $\vec{r}_{j}$ follow $|\vec{r}_{j} - \vec{r}_{i}| < r_{c} + r_{s}$ is in the neighborlist of particle.
i, the neighborlist data is stored in an array and kept until the next update is needed. The feasibility of the neighborlist method is shown in Figure. 2.3, the skin width $r_s$ is chosen reasonably large such that between neighborlist updates, a particle (for example, particle 5 in Figure. 2.3) which is not in the list of particle $i$, will not move through the skin area into the $r_c$ sphere.

In the implementations, faster neighborlist store and indexing methods are developed, including the cell index [110,111], the stenciled cell list [112], and the LBVH (linear bounding volume hierarchies) tree indexing [113]. Note that although most algorithms can detect the “dangerous” construction of neighborlist (for example, particle 5 moves into the $r_c$ sphere) and update the neighborlist automatically, it is still important to set the update interval $\delta t$ and skin width $r_s$ reasonably to reduce the number of “dangerous” constructions to save computational time. For example, if the $\delta t$ is large, the $r_s$ should be set relatively larger accordingly, and vice versa.

2.1.3.4 parallel computing

Most of the simulation algorithms or their variations can be executed in a parallel fashion, and in the implementation realm, parallel computing has became an important area of research in computer architectures and software systems. Simulation algorithms can be greatly accelerated using parallel computing techniques. The basic idea of parallel computing is upon the Amdahl’s law [114]:

$$S = \frac{T_s + T_p}{T_s + T_p/N}$$

(2.20)

where $S$ is the “speedup” of the whole computational task, $T_s$ is the portion of the computational task which cannot be parallelized, $T_p$ is the parallelizable portion, and $N$ is the number of processors. It is clear to see, in order to increase the speedup $S$, we may increase $N$ and/or wisely design the algorithm such that $T_p \gg T_s$. 
The hardware architectures to conduct parallel computing include the multi-core processor, in which two (dual), four (quad) or more processing units are assembled together, which is now commonly seen in personal computers; the computer cluster, which combines sets of computers together to a “node”, modern cluster can contain tens to hundreds of nodes; and the general-purpose computing on graphics processing units (GPGPU), compared to the multi-core processor, a GPU contains more GPU “cores”, but they have lower clock rate, simpler instruction and less cache memory, as illustrated in Figure 2.4, this figure is retrieved from CUDA toolkit documentation of NVIDIA on Feb. 15, 2018 [115], the orange blocks represent the dynamic random-access memory (DRAM) and the Cache memories, the yellow blocks represent the control units, and the green blocks represent the arithmetic logic units. GPGPU is very suitable for the data extensive (millions of particles) but simple arithmetic operation (Newton’s law/MC algorithm) computations such as MD or MC. In this thesis study, all the simulations are performed in GPU clusters.

![Figure 2.4 The GPU Devotes More Transistors to Data Processing.](image)

### 2.2 Analysis Methods

The raw data from the molecular simulation is nothing more than the kinetic and thermodynamic values of the particles as a function of time, such as the positions, velocities,
accelerations, pressure, temperature (kinetic energy), potential energy, etc. Analysis is required for understanding the structures and dynamics in the phase selection.

2.2.1 Structural analysis

This subsection describes analysis methods from the structural point of view.

2.2.1.1 pair correlation function

The pair correlation function \( g_{\alpha\beta}(r) \) (or \( g(r) \) for simplicity) is used to describe the distance correlations between each two particles (pair), the physical interpretation of \( g(r) \) is the probability of finding a pair of particles with distance \( r \) apart, relative to the probability expected for a complete random distribution at the same density [116]. The pair correlation function for binary system \( \alpha-\beta \) is defined as:

\[
g_{\alpha\beta}(r) = \frac{N}{4\pi r^2 \rho N_{\alpha} N_{\beta}} \langle \sum_{i=1}^{N_{\alpha}} \sum_{j=1}^{N_{\beta}} \delta (r - r_{ij}) \rangle \tag{2.21}
\]

where \( N \), \( N_{\alpha} \) and \( N_{\beta} \) are the number of total particles, \( \alpha \) particles and \( \beta \) particles, respectively, \( \rho \) is the particle density, \( r_{ij} \) is the distance from \( i \)th particle to \( j \)th particle. Thus \( g(r) \) provides the measurement of structure ordering of a system.

Figure. 2.5 is a typical pair correlation function, the \( g(r) \) of Al-Sm supercooled liquid at 800 K, \( r \)'s regions below the first peak describes the short-range order, regions from the first peak further to \( 8 \sim 9\text{Å} \) is a description of medium-range order, it is clearly shown that as \( r \) goes further to \( \sim 12\text{Å} \), the pair correlation function converge to 1, implying that in this supercooled liquid, there is no long-range order.

What’s more, we may transfer the real space information \( g_{\alpha\beta}(r) \) to the reciprocal space by performing the Fourier transformation, to compute the partial structure factor \( S_{\alpha\beta}(q) \):

\[
S_{\alpha\beta}(q) = 1 + \frac{4\pi \rho}{q} \int_{0}^{\infty} r \left[ g_{\alpha\beta}(r) - 1 \right] \sin(qr) dr \tag{2.22}
\]
In the experiment characterization, for example, the X-ray diffraction (discussed in Sub-section 2.2.5), the measurable value is whole components’ structure factor $S(q)$. In the simulation, $S(q)$ can be get by summing up the $S_{\alpha\beta}(q)$ weighted by the concentrations $c_\alpha$ ($c_\beta$) and the X-ray form factors $f_\alpha$ ($f_\beta$) of each particle type:

$$S(q) = \frac{\sum_{i,j} c_i c_j f_i(q) f_j(q) S_{\alpha\beta}(q)}{\sum_{i,j} c_i c_j f_i(q) f_j(q)}$$

(2.23)

hence, the structure factor $S(q)$ can bridge the experimental characterizations and the simulated results.

![Figure 2.5](image)

**Figure 2.5**  $g(r)$ of supercooled Al$_{90}$Sm$_{10}$ liquid.

### 2.2.1.2 bond orientation order

Apart from the two-body function such as pair correlation function $g(r)$, functions to describe many-body structure orders have also been developed, one widely used order is the
bond orientation order (BOO) [117, 118], which quantitatively measures the symmetry of a center particle and all the neighboring particles around it, defined as the \( q_l \) factor of particle \( i \):

\[
q_{lm}(i) = \frac{1}{N_i} \sum_{j=1}^{N_i} Y_{lm}(\theta(\vec{r}_{ij}), \phi(\vec{r}_{ij}))
\]

\[
q_l(i) = \sqrt{\frac{4\pi}{2l+1} \sum_{m=-l}^{l} |q_{lm}(i)|^2}
\] (2.24)

where \( N_i \) is the number of neighboring particles around particle \( i \), \( Y_{lm} \) is the spherical harmonics, \( \theta(\vec{r}_{ij}) \) is the inclination (polar angle) of the particle \( j \) to particle \( i \), and \( \phi(\vec{r}_{ij}) \) is the azimuth (azimuthal angle) of them. The value of \( l \) is related to the spacial symmetry of the system, for example, \( l = 4 \) measures the cubic symmetry, \( l = 6 \) measures the hexagonal symmetry. The \([q_4, q_6]\) parameter is often used to measure and check the structure ordering of a system. For example, the \([q_4, q_6]\) for FCC polyhedron is \([0.191, 0.575]\), that for icosahedron is \([0.000, 0.663]\). The \([q_4, q_6]\) values for local structures in non-crystal can also be solved, and compared with that of the known polyhedrons, in order to describe the structure ordering.

### 2.2.1.3 Cluster alignment score

Atomic cluster alignment method is developed by XW. Fang, Y. Sun, et al. [119, 120], aiming at characterizing local structure ordering of non-crystal systems including metallic glasses and/or liquids. It is a data mining algorithm, quantifying the similarities between any pair of atomic clusters, the alignment score is defined as:

\[
S_{ij} = \min_{0.8<\alpha<1.2} \left( \frac{1}{N} \sum_{k=1}^{N} \frac{(\vec{r}_{ik} - \alpha \vec{r}_{jk})^2}{(\alpha \vec{r}_{jk})^2} \right)^{1/2}
\] (2.25)

where \( N \) is the coordination number of the two clusters, \( \vec{r}_{ik} \) is the relative position of the \( k \)th atom in the \( i \)th cluster to the center of the \( i \)th cluster, \( \vec{r}_{jk} \) is the relative position of the \( j \)th atom in the \( j \)th cluster to the center of the \( j \)th cluster, \( \alpha \) is a tuning parameter for extending
or shrinking the bond length. The alignment score $S_{ij}$ measures the similarity of the two clusters. In the applications, the alignment score method can be applied in two ways: one is the collective alignment method, which performs the alignment algorithm between each two clusters in a system, this can be regarded as an unsupervised learning, aiming at finding the common structure features in a system; the other is the cluster-template alignment, which aligns the clusters in a system to a template cluster, this can be regarded as a supervised learning, aiming at evaluating the similarities of the structures in a system and a template cluster.

2.2.2 Dynamical analysis

This subsection describes analysis methods from the dynamical point of view.

2.2.2.1 particle trajectories

The output of the MD and/or MC simulations is the snapshots of particles in a simulation box at each time step, with their information including positions, velocities, accelerations, etc. The time step must be chosen wisely, if the time step is too small, the simulation time would be too short to observe interested phenomena, or it would cost very long time to observe it, on the contrary, if the time step is too large, the simulation would be unrealistic. Neither too large nor too small time step could give out realistic particle trajectories. For metallic system simulation, the time step are generally set at $1 \sim 10 fs$ ($10^{-14} \sim 10^{-15}s$), in the thesis study, we use $2 \sim 3 fs$ as the time step.

Once we have got the snapshots, visualization software packages can be very helpful in interpreting the data. In this thesis, VESTA (Visualization for Electronic and Structural Analysis) [121], VMD (Visual Molecular Dynamics) [122] and OVITO (Open Visualization Tool) [123] are used. VESTA is maintained at Japan National Museum of Nature and Science by Momma et al., VMD is maintained at University of Illinois at Urbana-Champaign
by Schulten et al., and OVITO is maintained at Technische Universität Darmstadt by Stukowski. The packages can visualize the three-dimensional atomic configurations, analyze the structural properties, what is more, VMD and OVITO can output the trajectories and analyze the dynamic properties, such that we can have a better understanding of the simulated results.

2.2.2.2 dynamical order parameter

Apart from the structural order parameters we described in Subsection 2.2.1, dynamical order parameters need to be defined, for that in the complicated devitrification process, phase selection is dependent on relationships between thermodynamic and kinetic factors, such as local chemical partitioning and atomic diffusion. Based on the different atom motion in solid and liquid, Yang Sun et al. proposed the dynamical order parameter [124,125]:

\[
\delta = |\vec{r}(t) - \vec{r}(t - \Delta t)|
\]  

(2.26)

where \(\vec{r}(t)\) and \(\vec{r}(t - \Delta t)\) are the position vectors of each atom at time \(t\) and \(t - \Delta t\), respectively. The dynamical order parameter is essentially a measurement of the displacement that an atom moves within a time interval \(\Delta t\). For more complicated systems, we may modify the definition of \(\delta\) to fit our needs, which is discussed further in Chapter 5.

2.2.3 Nucleation process modeling

The nucleation process in the liquid phase should be carefully studied, because 1) the supercooled liquid is the parent phase of the glass, the devitrification process of the glass phase can be decomposed to stages of nucleation and growth; 2) although classical nucleation model can be too crude to explain experimental or simulated results in some cases, but it grasps the important factors in the nucleation process, hence it is still valuable to review the classical nucleation theory first.
The classical nucleation theory (CNT) is a nucleation model in a Gibbs free energy point of view [126, 127], CNT models the phase transition between disordered phase (liquid) and the ordered phase (solid), by decomposing the Gibbs free energy change in this process to two terms: one is the free energy decrease from the disordered phase to the ordered phase, the other is the free energy increase due to the interface between the disordered and ordered phases. The rival effect leads a Gibbs free energy barrier, as schematically shown in Figure. 2.6, assuming the nuclei is sphere shaped with radius $r$, according to CNT, the free energy change is:

$$\Delta G(r) = \frac{4}{3}\pi r^3 \rho_s \Delta \mu + 4\pi r^2 \gamma$$  \hspace{1cm} (2.27)

where $\rho_s$ is the atomic density of the ordered phase (solid), $\Delta \mu$ is the chemical potential change from the disordered phase to the ordered phase, and $\gamma$ is the free energy density in the disordered-ordered interface. The free energy barrier $\Delta G^*$ and critical nuclei radius $r^*$ of the nucleation can be solved by taking the first derivative of $\Delta G$ and set it to 0, which are:

$$\Delta G^* = \frac{16\pi \gamma^3}{3(\rho_s \Delta \mu)^2}$$

$$r^* = \frac{2\gamma}{\rho_s |\Delta \mu|}$$  \hspace{1cm} (2.28)

the nuclei has to be large enough $r > r^*$ (to overcome the barrier $\Delta G^*$) to grow, otherwise it will shrink (melt).

Based on CNT, steady-state kinetic model [128, 129] is developed to calculate the nucleation rate, generally speaking, this model introduces an attachment rate $\omega^*$ of the atoms attaching to a cluster, such that the nucleation rate is a function of the attachment rate and the free energy barrier:

$$J(T) = \omega^* \Gamma Z e^{-\frac{\Delta G^*}{k_B T}}$$  \hspace{1cm} (2.29)

where $Z$ is the particle density in the steady state, $\omega^*$ is the attachment rate, and $\Gamma$ is the Zeldorich factor, these factors behave differently depending on the systems [130, 131].
Although the CNT model is easy to use, and is valid in some simple cases, however when it comes to details, the CNT model is too rough to explain complex nucleation process, critics upon CNT and improved models are proposed during the last hundred years [132–140], the structural and dynamical ordering study in the early stage of nucleation is also an interesting question [141,142].

