Multiscale modeling of homo- and heterogeneous system

Liu, Chao

2012

Liu, C. (2012). Multiscale modeling of homo- and heterogeneous system. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/51053>

<https://doi.org/10.32657/10356/51053>

Multiscale Modeling of Homo- and Heterogeneous System

Liu Chao

School of Chemical and Biomedical Engineering

A thesis submitted to Nanyang Technological University

In partial fulfillment of the requirement for the degree of

Doctor of Philosophy

2012

Acknowledgements

Praise is a debt that we owe unto the virtue of others. So I would like to begin by recording my gratitude to my advisor Prof. Lim Kok Hwa, and ex-advisor Prof. Kwak Sang Kyu and Prof. Santosh Ansumali not only for their instructive guidance to my project, but also for their continuous encouragement and support during this long process. We had a lot of useful discussions during the last few years. Prof. Kwak Sang Kyu gave me a lot of constructive suggestions and Prof. Santosh Ansumali gave me many helpful instructions and insightful views on how to build a readable and beautiful code, how to carry out core computer experiments and how to get crucial and useful literatures and books as well.

I am also grateful to my group members who have always been together with me throughout. I have enjoyed myself all the time when discussing with Huang Huancong, Chen Wenwen and Wahyu. They are very good companions in the laboratory.

Moreover, I shall express my gratitude to the school, for their providing the cluster for costive computation, and to Jawaharlal Nehru Centre for Advanced Scientific Research (JNCASR) in India for their supplying the short visit program. Last but not the least, I shall express my thanks to my family. Thank them for their perpetual support during my pursuit of PhD.

Table of Contents

Acknowledgements.....	I
Table of Contents	II
List of Figures	V
List of Tables.....	IX
Abstract.....	X
Chapter 1 Introduction	1
1.1 Molecular Dynamics.....	1
1.2 Language for Scientific Computing.....	2
1.3 Etomica	3
1.4 Monte Carlo	4
1.5 An Introduction to Direct Simulation Monte Carlo.....	5
1.6 Motivation and Objectives.....	7
Chapter 2 Fundamentals of Polyvacancy.....	9
2.1 Solid defects.....	9
2.1.1 Point defect	9
2.1.2 Other kinds of defects	12
2.2 Ensemble.....	12
2.2.1 Grand canonical ensemble	13
2.2.2 Isothermal-isobaric ensemble	13

2.3 Structure of crystalline materials	14
2.4 Number of vacancies	16
2.5 Wigner-Seitz cell	17
2.6 Single Occupancy Cell Method	18
Chapter 3 Kinetic Theory.....	19
3.1 Basic Concepts.....	20
3.1.1 The Hard Sphere Model.....	20
3.1.2 Mean Free Path	21
3.1.3 Mach Number	23
3.1.4 Collision Interval [74].....	23
3.1.5 Knudsen Number	24
3.1.6 The Grad-Boltzmann Limit [74].....	25
3.2 The Boltzmann Equation	26
3.3 The Chapman-Enskog Theory	29
3.4 Boundary Conditions	31
3.4.1 The Periodic Wall Boundary Condition	31
3.4.2 The Specular Wall Boundary Condition.....	31
3.4.3 The Diffuse Wall Boundary Condition [75]	32
3.5 Probabilistic Simulation Methods.....	32
Chapter 4 Theoretical calculation of polyvacancy.....	35
4.1 Theoretical calculation of monovacancy	35
4.1.1 Calculation of chemical potential μ	37
4.1.2 Calculation of the monovacancy free energy f_1	40

4.1.3 Calculation of monovacancy concentration.....	42
4.2 Theoretical calculation of divacancy	43
4.2.1 Calculation of adjacent vacancy free energy f_2	43
4.2.2 Calculation of divacancy concentration.....	46
4.3 Theoretical calculation of trivacancy.....	47
4.3.1 Trivacancy configurations for FCC and HCP crystals	47
4.3.2 Insertion probabilities of trivacancies in FCC and HCP.....	50
4.3.3 Deletion probabilities of trivacancies in FCC and HCP.....	53
4.3.4 Trivacancy concentrations of FCC and HCP.....	55
Chapter 5 Computational Experiment of DSMC.....	79
5.1 General Procedure of DSMC.....	80
5.1.1 DSMC initialization.....	80
5.1.2 DSMC time loop.....	83
5.2 Couette flow.....	91
5.3 Consistent Boltzmann Algorithm	93
5.4 Modified Enskog Method	99
Chapter 6 Conclusions and Outlook	108
6.1 Conclusions.....	108
6.2 Outlook	109
References.....	110
Publication List	122