![Figure 2.6 The classical nucleation theory model.](image)

### 2.2.4 Free energy computing

As discussed in the previous subsection (2.2.3), the free energy plays a fundamental or even crucial role in understanding phase selection, however, the free energy is difficult to calculate directly using a simple average of a physical quantity over the phase space. The origin of the difficulty is that both Metropolis MC or MD generate trajectories that are following the Boltzmann-like distributions, such that the sampling over higher-energy configurations is not enough, in other words, the sampling is not converge. However, the
free energy difference between system $X$ and $Y$ are much more tractable, and actually it is the difference which is more important to phase selection. In this subsection, we describe the methods for evaluating Helmholtz free energy difference $\Delta A$ for simplicity, those for Gibbs free energy difference $\Delta G$ can be easily derived upon this.

### 2.2.4.1 free energy perturbation

![Figure 2.7 The free energy perturbation method.](image)

The free energy is a state function, so the free energy difference $\Delta A$ between the reference state $X$ to the target state $Y$ is:

$$
\Delta A_{X \rightarrow Y} = -k_B T \ln \left( \frac{\int \int dp^N dq^N e^{-\beta H_Y(p^N,q^N)}}{\int \int dp^N dq^N e^{-\beta H_X(p^N,q^N)}} \right) \quad (2.30)
$$

where $k_B$ is the Boltzmann constant, $\beta = \frac{1}{k_B T}$, $H_X$ and $H_Y$ are the Hamiltonian of state $X$ and $Y$ respectively. If the sampling is converge, Equation. 2.30 can be evaluated with an ensemble average:

$$
\Delta A_{X \rightarrow Y} = -\beta^{-1} \ln \langle e^{-\beta [H_Y(p,q) - H_X(p,q) - H_Y(p,q) - H_X(p,q)]} \rangle_X \quad (2.31)
$$
where the $\langle \ldots \rangle_X$ denotes an ensemble average over configurations representative of the reference state $X$. Note that this equation is free of error only if the sampling of $X$ covers sampling of $Y$, but in most cases, they do not overlap at all, as shown in Figure. 2.7 (a).

The free energy perturbation method [143–151] introduces $N$ intermediate states to connect reference state $X$ to the target state $Y$, so effectively, each intermediate state is a small “perturbation” to its reference state, forming a pathway characterized by the general extent parameter $\lambda$ [143, 148, 152, 153], as shown in Figure. 2.7 (b). The overlapping of neighboring intermediate states is very significant, hence the free energy difference can be solved by summing up the free energy sub differences along the pathway:

$$\Delta A_{X \rightarrow Y} = -\beta^{-1} \sum_{i=1}^{N} \ln \left\langle e^{-\beta \left[ H_{\lambda_{i+1}}(p,q) - H_{\lambda_i}(p,q) \right]} \right\rangle_{i}$$

(2.32)

### 2.2.4.2 thermodynamic integration

Instead of evaluating the free energy differences along the intermediate states, we can alternatively parameterize the Hamiltonian itself with a general extent parameter $\lambda$, to connect the reference state $X$ and the target state $Y$, this strategy is called the thermodynamic integration (TI) [72, 153, 154]. Start from the definition of the free energy:

$$A = -\beta^{-1} \ln \left( \frac{1}{N!h^{3N}} \int \int dp^N dq^N e^{-\frac{H(p^N,q^N)}{k_BT}} \right)$$

(2.33)

To simplify the following discussion, let $Q = \frac{1}{N!h^{3N}} \int \int dp^N dq^N e^{-\frac{H(p^N,q^N)}{k_BT}}$, since $H$ is parameterized, $Q$ and $A$ are all functions of $\lambda$, we can take the derivative of $A$ with respect to $\lambda$:

$$\frac{\partial A(\lambda)}{\partial \lambda} = \frac{1}{N!h^{3N}Q(\lambda)} \int \int dp^N dq^N \frac{\partial H(p^N,q^N)}{\partial \lambda} e^{-\frac{H(p^N,q^N)}{k_BT}}$$

$$= \int \int dp^N dq^N \frac{\partial H(p^N,q^N)}{\partial \lambda} \cdot \left( e^{-\frac{H(p^N,q^N)}{k_BT}} \right) \frac{1}{Q(\lambda)}$$

$$= \left\langle \frac{\partial H(p^N,q^N)}{\partial \lambda} \right\rangle_{\lambda}$$

(2.34)
hence, the free energy difference $\Delta A_{X \rightarrow Y}$ can be find by integrating Equation. 2.34:

$$\Delta A_{X \rightarrow Y} = \int_{\lambda_0}^{\lambda_1} \langle \frac{\partial H(p^X, q^X)}{\partial \lambda} \rangle_{\lambda} d\lambda$$

(2.35)

In particular, the intermediate parameterized Hamiltonian can be constructed:

$$H(\lambda) = (1 - \lambda)U_X + \lambda U_Y$$

(2.36)

where $U_X$ and $U_Y$ are the inter-atomic potential of system $X$ and $Y$, respectively, Equation. 2.35 has a simpler form:

$$\Delta A_{X \rightarrow Y} = \int_{0}^{1} (U_Y - U_X)_{\lambda} d\lambda$$

(2.37)

TI algorithm is constructed based on Equation. 2.37. In the application of metallic alloys, system $X$ and $Y$ may have different concentrations, the TI need to be modified to consider the atom replacement operations, which is discussed in detail in Chapter. 3.

### 2.2.5 Experimental techniques

Although we focus on the simulation and theoretical study, it is also necessary to briefly describe the experimental characterization methods of the phases, since it is the first step to analyze any phase. In this thesis work, the X-ray diffraction (XRD) and the scanning transmission electron microscope (STEM) are used. The theory of XRD is Bragg’s law for any wave:

$$2dsin\theta = n\lambda \quad n \in \mathbb{Z}$$

(2.38)

where $d$ is the spacing between diffracting planes, $\theta$ is the incident angle, and $\lambda$ is the wavelength of the beam. The beam source of XRD is the electromagnetic wave: X-ray, whose wavelength $\lambda$ is much smaller than visible light, to $\lambda = 1 \sim 100\,\text{Å}$, which is compatible to the lattice spacing of materials. By studying the diffraction pattern, we are able to characterize the structures of materials, for example, study the differaction patter for crystal, or the
structure factor $S(q)$ for non-crystal. Since the invention of XRD technique in the early 20th century [155], the XRD characterization has become a very powerful tool in material analysis.

Thanks to the wave-particle duality nature of electrons, the electrons can also been used as the beam source, and since the wavelength of electrons are even much smaller, we are able to “cast light (beam) on” the atoms of the specimen directly, and detect the scattering, diffracted, reflected, transmitted and generated electrons. Scanning electron microscope (SEM) detects the secondary electrons emitted from atoms excited by the electron beam; transmission electron microscope (TEM) and scanning transmission electron microscope (STEM) detects transmitted electron beams (bright field), or the diffracted electron beams (dark field), or the scattered electron beams (high angle dark field); whereas scanning tunneling microscope (STM) is a little different because it makes use of the quantum tunneling effect. In this thesis study, high angle annular dark field (HAADF) imaging of STEM is used, HAADF has a high resolution which is strongly dependent on the atomic number, by ruling out the diffraction electrons, collecting only the scattering electrons from Rutherford scattering [156–158].
CHAPTER 3. A PROCEDURE TO EVALUATE THE FREE ENERGY OF LIQUID AND SOLID PHASES OF AN ALLOY SYSTEM

In this chapter, the work of a self-contained procedure to evaluate the free energy of liquid and solid phases of an alloy system is reported. The free energy of a single-element solid phase is calculated with thermodynamic integration using the Einstein crystal as the reference system. Then, free energy difference between the solid and liquid phases is calculated by Gibbs-Duhem integration. The central part of our method is the construction of a reversible “alchemical” path connecting a pure liquid and a liquid alloy to calculate the mixing enthalpy and entropy. We have applied the method to calculate the free energy of solid and liquid phases in the Al-Sm system. The driving force for fcc-Al nucleation in Al-Sm liquid and the melting curve for fcc-Al and Al$_3$Sm are also calculated. This work is valuable to the thesis study because an accurate evaluation of the free energy is the foundation in understanding the phase selections.

3.1 Introduction

The free energy is the key parameter to determine the phase selections, it is fundamental to achieving microscopic understandings of freezing and melting phenomena [159]. However, reliable free energy evaluation, especially for the alloy liquid, remains a significant challenge in condensed matter physics and material science. The origin of the difficulty in free energy calculation is that it cannot be expressed as a simple average of a physical quantity over the phase space, and thus cannot be evaluated in a single simulation with a
standard sampling technique, such as Monte Carlo (MC) or molecular dynamics (MD) simulations [72]. A variety of methods have been proposed to compute the free energy, which almost all aims at computing the free energy difference between the reference system whose free energy is known and the target system whose free energy is unknown. One probability most widely used approach is the free-energy perturbation (FEP) [151], which can be encompassed in the applications of Jarzynski equation [160]. Although these methods are basically simple to apply, they often incur inaccuracy if the path connecting the reference system to the target system is not wisely constructed, the key to improve the accuracy is to make the sampling procedure converge, improved methods include combining forward and reverse FEP averages [161], path-sampling scheme with a built-in criterion for diagnosing the convergence [162], path-sampling averaged by a biased distribution of non-equilibrium paths [163, 164], and the waste-recycling Monte Carlo with optimal estimates [165, 166]. An alternative approach is to compute the chemical potential of the liquid mixture directly by gradual insertions and removals of a test particle in the fluid [102, 167]. However, these improved methods are mostly more invasive and may be more difficult to implement with software like LAMMPS and HOOMD-blue. On the other hand, since the derivatives of the free energy are usually measurable quantities in direct MC or MD simulations. Based on this, the thermodynamic integration (TI) method [72, 153] outlines a practical way of computing the free energy by evaluating its derivatives along a reversible path connecting the target system and a reference system, the integration of the derivatives along the path gives the free energy difference between the two systems.

In principle, one can obtain the absolute free energy of solid and liquid phases by referencing to a harmonic crystal and the ideal gas, respectively, whose free energy can be analytically derived. However, it is generally not a good idea to treat the liquid and solid phases in separate frameworks when it is the free energy difference that controls phase stability [168], and many phenomena of interest such as crystal nucleation and growth occur
when the liquid becomes supercooled, when it behaves so differently from the ideal gas that one need to be very careful to obtain the required accuracy by using the ideal gas as the starting point. Here, we choose the harmonic crystal, such as the Einstein crystal [169,170] as a global initial reference system, since it can provide a reliable reference for pure solid phases in most cases. The general strategy is as follows: first, we calculate the absolute free energy of the solid phase directly using an Einstein-crystal reference; next, we determine the free energy difference between the solid and liquid phases at a specific state point; and finally, we use Gibbs-Duhem integration [171] to extend to other state points, such that all the free energy calculations are based on the same initial reference system. While the free energy difference at an arbitrary state point can be calculated by methods such as pseudo-supercritical path integration [172,173], in this work, we choose a special state point: the melting point, at which the free energy difference is zero. The accurate melting point is determined by monitoring the migration of a solid-liquid interface.

![Flowchart of the algorithm](image)

**Figure 3.1** The flowchart of our algorithm.
Next, we construct an “alchemical” path to transform a pure liquid to a liquid alloy, and apply TI to evaluate the mixing enthalpy and entropy during the process. Similar methods were frequently used previously to analyze affinity change upon substitution of certain atoms or functional groups in chemical or biochemical systems [154, 174, 175]. This strategy, together with a reliable method of determining solid/liquid free energy difference in single-element systems, forms a self-contained way of establishing phase equilibria in alloys. A flowchart of this algorithm is shown in Figure 3.1: the systems are represented with rectangles, the path to compute the free energy difference of the systems are represented with pentagons, the system located relatively to the upper is the reference system, whereas the one located relatively to the lower is the target system, every system has a common initial reference system: the Einstein crystal. $G(T)$ means free energy as a function of temperature, $G(T, x)$ means free energy as a function of temperature and composition.

### 3.2 Computational Details

All simulations are performed using the MD technique with a timestep of 2 fs, as implemented in LAMMPS GPU-accelerated package [98, 176]. Systems are fully equilibrated in 500,000 timesteps in canonical ensemble (NVT) or isothermal-isobaric ensemble (NPT) with the Nose-Hoover thermostat [109, 177]. The main purpose of performing MD simulations in this work is to calculate the ensemble average of certain quantities (details are shown below), which is equivalent to the temporal average under the ergodic hypothesis. The average is collected in another 500,000 timesteps after the equilibrium is reached. For efficient energy and force calculations, we use semi-empirical inter-atomic potential in the Finnis-Sinclair form [85], which was developed to reproduce pure Al properties, energetics of Al-Sm intermetallic alloys and Al-Sm liquid structures [178]. This potential was particularly designed to treat Al-rich alloys (at.% Sm < ~10%).
3.2.1 Pure fcc-Al and Al liquid

We start with the calculation of free energy of the fcc-Al phase with TI, using the Einstein crystal as a reference system. The Helmholtz free energy of a classical Einstein crystal can be determined analytically as $F_0 = 3Nk_BT\ln(h\nu/k_BT)$, with $N$ the number of atoms, $h$ the Planck constant, $\nu$ the vibrational frequency and $k_B$ the Boltzmann constant. To implement TI, one generates intermediate systems with potentials $U(\lambda) = (1-\lambda)U_E + \lambda U_{Al}$, where $U_E$ and $U_{Al}$ stand for the potentials for the Einstein crystal and the real Al system, respectively. Then, the difference in Helmholtz free energy between the two systems can be expressed as

$$F_{Al,s} - F_0 = \int_0^1 \langle U(\lambda) \rangle_{\lambda,NVT} d\lambda - \int_0^1 \langle U_E - U_{Al} \rangle_{\lambda,NVT} d\lambda$$  \hspace{1cm} (3.1)
In Equation 3.1, the subscript s stands for solid, and $\langle \ldots \rangle_{\lambda,NVT}$ denotes the canonical ensemble (NVT) average of fcc-Al with respect to the intermediate potential $U(\lambda)$. The volume is fixed at the equilibrium volume at ambient pressure, which is determined separately via MD simulation with the real FS potential for Al under NPT conditions. In this way, the Helmholtz free energy is equal to the Gibbs free energy at the same temperature.

As an example, we show in Figure 3.2 the integrand of Equation 3.1 for the implementation of TI at 800 K, open circles are data points collected in separated MD runs. The solid line is a cubic spline interpolation. The vibrational frequency $\nu$ for the Einstein crystal is chosen to be 5 THz, which is close to the principal peak of Al phonon density of states [179]. The integration, performed based on cubic spline interpolation of discrete data points collected by separate MD runs (red open circles), gives the free energy difference between fcc-Al and Einstein crystal reference $\Delta F = -3.872$ eV/atom.

Figure 3.3 The solid-liquid interface velocity for pure Al in the [100] direction as a function of temperature.
To calculate the free energy of Al liquid, we first determine the melting point \( T_m \) of fcc-Al under ambient pressure, at which the difference in Gibbs free energy between the solid and liquid phases \( \Delta G = 0 \). Following the method described in Ref. [180], we plot the solid-liquid interface (SLI) velocity, obtained from MD simulation for the [100] direction, as function of temperature (see Figure. 3.3). The melting temperature determined from these data is 915.7 ± 0.5 K, which is slightly lower than the experimental value (933 K). The Gibbs free energy difference at other temperatures is readily available by integrating the Gibbs-Helmholtz equation

\[
\left[ \frac{\partial (\Delta G/T)}{\partial T} \right]_P = -\frac{\Delta H}{T^2}
\]

(3.2)

where \( \Delta H \) is the enthalpy change in the liquid and solid phases, or, the latent heat. The absolute free energy for Al liquid can be obtained by combining the information on solid-liquid free energy difference and the absolute free energy for the solid fcc-Al calculated previously. The final results are shown in Figure. 3.4.

![Figure 3.4 Absolute Gibbs free energy of fcc and liquid Al as a function of temperature under the ambient pressure.](image)
3.2.2 \textit{Al}_{1-x}\textit{Sm}_x$ liquid

First, we introduce an auxiliary system $\text{Al}_{1-x}\text{Al}'_x$, in which the $\text{Al}'$ is a factitious isotope of $\text{Al}$ atom, $\text{Al}'$ atom has the mass of $\text{Sm}$ but interacts with other atoms as $\text{Al}$ atom does. Thus, the Hamiltonian of $\text{Al}_{1-x}\text{Al}'_x$ can be written as

$$H(p, q) = \sum_{i=1}^{3N(1-x)} \frac{p_i^2}{2m_{\text{Al}}} + \sum_{i=1}^{3Nx} \frac{p_i^2}{2m_{\text{Sm}}} + U_{\text{Al}}(q)$$

(3.3)

where $N$ is the number of atoms in the system, $m_{\text{Al}}$ and $m_{\text{Sm}}$ are the mass for $\text{Al}$ and $\text{Al}'$ atoms, respectively, and $(p, q)$ refers to a point in the phase space $(p_1, p_2, \ldots, p_{3N}, q_1, q_2, \ldots, q_{3N})$.

The Helmholtz free energy for the auxiliary liquid system is defined as

$$F_{\text{AlAl'}} = -k_B T \ln \left\{ \frac{\int dp dq e^{-\beta H(p,q)}}{[N(1-x)]!(Nx)!h^{3N}} \right\}$$

$$= -k_B T \ln \left\{ \frac{\int dq e^{-\beta U_{\text{Al}}(q)}}{[N(1-x)]!(Nx)!\Lambda_{\text{Al}}^{3N(1-x)} \Lambda_{\text{Sm}}^{3Nx}} \right\}$$

$$= Nk_B T \left\{ \ln \left( \Lambda_{\text{Al}}^{3(1-x)} \Lambda_{\text{Sm}}^{3x} \right) + (1-x) \ln(1-x) + x \ln(x) \right\}$$

(3.4)

$$+ \ln N - 1 + \frac{\ln \int e^{-\beta U_{\text{Al}}(q)} N}{N}$$

where $\beta = 1/k_B T$, and $\Lambda_{\alpha}$ is the de Broglie wavelength for species $\alpha$, which is defined as $\Lambda_{\alpha} = (\frac{k^2}{2\pi m_{\alpha} k_B T})^{1/2}$. For pure Al liquid:

$$F_{\text{Al}} = -k_B T \ln \left\{ \frac{\int dp dq e^{-\beta H(p,q)}}{N!h^{3N}} \right\}$$

$$= -k_B T \ln \left\{ \frac{\int e^{-\beta U_{\text{Al}}(q)}}{N!\Lambda_{\text{Al}}^{3N}} \right\}$$

$$= Nk_B T \left\{ \ln \left( \Lambda_{\text{Al}}^3 \right) + \ln N - 1 + \frac{\ln \int e^{-\beta U_{\text{Al}}(q)} N}{N} \right\}$$

(3.5)

Note that in last steps of both Equation. 3.4 and 3.5, Stirling’s approximation is used since $N$ is a large number: $\ln N! = N\ln N - N + O(\ln N)$. Comparing Equation. 3.4 and 3.5, one can obtain the Helmholtz free energy difference between $\text{Al}_{1-x}\text{Al}'_x$ liquid and pure
Al liquid:

\[ F_{Al_{1-x}Al'_x} - F_{Al} = Nk_BT\left\{ \ln \left( \frac{\Lambda_{Al}^{3(1-x)} \Lambda_{Sm}^{3x}}{\Lambda_{Al}^3} \right) - \ln \left( \Lambda_{Al}^3 \right) + (1-x)\ln(1-x) + x\ln x \right\} \]

\[ = Nk_BT\left[ \frac{3}{2} x\ln \frac{m_{Al}}{m_{Sm}} + x\ln x + (1-x)\ln(1-x) \right] \]  \hspace{1em} (3.6)

Since Al\(_{1-x}Al'_x\) liquid and pure Al liquid share the same interaction potential, the equilibrium volume should also be the same under the same pressure. Thus, Equation. 3.6 also describes the Gibbs free energy difference between the two systems (the \(PV\) term cancels out).

![Integrand of Equation. 3.7 for \(x = 0.25\) and \(T = 1500\) K.](image)

Figure 3.5  The integrand of Equation. 3.7 for \(x = 0.25\) and \(T = 1500\) K.

Next, we use TI to transform the factitious Al\(_{1-x}Al'_x\) system to the real Al\(_{1-x}Sm_x\) system. To do that, we introduce intermediate systems interacting as \(U(\lambda) = (1 - \lambda)U_{Al} + \lambda U_{AlSm}\).
Then, the difference in Gibbs free energy between the two systems can be expressed as:

$$G_{\text{AlSm}} - G_{\text{AlAl}'} = \int_0^1 \langle \frac{dU(\lambda)}{d\lambda} \rangle_{\lambda,NPT} d\lambda = \int_0^1 \langle U_{\text{AlSm}} - U_{\text{Al'}} \rangle_{\lambda,NPT} d\lambda$$

(3.7)

where $\langle \ldots \rangle_{\lambda,NPT}$ stands for the isothermal-isobaric (NPT) ensemble average with respect to the intermediate potential $U(\lambda)$.

Here we show the calculation of $x = 0.25$ and $T = 1500$ K as an example to describe the free energy calculation of liquid Al$_{1-x}$Sm$_x$ alloys. The transformation from Al liquid into the factitious Al-Al' liquid results in a free energy change $\Delta F = -0.156$ eV/atom, as calculated according to Equation. 3.6. The implementation of TI to transform Al-Al' into the real Al-Sm system is shown in Figure. 3.5, which gives $G_{\text{AlSm}} - G_{\text{AlAl}'} = -0.404$ eV/atom. In Figure. 3.5, open circles are data points collected in separated MD runs. The solid line is a cubic spline interpolation. Thus, the net difference of Gibbs free energy between the Al$_{0.75}$Sm$_{0.25}$ liquid and pure Al liquid is $-0.560$ eV/atom.

In this work, free energies of liquid alloy Al$_{1-x}$Sm$_x$ system, $x = 0.00 \sim 0.25$, $T = 700 \sim 1500$ K have been computed.

### 3.3 Applications

In this section, we demonstrate two applications of free energy calculations outlined in the above, namely, the determinations of driving force for fcc-Al nucleation in supercooled Al$_{1-x}$Sm$_x$ liquid and the melting curve for fcc-Al and Al$_3$Sm crystals.

#### 3.3.1 Driving force for nucleation of fcc-Al in supercooled Al$_{1-x}$Sm$_x$ liquid

Crystal nucleation in supercooled liquid is an important process in numerous areas of physical science [181]. It is also an important factor for glass formation, since glass is formed by suppressing crystal nucleation during fast quenching. As a marginal glass former, the
glass formability of Al-Sm has a strong dependence on the Sm concentration [182]. When as-quenched Al-Sm glass is gradually heated, the devitrification process often starts with the deposit of Al nanocrystals [183, 184]. Thus, study of the effect of Sm concentration on Al nucleation in supercooled Al-Sm liquids can provide useful information for both glass formation and devitrification processes [185].

Figure 3.6 Driving force for nucleation of fcc-Al in $\text{Al}_{1-x}\text{Sm}_x$ liquid at 700 K as a function of the Sm composition.

The driving force is a fundamental parameter that describes the net bulk free energy gain upon the formation of a crystalline nucleus. For fcc-Al nucleation in supercooled $\text{Al}_{1-x}\text{Sm}_x$ liquid, the phase transition for one Al atom is:

$$N \text{Al}_{1-x}\text{Sm}_x \rightarrow \text{Al}_{\text{fcc}} + (N - 1) \text{Al}_{1-x-\delta x}\text{Sm}_{x+\delta x}$$

(3.8)
where $N$ is the total number of the $\text{Al}_{1-x}\text{Sm}_x$ molecules, $\delta x = \frac{x}{N-1}$ using conservation laws, which is a small value. The driving force can be expressed as the free energy change in this transition:

$$\Delta \mu(x) = G(\text{Al}_{fcc}) + (N-1)G(\text{Al}_{1-x-\delta x}\text{Sm}_{x+\delta x}) - NG(\text{Al}_{1-x}\text{Sm}_x)$$

$$= G_s + (N-1)G_l + (N-1)\frac{\partial G_l}{\partial x} \delta x - NG_l$$

$$= x \frac{\partial G_l}{\partial x} + G_s - G_l$$

(3.9)

where $G_l$ and $G_s$ refer to the Gibbs free energy of $\text{Al}_{1-x}\text{Sm}_x$ liquid and fcc-Al, respectively.

In Figure. 3.6, we plot $\Delta \mu$ as a function of $x$ at a temperature of 700 K, where one can see that the driving force decreases as the Sm composition increases, but remains negative within the range of $x < 0.12$, showing that nucleation of fcc-Al is thermodynamically favoured within this composition range. However, it should be noted that fcc-Al is the only solid phase considered in Figure. 3.6. When $x$ becomes large (before reaching 0.12), nucleation of other solid phases such as $\text{Al}_3\text{Sm}$ will become thermodynamically more favourable than fcc-Al.

### 3.3.2 Melting curve (liquidus line) for fcc-Al and $\text{Al}_3\text{Sm}$

We also perform the free-energy calculation for the hexagonal $\text{Al}_3\text{Sm}$ phase, and traced out the melting curve (liquidus line) for both fcc-Al and $\text{Al}_3\text{Sm}$. We focus on the technologically important Al-rich region for this system, in which fcc-Al and $\text{Al}_3\text{Sm}$ are the only two relevant solid phases according to the Al-Sm phase diagram [186,187]. Each coexistence point on the melting curve $(x, T)$ of a solid phase denotes a coexistence state, which satisfies:

$$(x - x_s) \frac{\partial G_l(x, T)}{\partial x} + G_s(T) = G_l(x, T)$$

(3.10)

where $G_l$ and $G_s$ are the the Gibbs free energy of the liquid and solid phases, respectively, and $x_s$ is the Sm composition in the solid phase. Mathematically, the coexistence composition at a specific temperature can be determined by the “tangent” construction as shown in
Figure 3.7, in which the formation Gibbs free energy of $\text{Al}_{1-x}\text{Sm}_x$ liquid $G_f$ is plotted as a function of the Sm composition $x$, at a supercooled temperature of 880 K.

$G_f$ is calculated using the Gibbs free energy of fcc-Al and $\text{Al}_3\text{Sm}$ at the same temperature as reference states. In this way, $G_f$ for the two solid phases is zero (see Figure. 3.7). We construct tangential lines from the fcc-Al and $\text{Al}_3\text{Sm}$ phases to the liquid curve, shown as the red and blue lines in Figure. 3.7, respectively. The composition data is computed per $\Delta x = 0.01$, note that we do not have to start from pure Al liquid at each calculation, following the “alchemical” path, thermodynamic integrations are run with intermediate potentials, which can be embarrassingly parallelized. The red and blue lines denote the tangential lines from fcc-Al and $\text{Al}_3\text{Sm}$ to the liquid curve, respectively. The tangential points give the
coexistence liquid composition with the two solid phases, respectively. From the position of the tangential points, one can determine the coexistence liquid composition with the two solid phases to be 0.039 and 0.054, respectively.

![Melting curve for fcc-Al and Al3Sm from experiments and the current calculations for Al\(_{1-x}\)Sm\(_x\) in the Al-rich region of the Al-Sm system.](image)

Figure 3.8 Melting curve for fcc-Al and Al\(_3\)Sm from experiments and the current calculations for Al\(_{1-x}\)Sm\(_x\) in the Al-rich region of the Al-Sm system.

The above procedure is repeated for various other temperatures to map out the melting curve for fcc-Al and Al\(_3\)Sm, as shown in Figure 3.8. Our calculations predict a eutectic point at \(T = 863\) K and \(x = 0.051\), while the eutectic point from previous experiments is located at \(T = 908\) K and \(x = 0.03\). At compositions away for the calculated eutectic point, our calculations generally underestimate the liquidus temperature by less than 100 K. Since the only energetics data used in fitting the Al-Sm FS potential was generated by density-functional theory (DFT) calculations at 0 K [178], we do expect some discrepancy
with experiments in thermodynamic properties at finite temperatures. In this regard, a systematic way of determining solid-liquid phase equilibria, as outlined in the current chapter, is valuable if one wants to refine a classical potential in order to more faithfully reproduce experimental thermodynamic information.

3.4 Conclusion

In this chapter, we described the work in establishing a self-contained algorithm to rigorously evaluate the free energy for solid and liquid phases of an alloy system, based on thermodynamic integration. The algorithm starts from calculating the free energy of a single-element solid phase by referencing to a harmonic crystal. By monitoring the solid-liquid interface migration at different temperatures, we determine the melting point of the solid phase, which establishes a state of equality between the solid and liquid free energy. The free energy difference at other state points between solid and liquid phases can be obtained by integrating $\Delta H/T^2$ with temperature, where $\Delta H$ is the latent heat during melting. Then, we generate an alchemical path connecting a pure liquid to a liquid alloy, and use thermodynamic integration to evaluate the mixing enthalpy and entropy. As an example, we apply this method on the Al-Sm system to determine the driving force for Al nucleation in Al-Sm liquid and the melting curve for the solid phases Al and Al₃Sm.
CHAPTER 4. KINETIC EFFECT IN THE DETERMINATION OF THE NON-STOICHIOMETRIC SITE OCCUPATIONS

In Chapter 3, the work in establishing a self-contained algorithm to rigorously evaluate the free energy for solid and liquid phases of an alloy system has been described, although the free energy is fundamental to achieving microscopic understandings of freezing and melting phenomena, however, these phenomena can be more complex, making it difficult to be explained solely by the structures’ free energy profile itself. In this chapter, we report a work in observing and simulating the spatially-correlated site occupancy in the non-stoichiometric meta-stable $\varepsilon$-Al$_{60}$Sm$_{11}$ phase during devitrification of Al-10.2 at.% Sm glasses, this work reveals the kinetic effects in the phase selection.

A meta-stable $\varepsilon$-Al$_{60}$Sm$_{11}$ phase appears during the initial devitrification of as-quenched Al-10.2 at.% Sm glasses. The $\varepsilon$ phase is non-stoichiometric in nature since Al occupation is observed on the 16$f$ Sm lattice sites. Scanning transmission electron microscopic (STEM) images reveal profound spatial correlation of Sm content on these sites, which cannot be explained by the “average crystal” description from Rietveld analysis of diffraction data. Thermodynamically favourable configurations, established by Monte Carlo (MC) simulations based on a cluster-expansion model, also give qualitatively different correlation functions from experimental observations. On the other hand, molecular dynamics simulations of the growth of $\varepsilon$-Al$_{60}$Sm$_{11}$ in undercooled liquid show that when the diffusion range of Sm is limited to $\sim 4\text{Å}$, the correlation function of the as-grown crystal structure agrees well with that of the STEM images. Our results show that kinetic effects, especially the limited
diffusivity of Sm atoms plays the fundamental role in determining the non-stoichiometric site occupancies of the $\epsilon$-$\text{Al}_{60}\text{Sm}_{11}$ phase during the crystallization process.

### 4.1 Introduction

Al alloyed with $\sim 10$ at.$\%$ Sm represents a typical Al-rare earth (RE) system that can undergo deep undercooling from liquid and form amorphous solids or nanocrystalline composite materials with much improved mechanical properties compared with pure Al [182, 188–190]. When the as-quenched amorphous Al-10.2 at.$\%$ Sm melt-spun ribbons are heated, a metastable cubic phase is usually the first phase to appear in the multiple-step devitrification process [63].

![Figure 4.1](image)  
*Figure 4.1  The $\epsilon$-$\text{Al}_{60}\text{Sm}_{11}$ phase.*

Although this cubic phase was reported more than 20 years ago [183, 191], not until recently has its atomic structure been solved by an approach integrating experimental diffraction data, a genetic algorithm for crystal structure prediction, and molecular dynamics (MD) simulations [63]. The solved phase, labelled as $\epsilon$-$\text{Al}_{60}\text{Sm}_{11}$, has a body-centered cubic (BCC)
unit cell (space group $Im\bar{3}m$, No. 229) with a lattice constant of 13.9 Å. For the stoichiometric $\epsilon$-Al$_{60}$Sm$_{11}$ crystal, each cubic unit cell contains 120 Al atoms and 22 Sm atoms. However, both Rietveld analysis and MD simulations show that, among the 22 Sm sites, only 6 sites are fully occupied by Sm, and the remaining sites with a Wyckoff notation 16$f$ are shared with Al atoms, resulting in a non-stoichiometric phase Al$_{60+x}$Sm$_{11-x}$ with $x \sim 4$ (see Figure. 4.1). In Figure. 4.1, (a) is the unit cell of the $\epsilon$-Al$_{60}$Sm$_{11}$ phase. Red and blue denote Sm and Al atoms, respectively. The half red and half blue sites can be occupied by either Al or Sm (PO Sm sites). (b) is the top view of the FO and PO Sm sites of the [001] plane. The orange square shows a unit cell of the projected 2D lattice.

In Ref. [63], the authors argue that the tolerance to Al occupation on the PO sites is an important reason that the phase can readily nucleate and grow. However, it remains obscure what is the underlying reason for the shared occupation. Traditionally, nonstoichiometry in inorganic compounds is understood from thermodynamic factors such as enthalpy of formation and configurational entropy [192], while the kinetics associated with processing conditions is often overlooked. In this chapter, we collect detailed information about how the Sm/Al atoms occupy the PO sites in the $\epsilon$ phase from scanning transmission electronic microscopy (STEM), which clearly shows that the occupancies of the PO sites are spatially correlated across multiple unit cells. Monte Carlo (MC) and molecular dynamics (MD) simulations show that the spatial correlation is not originated from the thermodynamically favourable configurations. Instead, kinetic effects, especially the limited diffusivity of Sm atoms plays an important role in creating the observed spatial correlations.

### 4.2 Experiments

To further characterize the morphology of the phase, the specimens are quenched to room temperature, then they are prepared for scanning transmission electron microscopy (STEM)
characterization by focused ion beam (FIB) milling until the specimen thickness is as thin as around 50~100 nm. The FIB instrument in use is the FEI dual-beam Helios NanoLab G3 UC, and a FEI Titan Themis 300 Cube aberration corrected scanning transmission electron microscope. The high angle annular dark-field scanning transmission electronic microscopic (HAADF-STEM) is used to characterize the atoms projected onto the [001] crystal plane of the specimens, which has a square lattice with \(a = 9.83\AA\), as shown in orange box in Figure. 4.1 (b).

![Figure 4.2 Raw and processed HAADF-STEM image.](image)

In Figure. 4.2, (a) is an example of raw HAADF-STEM image; (b) is the zoom-in picture of the gray shaded area in (a), the orange box denotes a unit cell in the projected 2D lattice, the FO and PO sites along a \(\langle 110 \rangle\) line are labelled; and (c) is as the same as figure (b) with normalized intensities on all the atomic sites.
Figure 4.2 (a) shows a typical HAADF-STEM image, where only Sm atoms are clearly seen. This is reasonable since Sm has a much higher electron density than Al. Figure 4.2 (b) zooms in a portion of Figure 4.2 (a), in which two PO sites (PO1 and PO2) and one FO site within a unit cell are labelled along the $\langle 110 \rangle$ direction. The intensity at each Sm site is generally proportional to the total number of Sm atoms in the projected column. However, the raw image of Figure 4.2 (b) cannot be directly used to compare Sm occupancies at different sites, because of the non-uniform background intensities, which is indicative of fluctuations of thickness and composition in the sample. On the other hand, it is reasonable to assume that such fluctuation is negligible within one unit cell. Thus, one can use the intensity at the FO site to calibrate the intensity at other sites of the same unit cell along a $\langle 110 \rangle$ line. More specifically, we define a normalized intensity as $I_n = I/(2I_{FO})$, where $I$ is the absolute intensity of an atomic site directly read from the raw image such as Figure 4.2 (a), and $I_{FO}$ is the absolute intensity of the FO site in the same unit cell. A factor of 2 is added in the denominator because the number of PO sites doubles that of the FO sites in a projected column. Figure 4.2 (c) shows the normalized intensity along $\langle 110 \rangle$ lines of the same region as that of Figure 4.2 (b).

Figure 4.3 (a) shows the normalized intensity along a $\langle 110 \rangle$ crystal line (circled by a yellow box in Figure 4.2 (a)), which indicates a large variation of the occupancy on the PO sites. Interestingly, if the difference of $I_n$ between two neighboring PO sites, PO1 and PO2 as shown in Figure 4.2 (c), $I_d = I_{n,PO1} - I_{n,PO2}$ is plotted versus the unit cell number along the same line in Figure 4.3 (b), one can see an oscillatory behavior with a period of multiple unit cells, suggesting some spatial correlations in the PO sites occupancy.
In Figure 4.3, (a) is the normalized intensity \( I_n \) along the \( \langle 110 \rangle \) crystal lines circled by a yellow box in Fig. 2(a), the red circles denote the FO sites, whose intensities are set to unity, the green and blue circles denote PO1 and PO2 sites, respectively; and (b) is the difference of the intensity \( I_d \) between two neighboring PO sites versus the unit cell number.

![Figure 4.3 Intensity along the \( \langle 110 \rangle \) crystal lines.](image)

To better capture the spatial correlation, we define the correlation function between two unit cells along the \( \langle 110 \rangle \) direction as:

\[
g(R) = \langle I_d(r)I_d(r + R) \rangle_r
\]  

(4.1)

where \( r \) is the position of a PO site, \( R \) is a lattice vector along the \( \langle 110 \rangle \) direction in the projected 2D lattice (see Figure. 4.2 (b)), and \( \langle \ldots \rangle_r \) denotes the average over all different positions \( r \). In Figure. 4.4, we show the correlation function \( g(R) \) averaged over all \( \langle 110 \rangle \) lines in six STEM images collected from 2 Al\textendash{}Sm specimens. For comparison purpose, we also show the correlation function generated by assuming the occupancy of Sm atoms on the PO sites is completely random. The solid lines are fittings to the decay function \( g(R) = Ae^{-R/\xi} \). According to Equation. 4.1, \( A = g(0) = \langle I_d^2(r) \rangle \), which essentially
measures the variation \((\langle I_d^2 \rangle - \langle I_d \rangle^2)\) of the distribution of Sm occupancy in all the PO sites since \(\langle I_d \rangle^2\) is very small \(< 10^{-4}\) in all the samples studied in this chapter. \(\xi\) is a parameter characterizing the correlation length. The value of \(A\) is 0.020 ± 0.002 and 0.0062±0.0002 for the STEM and random samples, respectively, indicating that the variation of Sm occupancy in experimental samples is significantly larger than that from a random sample. The correlation length \(\xi\) for the experimental sample is 0.98\(a\)±0.25\(a\), demonstrating non-vanishing correlations over multiple lattice constants. In contrast, \(\xi\) for the random sample is merely 0.25\(a\) ± 0.11\(a\), showing that the correlation is essentially zero since \(R/a\) is a positive integer number.

Figure 4.4 The averaged correlation function along \(\langle 110 \rangle\) lines of STEM samples (blue), and randomly generated samples (red).
4.3 Simulation Details

To understand the origin of the correlation of site occupancy observed in experiments, we perform Monte Carlo (MC) simulation [71,193] based on a cluster expansion (CE) model [194,195], as well as classical molecular dynamics (MD) simulation [72,196]. MC simulations emphasize on establishing the thermodynamically favorable configurations; while kinetics especially the effects Sm mobility is addressed in MD simulations.

4.3.1 Monte Carlo simulation

Since there is a well-defined underlying BCC lattice, and the partial occupancy is only associated with certain fixed lattice sites (i.e., the 16f PO sites), the free energy of the system can be regarded as a function determined only by the configurations of different chemical elements on the PO sites. Cluster expansion (CE) is an appropriate method to establish the free energy model for this system [194]:

\[
F = F_0 + \sum_i J_i S_i + \sum_{i<j} J_{ij} S_i S_j + \sum_{i<j<k} J_{ijk} S_i S_j S_k + \ldots
\]  

(4.2)

where \( S_i \) is the chemical occupation of the partially occupied sites, which is set to be 1 (−1) if the site \( i \) is occupied by Al (Sm), \( J_i \), \( J_{ij} \), \( J_{ijk} \), ... are the cluster expansion (CE) coefficients, and \( F \) is the total free energy of the system which includes the internal energy and a contribution from the vibrational entropy. \( F \) in Equation 4.2 is expanded as a summation of contributions from all “clusters” including singlets, doubles, triplets, etc. In practice, only the first few terms are included in Equation 4.2 assuming that higher order clusters have negligible contributions to \( F \).

To fit the cluster expansion coefficients, we first calculate the free energy of 18 training structures with various occupation of the PO sites. The free energies of the training structures are calculated within the quasi-harmonic approximation [197,198] at two different
temperatures $T = 500$ K and 800 K. To further expedite the calculations, we have employed a classical potential in the Finnis-Sinclair (FS) form [85], which was carefully fitted to the ab-initio energetics of a series of Al-Sm compounds with $\sim 10$ at.% Sm as well as the pair correlation functions of Al$_{80}$Sm$_{10}$ liquid [178]. Using this potential, we were able to identify the $\epsilon$-Al$_{60}$Sm$_{11}$ phase in a genetic algorithm search. Moreover, the X-ray diffraction pattern of the $\epsilon$ phase with built-in antisite defects directly grown in molecular dynamics simulations using the same potential matches excellently with experiments, suggesting that this interatomic potential also correctly captures the nature of defects in $\epsilon$-Al$_{60}$Sm$_{11}$ [63]. With these data in hand, we then use the Alloy Theoretic Automated Toolkit (ATAT software) [106] to fit the CE coefficients.

![Figure 4.5 Comparison of the formation free energy calculated using CE model ($E_{CE}$) and FS potential ($E_{FS}$) for 851 Al$_{120+y}$Sm$_{22-y}$ structures, at (a) $T = 500$ K and (b) $T = 800$ K.](image)

As a validation of the CE model, we apply our model to compute the formation free energy of the structures of a unit cell (Al$_{120+y}$Sm$_{22-y}$) with different compositions and PO sites configurations, using the two end compositions $y = 0$ and $y = 16$ as references. These
values are compared with those directly calculated by quasi-harmonic approximation based on the FS potential. A total of 851 inequivalent structures containing 142 atoms in a cubic unit cell are used in this step. The comparison is shown in Figure 4.5 (a) and Figure 4.5 (b), for $T = 500$ K and $800$ K, respectively. We define the root mean square discrepancy ($RMSD$) to quantitatively evaluate the error of the CE model:

$$RMSD = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (E_{CE}^i - E_{FS}^i)^2} \tag{4.3}$$

where $N$ is the total number of the structures (in this case, $N = 851$), $E_{CE}^i$ and $E_{FS}^i$ are the formation free energy of structure $i$ computed using the CE model and directly from the FS potential, respectively. $RMSD = 0.34(0.48)$ meV/atom for $T = 500(800)$ K. Both values are several orders of magnitude smaller than the thermal fluctuation $k_B T$ at corresponding temperatures, indicating that the error of the CE model is well controlled.

Figure 4.6 MC simulation results.
Based on the validated CE model, we simulate the equilibrium configuration of the PO sites of the $\epsilon$-Al$_{60}$Sm$_{11}$ phase at 500K and 800K, using the Metropolis MC algorithm [71]. 500 K is close to the temperature (468 K) at which the spatial correlation is observed in experiments, and 800 K is used for comparison with MD simulations that will be shown later. Since we only need to explicitly consider the 16 PO sites in a unit cell with size $\sim$ 1.4 nm, we can create a fairly large simulation box containing 20 $\times$ 20 $\times$ 20 = 8,000 unit cells (128,000 PO sites), with a box size of 27.8 nm. Initially, equal number of Sm and Al atoms are randomly distributed in the PO sites, corresponding to an overall composition of Al$_{64}$Sm$_7$, which is close to that of the experiments. Figure 4.6 (a) and (b) show the free energy as a function of Monte Carlo (MC) step at 500 K and 800 K, respectively. Both figures level off at the end of the simulation, indicating that simulations have reached the thermodynamic equilibrium.

To generate a “simulated” STEM image, we project the MC samples along the $\langle 001 \rangle$ direction, and make the intensity at each PO site directly proportional to the number of Sm atoms in the projected column (see Figure. 4.6 (c) and (d)). Then, the $\langle 110 \rangle$ lines in the projected image are collected to analyze the correlation of Sm occupancies in the same way as we did on the experimental samples. However, the samples in our MC simulations are still significantly smaller than the experimental samples: the thickness of the MC samples along the projection direction is 27.8 nm, while that of experimental samples ranges from 50 to 100 nm. Here, a bootstrapping technique [199] is used to better match the experimental parameters. That is, two $\langle 110 \rangle$ lines randomly collected from the MC sample are stacked together to reproduce a “bootstrapped” sample with a thickness of 55.6 nm that is directly comparable with experiments. The number of total $\langle 110 \rangle$ lines that can be used for statistical analysis is also increased to $\left( \begin{align*} N \\ k \end{align*} \right)$, where $N$ is the number of lines originally collected from the MC sample.

The correlation functions $g(R)$ as defined in Equation. 4.1 of the bootstrapped MC samples at both 500 K and 800 K are given in Figure. 4.6. Again, the decay function
$g(R) = Ae^{-R/\xi}$ is used to fit the data, shown in Figure. 4.7, solid lines are fittings to the function $g(R) = Ae^{-R/\xi}$. The fitting parameters at both temperatures are much closer to the randomly generated sample than to the STEM samples (see Figure. 4.2). This clearly demonstrates that the thermodynamically equilibrated configuration does not reproduce the observed spatial correlation in the experiments.

![Figure 4.7 Averaged correlation function along ⟨110⟩ lines of the MC samples, at 500 K (blue) and at 800 K (red).](image)

4.3.2 Molecular dynamic simulation

We then run MD simulations using the GPU-accelerated LAMMPS code [98, 176] to simulate the growth of the $\epsilon$-phase. The same classical potential as the one used to construct the cluster expansion model in MC simulations is used in the simulations [178]. Unlike MC, MD directly integrates Newtonic equations of motion and automatically include the dynamical effects. On the other hand, unlike our MC simulations which only consider the
atoms occupying the PO sites, all atoms must be explicitly treated in MD simulations. Therefore, we can only afford to simulate a much smaller system.

![Figure 4.8 Atomic configurations during molecular dynamic (MD) simulations.](image)

In our simulations, the initial configuration is constructed by inserting a seed layer of the \( \epsilon \) phase into the undercooled liquid (see Figure. 4.8 (a)), which has a Sm concentration of 10.2 % and is pre-equilibrated. Periodic boundary conditions are also used in the initial configuration. The seed layer contains \( 4 \times 4 \times 1 \) unit cells, in which the PO sites are randomly occupied by equal number of Sm and Al atoms. Since at the experiment temperature of \( \sim 500 \) K, the growth kinetics is too slow to be accessed within MD time scales, we run MD simulations at an elevated temperature of 800 K. At the same time, to better account for the limited Sm diffusivity at lower temperatures, we also run the “constrained” MD simulations, in which the diffusion of a Sm atom is limited to within a sphere of radius \( r_c \) from its initial position in the liquid phase, by applying an additional potential \( V(r) = k|\mathbf{r} - \mathbf{r}_0|^6 \). \( \mathbf{r} \) is the position of a Sm atom and \( \mathbf{r}_0 \) is its initial position in the liquid phase. A large exponent 6 is used to ensure a flat bottom of \( V(r) \) when \( \mathbf{r} \) is close to \( \mathbf{r}_0 \). \( k \) is varied to give two different values of \( r_c \): 4.2 Å and 2.6 Å for the current work. These results are compared with those from unconstrained MD simulations (\( r_c = \infty \)). The final atomic structures for
the MD simulations are shown in Figure 4.8 (b-d): [(b) final configuration of unconstrained MD ($r_c = \infty$); (c) final configuration for $r_c = 4.2\,\text{Å}$; (d) final configuration for $r_c = 2.6\,\text{Å}$.] In all the cases, most of the liquid region has been transformed into the $\epsilon$ phase. The growth essentially stops when the two growth fronts start to merge due to the periodic boundary conditions.

The thickness of the MD samples along the $\langle 001 \rangle$ direction is 5.56 nm. Again, the bootstrapping technique is used to generate a fair comparison with experiments. The correlation functions $g(R)$ of the three MD samples with different $r_c$ are shown in Figure 4.9, solid lines are fittings to the function $g(R) = Ae^{-R/\xi}$.

![Figure 4.9](image)

Figure 4.9  Averaged correlation function along $\langle 110 \rangle$ lines of the MD simulations, with $r_c = \infty$ (blue), $r_c = 4.2\,\text{Å}$ (red) and $r_c = 2.6\,\text{Å}$ (black).
4.4 Conclusion

We show the estimated fitting parameters $A$ and $\xi$, together with their standard errors for all the samples studied in this chapter in Figure. 4.10. One can see that only the sample generated in MD simulations with $r_c = 4.2\text{Å}$ has a significant overlap with the STEM samples, indicating that limited Sm mobility plays an important role in producing the special correlation in the PO site occupancy. However, when the Sm movement is over confined ($r_c = 2.6\text{Å}$), the Sm distribution on the PO sites carries too much legacy from the random liquid structure, which is characterized by much smaller variance and essentially no inter-unit cell correlation. On the other hand, the MD sample with $r_c = \infty$ clearly overestimates the correlation length. It should be noted that even with $r_c = \infty$, the diffusion of Sm atoms in MD simulations is still local in nature due to its intrinsic low diffusivity. This is different from the MC simulations at the same temperature $T = 800 \text{ K}$, in which long range diffusion of Sm atoms is needed in order to establish the thermodynamic equilibrium. As a result, one can see strong contrast of the fitting parameters between the MC sample at 800 K and the MD sample with $r_c = \infty$.

By studying all the samples in Figure. 4.10, it comes to the conclusion: When melt-spun Al-10.2 at.% Sm glass is heated, it first devitrifies into a cubic $\epsilon$-Al$_{60}$Sm$_{11}$ phase. STEM images reveal profound spatial correlations of the Sm occupancy on the 16$f$ Wyckoff positions in the non-stoichiometric $\epsilon$ phase. Such spatial correlations cannot be reproduced by random occupation of these lattice sites, as suggested by Rietveld analysis of the X-ray diffraction spectrum based on. We perform both MC and MD simulations to try to interpret such spatial correlations. In MC simulations, a CE model is constructed to compute the free energy as a function of the Sm configuration on the partially occupied sites. The free energy predicted by the CE model is validated to be accurate within thermal fluctuations. The MC simulations successfully generate the thermodynamically favorable configurations, which do
not reproduce the spatial correlations observed in experiment. With MD, we simulate the
growth of the $\epsilon$ phase by combining the liquid together with a crystal seed at an undercooled
temperature. In addition to the conventional MD, we also run constrained MD in which
the diffusion of Sm atoms is limited within a certain range $r_c$ from its initial position by an
external potential. Our results show that when $r_c \sim 4 \text{Å}$, the as-grown phase shows spatial
correlations that matches well with experiments, implying that the limited diffusivity of Sm
is crucial for the appearance of the correlations of Sm content on the partially occupied sites.

Figure 4.10 The estimated fitting parameters $A$ and $\xi$, together with their standard
erors represented by a rectangular box for all the samples studied in
this chapter.
CHAPTER 5. PERSISTENT-EMBRYO METHOD
SIMULATION OF DEVITRIFICATION PROCESS OF
$\epsilon$-Al$_{60}$Sm$_{11}$ PHASE

In Chapter 4, the meta-stable $\epsilon$-Al$_{60}$Sm$_{11}$ phase appears during the initial devitrification of as-quenched Al-10.2 at.% Sm glasses has been simulated, using MC and MD simulation. In the MC simulation, the equilibrium configuration of the PO sites of the $\epsilon$-Al$_{60}$Sm$_{11}$ phase is simulated with a cluster expansion model based on free energy, as a prerequisite, the MC simulation assumes that there exists a well-defined underlying big cubic lattice. In the MD simulation, initial configuration is constructed by inserting a seed layer of the $\epsilon$ phase into the undercooled liquid, as a prerequisite, the MD simulation assumes that there exists a well-developed layer of non-stoichiometric $\epsilon$-Al$_{60}$Sm$_{11}$ phase. Both of these two assumptions, although reasonably model the growth stage of the $\epsilon$ phase’s expansion, but may be lack of descriptions of the embryonic stage of the $\epsilon$ phase’s development.

However, conventional molecular dynamics (MD) simulation fails to access the embryonic stage even within a very long time scale, the fundamental limitation is that the crystal nucleation from liquid in most cases is too rare to be accessed, unfortunately, the $\epsilon$-Al$_{60}$Sm$_{11}$ nucleation is a typical inaccessible process due to its slow dynamics and complicated structures. The “persistent-embryo” method (PEM) method aims at accessing the rare early nucleation case(s) within the limited time scales, to facilitate crystal nucleation in MD simulations by preventing small crystal embryos from melting using external spring forces. This method was developed in collaboration with Yang Sun, Huajing Song, Feng Zhang, et al. [200], and have found great success in studying the nucleation rate of pure Ni nucleation in a moder-
ate undercooling liquid, and the B2 phase nucleation in a strong glass-forming Cu-Zr alloy. PEM opens a new avenue to study solidification under realistic experimental conditions via atomistic computer simulation, in this chapter, we have applied PEM to model the nucleation and growth process of the $\epsilon$-Al$_{60}$Sm$_{11}$ phase in the undercooled Al$_{90}$Sm$_{10}$ liquid, and got some preliminary results, which are also reported here.

5.1 Persistent-Embryo Method

The inefficiency to sample nucleation events in conventional MD simulations can be explained using the classical nucleation theory (CNT), in CNT, the change in the free energy can be written as Equation 2.27, which is a function of nucleus radius $r$. Rewrite it by introducing the parameter $N$ atoms contained in the nucleus with any possible shape, Equation 2.27 can be transferred to the form as a function of $N$:

$$\Delta G(N) = N\Delta \mu + s \left( \frac{N}{\rho_s} \right)^{2/3} \gamma$$

(5.1)

where $\rho_s$ is the atomic density of the ordered phase (solid), $\Delta \mu < 0$ is the chemical potential change from the disordered phase to the ordered phase, $\gamma > 0$ is the free energy density in the disordered-ordered interface, and $s$ is a factor to account for the nucleus shape. Intrinsically, the nucleation is a competition between the energy gain associated with the transformation of the bulk liquid into a crystal phase ($N\Delta \mu$ term) and the energy cost of creating a solid-liquid interface ($s \left( \frac{N}{\rho_s} \right)^{2/3} \gamma$ term), such that Figure 2.6 can be re-plotted with respect to nucleus cluster size $N$, as schematically shown in Figure 5.1 (a), this competition between the bulk and interface terms leads to a critical barrier $\Delta G^*$ where the nucleus reaches the critical size $N^*$. The low probability of overcoming this free energy barrier $\Delta G^*$ makes it inefficient to sample nucleation events in conventional MD simulations [201]. By studying 5.1 (a) in detail, conclusion can be made that the composition of $\Delta G^*$ contains a large part of free energy that need to grow and keep a small crystalline cluster (embryo), the basic idea
of PEM is to adopt external perturbations to keep the embryo from melting, and remove all perturbations before the nucleus reaches the critical size $N^*$. In the initial simulation step, the crystalline embryo with $N_0$ atoms ($N_0 \ll N^*$) is created and inserted into liquid, a variable harmonic potential is applied to constrain the embryo from melting, the potential is a piecewise function of $N$ with a subcritical threshold $N_{sc}$ ($N_{sc} < N^*$) as the interval separator, in the interval $N < N_{sc}$, the constrained potential is a mono-decreasing function with respect to $N$, whereas in the interval $N \geq N_{sc}$, the constrained potential is removed. In our simulation, the harmonic spring potential is used, whose spring constant reads:

$$k(N) = \begin{cases} 
  k_0 \frac{N_{sc} - N}{N_{sc}} & N < N_{sc} \\
  0 & N \geq N_{sc}
\end{cases}$$  \hspace{1cm} (5.2)

Figure 5.1  The the persistent-embryo method.

The PEM simulation is schematically shown in Figure 5.1, (a) plots the excessed free energy (black) and spring constant (red) as a function of the crystalline cluster size $N$, where $N_0$ is the number of atoms in the constrained embryo, $N_{sc}$ is the subcritical threshold, and $N^*$ is the critical size. The PEM tune the spring constant to zero before reaching the critical
nucleus size to ensure the dynamics of the system is unbiased at the critical point. (b) is a sketch of the as-grown crystalline cluster around embryo at \( N < N_{sc} \). The yellow atoms with spring icon are the persistent embryo, the red are the as-grown atoms, showing the crystalline packing, and the gray are the liquid atoms. Note that when \( N \) reaches \( N_{sc} \), the springs are removed, so the overall process simulates the nucleation free of external potential, which is an advantage of this approach compared to others such as the lattice mold method [202]. Using the PEM simulation, the nucleation process of pure Ni and B2 phase in the \( \text{Cu}_{50}\text{Zr}_{50} \) liquid alloys is studied, the nucleation rates are computed, which agrees well with experiment.

5.2 Simulation Details

In the Al-Sm system, we focus on the structures and dynamics of the Sm atoms only, actually we have observed that, once the Sm framework has been formed to enough size, the \( \epsilon \)-\( \text{Al}_{60}\text{Sm}_{11} \) phase can be formed. So Sm atom is a suitable indicator which provides enough information. PEM spring potential is redefined accordingly:

\[
k^{Sm}(N^{Sm}) = \begin{cases} 
    k^{Sm}_{0} \frac{N^{Sm} - N^{Sm}_{sc}}{N^{Sm}_{sc}} & N^{Sm} < N^{Sm}_{sc} \\
    0 & N^{Sm} \geq N^{Sm}_{sc}
\end{cases}
\]  \hspace{1cm} (5.3)

where the superscript \( Sm \) means that the counting for solid-like atoms considers Sm atom only. But before we apply PEM to Al-Sm system, two crucial questions have to be solved: what order parameter should we use to identify the solid-like atoms \( N \); and what is the embryo looks like in this system.

5.2.1 Identifying solid/liquid-like atoms

In Ref. [200], we employ bond-orientational order (BOO) parameter [117, 118] (Subsection 2.2.1.2) to quickly identify the solid-like and liquid-like atoms during the MD simulation, by calculating \( S_{ij} = \sum_{m=-6}^{6} q_{6m}(i)q^{*}_{6m}(i) \) between two neighboring atoms based on the
Steinhardt parameter \( q_{6m}(i) = \frac{1}{N_b(i)} \sum_{j=1}^{N_b(i)} Y_{lm}(\vec{r}_{ij}) \), where \( Y_{lm}(\vec{r}_{ij}) \) is the spherical harmonics and \( N_b(i) \) is the number of nearest neighbors of atom \( i \), two neighboring atoms \( i \) and \( j \) are considered to be connected when \( S_{ij} \) exceeds a threshold (pre-determined in Ref. [203]) and the atoms with 6 connected neighbors are recognized as solid-like. However, the \( \epsilon \)-Al\(_{60}\)Sm\(_{11} \) phase is a non-stoichiometric phase, there are 16\( f \) lattice sites with partially occupation, so the symmetry of a center atom and all the neighboring atoms around it would be ill defined, and we do not have the pre-determined threshold for the Al-Sm system either. So we have to introduce other order parameters, such as the dynamical order parameter, based on its definition of Equation. 2.26, a modified dynamical order parameter is proposed:

\[
\delta_m = \langle \vec{r}^2 \rangle_{\Delta t} - \langle \vec{r} \rangle_{\Delta t}^2
\]  

(5.4)

where \( \vec{r} \) is the position vector of the atom, \( \langle \ldots \rangle_{\Delta t} \) means the average of the sampling positions over a time range \( \Delta t \), in the Al-Sm system, \( \Delta t \) is set to be 1.2 ns. Effectively, it measures the variance of an atom’s trajectory within a time range \( \Delta t \), the atoms with low \( \delta_m \) are labeled as “slow” or “solid-like” atoms, the atoms with high \( \delta_m \) are labeled as “fast” or “liquid-like” atoms.

To determine the \( \delta_m \) threshold, we plot the histogram of \( \delta_m \) in the solid and liquid simulated samples.

- **Solid sample:** Start from the non-stoichiometric \( \epsilon \)-Al\(_{60}\)Sm\(_{11} \) phase which contains 30,672 atoms, first replace the Sm atoms in the 16\( f \) sites by Al atoms randomly, to tune the composition to Al \( \sim \) 10 at.\% Sm. Then perform 100 ns MDMC simulations at 800 K to anneal the sample to the equilibrium state. Finally, anneal the sample at 800 K for 2.4 ns to compute the \( \delta_m \) of each atom.

- **Liquid sample:** First take the solid sample in the last step, melt it at 2,200 K for 50 ns. Then anneal it at 800 K for 50 ns to equilibrium. Finally, anneal the sample at 800 K for 2.4 ns to compute the \( \delta_m \) of each atom.
Both of the simulations use an inter-atomic potential in the Finnis-Sinclair form \[85\], this potential was particularly designed to treat Al-rich alloys (at.% Sm < ∼10%) \[178\].

The histogram of $\delta_m$ is shown in Figure 5.2, the $\delta_m$ of Sm in solid sample is plotted in black, that of Sm in liquid sample is plotted in red, and that of Al in liquid sample is plotted in blue. By comparing the $\delta_m$ histogram of Sm in solid and in liquid samples, we can see the modified dynamical order parameter can identify solid/liquid-like Sm atoms very well, any $\delta_m$ value between $0.4 \text{Å}^2 \sim 1.6 \text{Å}^2$ is capable as a threshold, in the PEM application, we set the threshold to $\delta_m^{th} = 1.0 \text{Å}^2$. By comparing the $\delta_m$ histogram of Sm and Al in liquid samples, we can see the dynamical behaviors of Sm and Al are very different in the supercooled liquid, the Sm is much “slower” with a peak distribution of Sm is at $\delta_m \sim 6.2 \text{Å}^2$, while that of Al is at $\delta_m \sim 9.5 \text{Å}^2$.

Figure 5.2  The histogram of $\delta_m$. 
5.2.2 Embryo of $\epsilon$-Al$_{60}$Sm$_{11}$

The number of embryo atoms $N_0$ should be either too large nor too small, if $N_0$ is too large, we would miss the information of the embryonic stage, if $N_0$ is too small, the advantage of PEM simulation would not be benefit from. We have done an extensive tests and choose $N_0 = 21$ as the embryo, the embryo structure is shown in Figure. 5.3.

![Embryo of $\epsilon$-Al$_{60}$Sm$_{11}$ phase](image)

Figure 5.3 The embryo of $\epsilon$-Al$_{60}$Sm$_{11}$ phase.

In Figure. 5.3 (a), the $\epsilon$-Al$_{60}$Sm$_{11}$ phase is shown, where the red spheres represent Sm atoms, the silver spheres represent Al atoms, and the yellow spheres represent partially occupied sites which can be either Sm or Al atoms. We put Sm atoms in all the partially occupied sites, since the embryo is tiny (123 atoms in total, 21 Sm atoms) compared with the liquid environment it would be immersed in (30,672 atoms in total, $\sim$ at. 10% Sm atoms), this embryo will change little on the system’s composition. Using the face centered Sm atom as the center point, atoms within a radius of 6.98 Å sphere (shaded in light blue) are selected as the embryo, which is shown in Figure. 5.3 (b). Then we insert the embryo into the liquid
sample, and perform the MD simulation at 800 K in NPT ensemble, again, the inter-atomic potential in the Finnis-Sinclair form [85] is used, the MD time-step is chosen to be 2.5 fs, the temperature is controlled using the Nosé-Hoover thermostat [177], the coupling constant for the thermostat (tau) is set to be 0.1 ps and that for the barostat (tauP) is set to be 2.0 ps. The embryo Sm atoms are under an additional PEM potential with the form of Equation. 5.3, in which \( N_{Sm}^0 = 21 \), \( N_{Sm}^{sc} \) is chosen to be \( N_{sc}^{Sm} = 70 \).

### 5.3 Results and Analysis

Figure. 5.4 presents the simulation results. Figure. 5.4 (a) plots the number of solid-like Sm atoms (\( N^{Sm} \)) with respect to the simulation time, Figure. 5.4 (b) shows one snapshot from the slowly growth stage, as labeled by the dashed vertical line in (a). In addition to the dynamical order parameter is used to analyze the solid/liquid-like Sm atoms, the cluster alignment score is also adopted [119, 120], using the \( \epsilon\text{-Al}_{60}\text{Sm}_{11} \) crystal as the template, the details of the cluster alignment score method has been discussed in Subsection. 2.2.1.3. Generally speaking, the dynamical order parameter identify the solid-like Sm atoms from a
dynamical point of view: if an atom moves as slow as those in the solid sample, it is identified as the “solid-like” atom; the cluster alignment score method identify the solid-like Sm atoms from a structural point of view: if an atom is on or close to the $\epsilon$-Al$_{60}$Sm$_{11}$ lattice sites, it is identified as the “solid-like” atom. From Figure. 5.4 (a), we can see the number of solid Sm atoms identified by the cluster alignment method is always lower than that identified by the dynamical order parameter, which means that there are some “slow” Sm atoms which are not in the crystalline phase’s lattice sites. In Figure. 5.4 (b), we studied the morphology of these atoms, all the atoms in (b) are the solid Sm atoms identified by dynamical order parameter, among which only the blue ones are identified as the solid Sm atoms by cluster alignment as well, the red ones are not in or close to the crystalline lattice sites. It need to be mentioned that (b) is a two-dimensional illustration of a three-dimensional morphology, actually all the red atoms are around the blue atoms’ surface.

The detailed nucleation and growth path(s) are still under investigation, but at least we can make the following findings:

- The dynamics of Sm atoms in the Al $\sim$ 10 at.% Sm glasses are very slow. [Figure. 5.2]
- The Sm atoms are the “frame” atoms in the formation of $\epsilon$-Al$_{60}$Sm$_{11}$ phase, once the Sm frame is constructed, non-stoichiometric $\epsilon$ phase would grow. [Figure. 5.4 (a)]
- In the devitrification process of the Al $\sim$ 10 at.% Sm glass to $\epsilon$-Al$_{60}$Sm$_{11}$ phase, the slow Sm atoms could form a large interstitial layer [Figure. 5.4 (b)], this layer plays a negative role in the growth of crystalline phase [Figure. 5.4 (a)]. Different from the free energy barrier, this is a kinetic effect due to the limited diffusivity of Sm atoms.

These findings solidify our conclusion in Chapter. 4, that kinetic effects plays an important role in determining the $\epsilon$-Al$_{60}$Sm$_{11}$ phase during the devitrification process. What is more, the morphology of the interstitial layer due to the kinetic effect is shown.
CHAPTER 6. IMPLEMENTATION OF EAM/FS
POTENTIALS IN HOOMD-BLUE

In all the above chapters (Chapter. 3, 4, and 5), molecular dynamic simulation using classical potential is required, some simulation can be very extensive in time (∼ micro- or even milli- seconds) and/or in problem size (∼ millions of particles). To make these simulations possible, parallel computing must be adopted, in all the works reported in the thesis, the GPU accelerated parallel computing is used.

In this chapter, we present an implementation of EAM and FS interatomic potentials, which are widely used in simulating metallic systems, in HOOMD-blue, a software designed to perform classical molecular dynamics simulations using GPU accelerations. We first discuss the details of our implementation and then report extensive benchmark tests. We demonstrate that single-precision floating point operations efficiently implemented on GPUs can produce sufficient accuracy when compared against double-precision codes, as demonstrated in test simulations of calculations of the glass-transition temperature of Cu_{64.5}Zr_{35.5}, and pair correlation function $g(r)$ of liquid Ni$_3$Al. Our code scales well with the size of the simulating system on NVIDIA Tesla M40 and P100 GPUs. Compared with another popular software LAMMPS running on 32 cores of AMD Opteron 6220 processors, the GPU/CPU performance ratio can reach as high as 4.6. The source code can be accessed through the HOOMD-blue web page for free by any interested user.
6.1 Introduction

Molecular dynamics (MD) [72] is a widely used method for simulating atomic or molecular systems. In MD simulations, the system is often coupled to a thermostat and/or barostat, and the forces on each particle are determined from the potential that describes the interaction among the particles. Then, the trajectories are produced by integrating Newton equations of motion. MD has been a powerful probe to investigate the dynamical evolution of atomic or coarse-grained systems. MD is also regularly applied to systems in equilibrium, where thermodynamic properties are calculated from temporal averages.

While ab-initio methods can be implemented within MD [86,87], they incur a high computational cost, which restricts its application to modest spatial and temporal scales. It is possible to simulate much larger systems (> 1,000 atoms) and longer time (> 1 ns), by including many quantum effects into a simpler classical potential, from which energy and forces are more efficiently calculated. Among these potentials, the embedded atom method (EAM) [84] and the Finnis-Sinclair (FS) model [85], developed almost at the same time, have found great success in describing structural and mechanical properties in metallic systems, including liquid, amorphous, crystalline phases [204–208]. In the EAM model, the total potential energy is expressed as a sum of a pair-wise term and an embedding term that parametrizes many-body effects. The embedding term is essential in describing the mechanical properties of the solid phases [84]. The FS model, on the other hand, was derived from the tight-binding approach based on the second moment approximation [85]. While independently developed, these two potentials share significant similarity in their formulation, and in this paper we present an implementation for both.

Standard MD algorithms can be parallelized on distributed memory clusters through Message Passing Interface (MPI) [209], which remains the dominant communication protocol in high-performance computing. Today, most popular MD codes (including LAMMPS [98],
GROMACS [210], NAMD [211], HOOMD-blue [212]) run very efficiently in clusters. The CUDA language, developed by NVIDIA in 2008, provides a parallel computing platform for programming on Graphic processing units (GPU)s, and has a far-reaching impact on large scientific computing [213]. HOOMD-blue was the first general purpose MD code to perform the simulations entirely on GPUs [99], resulting in a dramatic efficiency boost of two or more orders in magnitude as compared with traditional CPU clusters. Moreover, the object-oriented structure and the python interface of HOOMD-blue makes it easy for developers to add new features, and couple it to the vast library of existing python packages. For example, Anderson et al. [214] designed and implemented a scalable hard particle Monte Carlo simulation toolkit (HPMC) as a package of HOOMD-blue, Spellings et al. [215] implemented a GPU accelerated Discrete Element Method (DEM) molecular dynamics for conservative, faceted particle simulations, French et al. [216] implemented the second-moment approximation to the tight-binding (TB-SMA) potential within HOOMD-Blue. The first attempt to implement EAM/FS potentials in HOOMD-blue was done by Morozov et al. [217]. However, such implementation was never finalized, and the resulting code could only perform a very limited number of calculations, and even in those cases, the implementation was not free of errors. In this chapter, we provide full support for EAM and FS potentials in HOOMD-blue. We first describe all the details of our implementation and provide several benchmarks for speed and accuracy. It is our expectation that this chapter will serve to any future users as a reference.
6.2 Implementation of the EAM in HOOMD-blue

6.2.1 Overview of Alloy and FS models

The total potential energy within EAM [84] and FS [85] potentials can be expressed in the following general form:

\[ E_{tot} = \sum_i F_{\tau_i}(P_i) + \frac{1}{2} \sum_i \sum_{j \neq i} \phi_{\tau_i\tau_j}(r_{ij}) \]  

(6.1)

where \( r_{ij} \) is the distance between atoms \( i \) and \( j \), \( P_i = \sum_{j \neq i} \rho_{\tau_i\tau_j}(r_{ij}) \) is the total electron density at the location of atom \( i \), in which \( \tau_i \) is the type of atom \( i \), \( \rho_{\tau_i\tau_j}(r_{ij}) \) is the contribution to the electron density at the location of atom \( i \) from atom \( j \), the embedding function \( F_{\tau_i}(P_i) \) represents the energy required to “embedded” atom \( i \) into the electron cloud, and the pair function \( \phi_{\tau_i\tau_j}(r_{ij}) \) represents the pair-wise potential energy. To distinguish the EAM potential and the FS potential, we follow a general convention to call the EAM potential “EAM/Alloy” type or “Alloy” potential, and the FS potential “EAM/FS” type or “FS” potential. For the “Alloy” potential, the electron charge density \( \rho_{\tau_i\tau_j}(r_{ij}) \) depends only on atom \( j \)’s type \( \tau_j \), and thus can be reduced to \( \rho_{\tau_j}(r_{ij}) \). While within the “FS” potential, \( \rho_{\tau_i\tau_j}(r_{ij}) \) generally depends on both \( \tau_i \) and \( \tau_j \). The force on the atom \( i \) is calculated according to:

\[ \vec{F}_i = -\vec{\nabla}_{r_i} E_{tot} \]

\[ = -\sum_{j \neq i} \left[ \left( \frac{dF_{\tau_i}(P_i)}{dP} \cdot \frac{d\rho_{\tau_i\tau_j}(r_{ij})}{dr} + \frac{dF_{\tau_j}(P_j)}{dP} \cdot \frac{d\rho_{\tau_j\tau_i}(r_{ij})}{dr} + \frac{d\phi_{\tau_i\tau_j}(r_{ij})}{dr} \right) \frac{\vec{r}_i - \vec{r}_j}{r_{ij}} \right] \]  

(6.2)

6.2.2 Implementation

HOOMD-blue can run on both CPU and GPU, which is controlled by setting “--mode=cpu” or “--mode=gpu” in HOOMD-blue command line options. When the CPU mode is turned on, an object of EAMForceCompute class is instantiated to perform all the computations
of energy and forces exclusively on the CPU. One should note that the unparallelized CPU mode runs very slow; thus, it is mainly for debugging purpose and not for any real applications. By default, the GPU mode is switched on, and the computation is handled by the EAMForceComputeGPU class, which is parallelized using the CUDA language.

For either the CPU or GPU mode, the potential file is loaded by the member function loadFile in the EAMForceCompute class. The potential files follow the DYNAMO setfl format [218]. For the “FS” potential file, there are \( n \) density function blocks for each atom type \( \tau_j \), corresponding to \( \rho_{\tau_1 \tau_j}(r), \rho_{\tau_2 \tau_j}(r), \ldots, \rho_{\tau_n \tau_j}(r) \), where \( n \) is the total number of atom types in the alloy system. However, for the “Alloy” potential, there is only one such block since the function \( \rho_{\tau \tau_j} \) depends only on \( \tau_j \). Here, in order to make the computing consistent and easy to maintain, the type “Alloy” potential’s \( \rho \) block is duplicated for \( n \) times, so that the “Alloy” and “FS” potentials can be treated in the same fashion in the rest of the code. In the python script, the “type” keyword should be set to “Alloy” or “FS” according to the type of the potential file.

Once the tabulated electron density \( \rho \), embedding function \( F \), pair function \( \phi \) are read, their cubic interpolation parameters of these functions are computed and stored in the memory, so that the function value at any arbitrary data point can be efficiently calculated. Additional discussions about the interpolation scheme can be found in the next subsection. Then the algorithm goes to the energy and force computing. If the CPU mode is activated, the energy and forces are computed using member function computeForces of class EAMForceCompute, which computes Equation 6.1 and Equation 6.3, by looping over all atoms. If the default GPU mode is activated, the potential data is copied and allocated from the memory to the GPU texture memory variables in class EAMForceComputeGPU. To compute forces on GPU, class EAMForceComputeGPU calls two CUDA kernels defined in the CUDA source file EAMForceGPU.cu sequentially. The first kernel gpu_kernel_1 computes \( P_i \) of the atom recorded at the corresponding thread, and the second kernel gpu_kernel_2
computes the force of the atom recorded at the corresponding thread. By dividing the computation to two kernels sequentially, the total electron densities computed by kernel `gpu_kernel_1` are guaranteed to be synchronized before the kernel `gpu_kernel_2` is executed.

Figure 6.1 shows the flowchart of our algorithm, the matrix blocks in `gpu_kernel_1` and `gpu_kernel_2` represent two GPU threads, which can be executed in parallel. `gpu_kernel_1` block and `gpu_kernel_2` block are executed sequentially, because the total electron density $P_i$ must be synchronized before computing the forces.

For the requirement of further maintenances, a unit test is created using NumPy’s test support, which computes the potential energy and forces of a Ni$_3$Al crystalline system, the relative tolerance between the test value and the reference value is set to be $10^{-6}$ eV/atom and $1.6 \times 10^{-14}$ N/atom for energy and forces, respectively. Users will be alerted if the unit test fails. If that happens, users may need to check the software and hardware prerequisites, refer to the compiling manual, and/or change the building options according to software and hardware configurations, following HOOMD-blue’s documentation [219].

A Python module named `hoomd.metal` is provided following HOOMD-blue’s code convention, users can import the `hoomd.metal` module to load the support of EAM/FS potentials. Users can create the Python script, to custom simulation routines for controlling the simulations’ initial conditions, interactions and other parameters, together with Python’s powerful features, instructions and commands, which is very flexible. This version of our code is mainly developed on a single-GPU architecture, currently only the serial version of the EAM/FS implementation is supported. As we will demonstrate in Section 6.3 and Section 6.4, the serial code runs very efficiently for systems containing up to tens of thousands of atoms, which can address many research needs. For this reason, we decided to publish the serial code first. Recently, we have acquired resources, and are working on the parallel version of the code using MPI.
Figure 6.1 The flowchart of EAM/FS energy and force computing algorithm.
6.2.3 Interpolation

In practice, the embedding function $F(\rho)$, electron density function $\rho(r)$ and pair interaction function $\phi(r)$ are provided in a tabulated format, and an interpolation scheme is needed to calculate the energy and the forces at those points where the tabular values are not available. Although NVIDIA GPUs provide a very efficient linear interpolation at the hardware level, which can be around $2 \sim 3$ times faster than the software cubic interpolation, it has only 9 bits of internal precision, which might result in hard-to-foresee problems in the targeted application of highly accurate and reliable metallic simulations. So the hardware linear interpolation does not follow HOOMD-blue style conventions. In our implementation, the slower but safer software cubic interpolation is used.

6.2.4 Precision

Most NVIDIA CUDA enabled GPUs have much better performance in single precision floating point operations than double precision operations according to the CUDA toolkit documentation [115]. For example, on GPUs with compute capability (cc) 5.2 (e.g. Tesla M40), 128 single precision operations of addition, multiplication and multiplication-addition can be completed per clock cycle per multiprocessor, whereas that number for double precision operations is only 4. These GPUs are essentially not designed to perform double precision tasks. On the other hand, for GPUs of cc=3.5 (e.g. Tesla K40), cc=3.7 (e.g. Tesla K80), and cc=6.0 (e.g. Tesla P100), the efficiency ratio between single and double precision operations ranges from 3:1 to 2:1. Thus, double precision operations could be a viable choice on these GPUs. Our code supports both single and double precisions in the GPU mode. Considering the fact that at this moment GPUs of cc=5.0, 5.2 are still widely used, we need to particularly investigate the accuracy of our code in the single precision mode. A broad range of tests have been made, which prove that the simulation in single precision mode
can deliver accurate enough results. An elaborate discussion about these tests will be made in Section 6.3. For cc=5.0, 5.2, 3.0, 3.2 GPUs, users are recommended to use the single precision mode for metallic simulations by turning the SINGLE\_PRECISION flag on.

### 6.3 Accuracy testing

We have applied the newly developed EAM and FS code within HOOMD-blue on various real problems using the single precision mode, and its accuracy is checked against the CPU version of LAMMPS (MPI, 16 Feb 2016 stable version), which implements the standard double-precision computing. LAMMPS is chosen because it has been overwhelmingly the dominant software package for simulations with EAM/FS potentials. The broad tests include simulations with EAM and FS potentials, under the canonical (NVT) and the isothermal-isobaric (NPT) ensembles, which are the two most commonly used ensembles in MD simulations. We discuss two of our tests in this section, as summarized in table 6.1. In these simulations, the MD time-step is chosen to be 2.0 fs, the temperature is controlled using the Nosé-Hoover thermostat [177], the coupling constant for the thermostat (tau) is set to be 0.1 ps and that for the barostat (tauP) is set to be 2.0 ps, the neighbor list is reconstructed for every 5 MD time-step, and the buffer radius of neighbor list (r\_{buff}) is set to be 1.0 Å. In LAMMPS, users can set the length of thermostat chain and that on barostat by setting tchain and pchain values respectively, in the tests below, these values are set by default, which are tchain = 3 and pchain = 3.

<table>
<thead>
<tr>
<th>system</th>
<th>number of atoms</th>
<th>process</th>
<th>potential type</th>
<th>ensemble</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cu\textsubscript{64.5}Zr\textsubscript{35.5}</td>
<td>5000</td>
<td>rapid cooling</td>
<td>eam/fs</td>
<td>NPT</td>
</tr>
<tr>
<td>Ni\textsubscript{3}Al</td>
<td>4000</td>
<td>liquid annealing</td>
<td>eam/alloy</td>
<td>NVT</td>
</tr>
</tbody>
</table>
6.3.1 Rapid cooling of Cu$_{64.5}$Zr$_{35.5}$ alloy

Rapid cooling from metallic liquids is a widely used metallurgical technique for synthesizing novel alloys with fine-grained or amorphous microstructures. Under sufficiently fast cooling rate, nucleation of a crystalline phase in the liquid alloy can be kinetically blocked. As a result, the system is driven into a glassy state. Although the liquid-to-glass transition is usually not considered a phase transition since the glassy state is not a thermodynamical stable phase, the specific heat and the thermal expansion coefficient may display a discontinuity at the transition temperature, due to the different characters of $\alpha$ and $\beta$ relaxations in the liquid and glassy states, respectively [220,221].

![Figure 6.2 Dependence of the potential energy on temperature by HOOMD-blue and LAMMPS during the cooling process of Cu$_{64.5}$Zr$_{35.5}$ systems.](image)

In our tests, we simulate the rapid cooling of a strong binary glass former Cu$_{64.5}$Zr$_{35.5}$ [222–225]. The unit cell contains 5,000 atoms with periodic boundary conditions. The system is thermalized at 2000 K for 10 ns, and then cooled down to 500 K at a cooling rate of $10^{10}$ K/s. The NPT ensemble is used throughout the simulation. The energy and forces
are calculated with a Cu-Zr potential in the FS format [226] during the simulations. Our results are compared against the EAM LAMMPS implementation, using the exact the same initial configuration (including the same positions and velocities).

In Fig. 6.2, we show the potential energy per atom of the system as a function of the temperature during the cooling process. The glass transition occurs at the same temperature $T_g \sim 750$ K, as shown in the shaded area. HOOMD-blue is run on one NVIDIA M40 GPU (red), and LAMMPS is run on one AMD Opteron 6220 CPU with 32 cores (black). Results using HOOMD-blue and LAMMPS are indistinguishable within statistical error. The observed kink at $T \sim 750$ K corresponds to the glass-transition temperature $T_g$, this is consistent with the estimates from prior work [226].

6.3.2 Thermalising of Ni$_3$Al alloy

![Figure 6.3](image-url)

Figure 6.3 pair correlation functions $g(r)$ of Ni$_3$Al alloy at 1700 K by HOOMD-blue and LAMMPS are in agreement within statistical accuracy.
Nickel aluminide has been widely used for high-temperature coatings in turbine blades and jet engines due to its low weight, high strength and melting point and good chemical stability [227]. The Ni$_3$Al alloy can be prepared by melting in a high frequency vacuum induction furnace [228]. Understanding the structure of the Ni$_3$Al melt is essential in controlling the properties of the alloy as a coating material.

In this subsection, we calculate the pair correlation function $g(r)$ of Ni$_3$Al liquid at 1700 K, using a previously developed EAM potential for this system [229]. As shown in Figure. 6.3, the HOOMD-blue and LAMMPS are in agreement within statistical error. HOOMD-blue runs on one M40 GPU in Maxwell cluster in single precision (red), and LAMMPS runs on 32 cores of AMD Opteron 6220 CPU (black).

### 6.3.3 Accuracy testing summary

A series of other tests using EAM/FS potentials in the single precision mode have been made, with results showing that the calculated physical properties (energy, volume, short to medium range order, diffusion, etc.) are consistent within statistical error. The tests demonstrate that not only the code we developed is robust and accurate, but also the single precision mode is accurate enough for simulations of metallic systems using EAM/FS potentials.

### 6.4 Performance benchmarks

A series of Ni$_3$Al crystalline samples with sizes ranging from 4,000 atoms to 62,500 atoms are simulated using the canonical (NVT) ensemble to study the performance of our code. The simulation parameters are not changed from those in Section 6.3, and the EAM potential introduced in Section 6.3.2 [229] is used for energy and force calculations. The samples are annealed at 300 K for 1 ns (500,000) steps, based on which we can calculate the average
computing time required to perform one MD time-step (TRPT). The relative GPU to CPU performance ratio is computed by comparing TRPT collected for HOOMD-blue on a single GPU and LAMMPS on 32 CPU cores. The details of the computing systems and software used for benchmarks are as follows:

- **Ames Lab “Maxwell” GPU cluster**: 1 Tesla M40 (cc=5.2) is used
  CPU: Intel Xeon E5-2698v4, 2.20 GHz; GPU: NVIDIA Tesla M40; RAM: 512 GB; Software: HOOMD-blue; OS: CentOS 7.

- **NVIDIA “PSG” GPU cluster**: 1 Tesla P100 (cc=6.0) is used
  CPU: Intel Haswell E5-2698v3, 2.30GHz; GPU: NVIDIA Tesla P100; RAM: 256 GB; Software: HOOMD-blue; OS: CentOS 7.

- **Ames Lab “Brem12” CPU cluster**: 32 MPI processes in 2 nodes (32 cores) are used
  CPU: AMD Opteron(TM) Processor 6220, 3.00 GHz; RAM: 64 GB; Number of node: 12 nodes; Number of cores per node: 16 cores; Software: LAMMPS (MPI, 16 Feb 2016 stable version); OS: CentOS 6.

The HOOMD-blue is compiled with GCC-4.8.5, which is the minimum compiler requirement in the manual of HOOMD-blue [219], and NVIDIA CUDA-8.0 in the “Maxwell” and “PSG” clusters, and the LAMMPS (MPI) is compiled with GCC-4.8.5 and Open MPI 1.8.8. HOOMD-blue simulation in double precision is only tested on NVIDIA “PSG” GPU cluster’s Tesla P100 GPU (cc=6.0). LAMMPS performs the same simulations using Message Passing Interface (MPI) in a multi-processor CPU cluster “Brem12”. The results reported in here are based on runs on 32 cores. It should be noted that LAMMPS uses 64-bit double precision operations on CPUs.

TRPT is presented in Figure 6.4 as a function of the number of atoms for typical MD simulations’ problem size ($4,000 \leq N_{\text{atoms}} \leq 62,500$) in EAM/FS-type potentials: $\text{Ni}_3\text{Al}$
crystal is simulated using EAM potential on different computational systems: HOOMD-blue is run on one M40 GPU in Maxwell cluster in single precision (red square), on one P100 GPU in PSG cluster in single precision (blue triangular), on one P100 GPU in PSG cluster in double precision (blue circle), and LAMMPS is run on 32 cores of AMD Opteron 6220 CPU (black plus). Lines represent linear fitting. The lower the TRPT is, the faster the simulation runs. A quasi-linear dependence of TRPT on the system size can be observed at \( N > N_c \) on GPUs, where \( N_c \sim 10,000 \) for M40 and \( N_C \sim 13,000 \) for P100. The performance is worse when \( N < N_c \), showing the GPU is underloaded when the system is not large enough. On the CPU cluster, the TRPTs shows linear scaling with the problem size for all sizes.

![Figure 6.4](image-url)  
Figure 6.4 Computing time required to perform a single MD time-step (TRPT) as a function of the number of atoms in the system.

Although the architecture of a CPU cluster and a GPU node is vastly different, a comparison of computing speed on typical CPU and GPU based hardware can still show useful information about the acceleration effect of GPUs. In Figure 6.5, we show the GPU/CPU performance ratio, defined as the inverse ratio of TRPT, as a function of system size, com-
putational details are the same as those of Figure. 6.4. Again, TRPT on GPU is collected from a single NVIDIA M40 or P100 GPU, and TRPT on CPU is collected from 32 cores of AMD Opteron 6220 CPUs. When the system is underloaded for the GPU, a moderate acceleration is observed with the GPU/CPU performance ratio ranging from 1.5 to 2.5. For fully loaded systems, the GPU/CPU performance ratio can be as high as 4.6 and 4.1 for single and double precision operations, respectively. P100 also performs significantly better than M40, because P100 has more CUDA cores (3,584) than M40 has (3,072), and the base clock frequency of P100 (1,328 MHz) is much higher than that of M40 (948 MHz).

![Figure 6.5](image.png)

Figure 6.5 The GPU/CPU performance ratio as a function of the number of atoms in the system.

By analyzing the EAM/FS formulas for the energy and force calculations in Equation. 6.1 and Equation. 6.3, respectively, it is found that the computational cost for the embedding part and the pair part is comparable during energy calculation (see Equation. 6.1), the computational cost for the embedding part is much higher than that of the pair part during force calculation (7 operations v.s. 1 operation, see Equation. 6.3). To estimate the actual
time cost on the embedding functions as compared with the pair functions in EAM/FS potentials, we also perform benchmark on a simple Lennard-Jones (LJ) system, which only contains pair functions. A binary “A\textsubscript{3}B” Lennard-Jones (LJ) system is generated in such a way that the positions for minima of A-A, A-B, and B-B pair potentials match the positions of Ni-Ni, Ni-Al, and Al-Al pair correlation functions in Figure. 6.3, respectively. The $r_{\text{cut}}$ for LJ potentials is set to 6.28721 (Å), which is the same as that of the Ni-Al EAM potential [229]. The same set of parameters are also used in MD simulations using the LJ potential and the EAM potential. In this way, one can obtain a fair comparison of the computational cost for LJ and EAM potentials.

![Graph showing the computing time required to perform a single MD time-step as a function of the number of atoms for EAM and LJ systems.](image)

Figure 6.6 Computing time required to perform a single MD time-step as a function of the number of atoms for EAM and LJ systems.

The TPRT results for the EAM potential and the LJ potential are shown in Figure. 6.6: HOOMD-blue is run on one M40 GPU in Maxwell in single precision: Ni\textsubscript{3}Al crystal is simulated using EAM potential (red square), and A\textsubscript{3}B particles system is simulated using Lennard-Jones potential (black diamond). The EAM runs 6 $\sim$ 8 times slower than LJ on
an M40 GPU, which indicates that the actual cost of calculating the embedding function part consumes $83 \sim 88\%$ computational time of the EAM and FS implementation. This is consistent with direct analytic results from Equation 6.1 and 6.3.

### 6.5 Conclusion

We have implemented the support of the EAM and FS potentials as a package in the HOOMD-blue software, under the BSD 3-clause license. The source code is available for download in HOOMD-blue website [230]. As a package of HOOMD-blue, our code follows all the HOOMD-blue conventions, so users could refer to HOOMD-blue documentation page [219] for the software/hardware prerequisites, building and installing method, and code descriptions. We have also included unit tests of our package, which help users to test their builds.

The accuracy of the code has been verified in a variety of broad tests, among which, two tests: computing the glass transition temperature of strong glass former $\text{Cu}_{64.5}\text{Zn}_{35.5}$ and the pair correlation function of the $\text{Ni}_3\text{Al}$ liquid are shown in this chapter. All tests using this code give consistent results compared to those using LAMMPS within statistical errors.

The performance of the code has been studied in a practical way, by an overall consideration of efficiency and user friendliness. As shown in Figure 6.4 and Figure 6.5, the code performs significantly faster than LAMMPS running on a typical CPU cluster. For the user interface, our hoomd.metal module follows HOOMD-blue code convention, which allows it to be coupled to the extensive python libraries.
CHAPTER 7. SUMMARY AND FUTURE PLAN

The development of materials is the driving force in the civilization, people have always committed to make new materials to meet the requirements of the human society. The fabrication, and the understanding of a material construct a positive two-step feedback loop, after thousands of years of continuous exploration, a large number of knowledge about the fabrication methods, structures, and properties of thermodynamically stable materials have been known. Meanwhile, meta-stable materials, and non-stoichiometric materials have drawn people’s attention in two perspectives: one is the application, some meta-stable non-stoichiometric materials possess unique, better properties; the other is the understanding, these materials often exhibit complex structures and complicated phase selections, whose driving mechanism is not clearly known. Take the Al-Sm system shown in Figure 1.2 as an example again, we can see that for the same composition of metallic glass Al$_{90}$Sm$_{10}$, it can undergoes such diverse and distinct phase selection paths depending on the fabrication processes, novel meta-stable phases can be generated accordingly. Our goal is to have a comprehensive understanding of the nature of these novel phases and the phase transitions, in order to predict and control them. Motivated by this goal, my thesis study uses high performance computing technology, together with theory and experiment, investigates the structural and dynamical properties during the phase selection in the Al-Sm alloy. We summarize the thesis study works as follows:

1. We have established a self-contained algorithm to rigorously evaluate the free energy for solid and liquid phases of an alloy system, based on thermodynamic integration, which has good agreements with experimental data. The significance of this work lies
in (1) it can be used to determine the driving force for nucleation in supercooled liquid, which is fundamental to achieving microscopic understandings of freezing and melting phenomena; (2) it is self-contained, the accuracy only depends on the accuracy of classical inter-atomic potential. This is the foundation of our understandings on phase selection.

2. We have elucidate the kinetic effect, especially the limited diffusivity of Sm atoms’ role in the meta-stable non-stoichiometric $\epsilon$-Al$_{60}$Sm$_{11}$ phase forming. In Chapter. 4, Monte Carlo simulation with cluster expansion free energy model, and molecular dynamics simulation with “bootstrap” accumulation method make it possible to simulate samples whose (effective) size can be comparable to experimental ones; In Chapter. 5, the persistent-embryo method (PEM) opens a new avenue to study solidification under realistic experimental conditions via atomistic computer simulation. These methods link the simulation and experiment, provide a convincing scheme to verify our understandings of phase selection.

3. We have also developed GPU based, Pythonic molecular dynamics software package: the HOOMD-blue EAM/FS package. Our work is expected to significantly expand the scope of the HOOMD-blue software, which was traditionally used in the community of soft condensed matter physics. And its Python interface makes it more convenient to perform the simulation and analyze the results. This provide a powerful tool in understanding the phase selection.

The above works have yielded complete or preliminary achievements, the establish of various algorithms, models, methods and software packages lays foundations for future research. But of course, this thesis study is just one small step further to our goal, to have a deeper understanding of the phase selections in meta-stable metallic alloys, a lot of follow-up tasks need to be completed. These tasks include: (1) Use the free energy evaluation
data of our procedure as an descriptor, compared with experimental data, to refine the classical EAM/FS potential. (2) Incorporate the structural and dynamical properties of Sm atoms we have found in Chapter. 4 and Chapter. 5 into the coarse-grain model or phase field model, to simulate the phase selection in macroscopic scale. (3) Investigate the effect of non-stoichiometric persistent-embryo, to identify all the very early nucleation pathways and their associated probabilities. (4) MPI parallel the HOOMD-blue EAM/FS package, to perform multi-GPU computing in order to simulate larger samples.
BIBLIOGRAPHY


in Co2ZnO4 and Rh2ZnO4,” *Physical Review B - Condensed Matter and Materials Physics*, vol. 84, no. 6, pp. 1–11, 2011.


APPENDIX. PUBLICATIONS, MANUSCRIPTS AND REPORTS

My publications, manuscripts and reports during the Ph.D. study are listed:

Publications and Manuscripts


Reports