List of Figures

FIG. 2.1 VACANCY POINT DEFECT	10
FIG. 2.2 INTERSTITIAL POINT DEFECT	11
FIG. 2.3 IMPURITY POINT DEFECT.....	11
FIG. 2.4 BODY-CENTRED CUBIC[65]	15
FIG. 2.5 FACE-CENTRED CUBIC[65]	15
FIG. 2.6 HEXAGONAL CLOSE PACKED[65].....	15
FIG. 2.7 THE CONSTRUCTION OF WIGNER-SEITZ CELLS	17
FIG. 3.1 COLLISION CROSS SECTION OF HARD SPHERE MOLECULES	21
FIG. 3.2 MOLECULES IN A COLLISIONAL TUBE	21
FIG. 4.1 COMPARISON OF SIMULATION AND THEORETICAL VALUES OF PRESSURE FOR MONOVACANCY.....	38
FIG. 4.2 PLOT OF MONOVACANCY CONCENTRATION AS A FUNCTION OF NUMBER DENSITY	43
FIG. 4.3 ILLUSTRATION OF DIVACANCY SIMULATION WITHOUT PARTICLES.....	44
FIG. 4.4 GUI FOR DIVACANCY SIMULATION	45
FIG. 4.5 PLOT OF DIVACANCY CONCENTRATION AS A FUNCTION OF NUMBER DENSITY ...	47
FIG. 4.6 THE 4 TYPES OF FCC TRIVACANCIES: FCC3.1, FCC3.2, FCC3.3 AND FCC3.4, CORRESPONDING TO ANGLES OF 180° , 120° , 90° AND 60° , RESPECTIVELY. THE DARK CIRCLES REPRESENT THE VACANCY SITES. THE LINES IN FCC3.4 SHOW THAT THE THREE VACANCY SITES FORM AN EQUILATERAL TRIANGLE.	48
FIG. 4.7 THE 9 TYPES OF HCP TRIVACANCIES: HCP3.1, HCP3.2, HCP3.3, HCP3.4, HCP3.5, HCP3.6, HCP3.7, HCP3.8, HCP3.9. THE BLACK AND WHITE CIRCLES REPRESENT THE VACANCY SITES. THE THREE ROWS, FROM TOP TO BOTTOM, CORRESPOND TO TRIVACANCIES WITH 0, 2 AND 3 SITES IN THE SAME PLANE, RESPECTIVELY.	49

FIG. 4.8 (COLOR ON LINE) THE SYSTEMS A) (I.E. $M^{(0)}$), B) (I.E. $M^{(1)}$), C) (I.E. $M^{(2)}$), AND D)

(I.E. $M^{(3)}$) REPRESENT THE SYSTEMS OF MONO-, DI-,TRI-, AND TETRA-VACANCIES, RESPECTIVELY. THE SYSTEM A) IS THE SUBSYSTEM OF THE SYSTEM B), WHICH BECOMES THE SUBSYSTEM OF THE SYSTEM C), AND SO ON. LATTICE SITES ARE NOT EXPLICITLY SHOWN, BUT IT IS ASSUMED THAT VACANCIES AND PARTICLES (I.E. GRAY BACKGROUND) ARE LOCATED AT LATTICE SITES. [REPRINTED WITH PERMISSION FROM REF. [76]. COPYRIGHT 2011, TAYLOR & FRANCIS]..... 59

FIG. 4.9 (COLOR ON LINE) SCHEMATIC DIAGRAMS TO DEVELOP THE MONOVACANCY SYSTEM ($A \rightarrow B$), THE DIVACANCY SYSTEM, ($C \rightarrow D$), AND THE TRIVACANCY SYSTEM

($E \rightarrow F$), WHICH ARE PRESENTED WITH EACH TOTAL FREE ENERGY F_n , RESPECTIVELY. THE LATTICE SITES ARE OCCUPIED BY PARTICLES UNLESS EMPTIED BY VACANCIES. [REPRINTED WITH PERMISSION FROM REF. [76]. COPYRIGHT 2011, TAYLOR & FRANCIS] 62

FIG. 4.10 (COLOR ON LINE) SCHEMATIC OF TRIVACANCIES IN FCC WITH SPREAD ABC PLANES AND HCP WITH SPREAD ABA PLANES. FILLED AND OPEN CIRCLES

REPRESENT POSSIBLE VACANT SITES IN TRIVACANCY. THE FAR-LEFT PLANE (WITH THREE FILLED CIRCLES) IS AT THE TOP OF THE STACK, THE PLANE B (WITH SIX FILLED CIRCLES) IS IN THE MIDDLE, AND THE RIGHT PLANE (WITH THREE OPEN CIRCLES) FROM IT IS THE BOTTOM ONE. ON THE RIGHT HAND SIDE, TRIVACANCIES ARE SHOWN WITH THE MIDDLE PLANE AS A REFERENCE PLANE. IF FILLED CIRCLES ARE NOT ON VERTICES OF HEXAGON, THOSE ARE IN THE TOP PLANE. OPEN CIRCLES, WHICH ARE ON VERTICES OF HEXAGON, ARE IN THE BOTTOM PLANE. [REPRINTED WITH PERMISSION FROM REF. [76]. COPYRIGHT 2011, TAYLOR & FRANCIS] 69

[REPRINTED WITH PERMISSION FROM REF. [76]. COPYRIGHT 2011, TAYLOR & FRANCIS] 71

FIG. 4.13 (COLOR ON LINE) PLOT OF $-\log x$ VS. THE COMPRESSIBILITY FACTOR,

PV_0/NkT FOR THE HCP HS CRYSTAL. BENNETT AND ALDER IS REF. [68]. LINES ARE

GUIDES TO AN EYE. ERRORS ARE SMALLER THAN SYMBOLS.[REPRINTED WITH PERMISSION FROM REF. [76]. COPYRIGHT 2011, TAYLOR & FRANCIS].....	75
FIG. 4.14 (COLOR ON LINE) PLOT OF THE FORMATION FREE ENERGY OF TRIVACANCY $-f_3$ VS. THE DIMENSIONLESS DENSITY ρ^* FOR THE HCP HS CRYSTAL. LINES ARE GUIDES TO AN EYE. ERRORS ARE SMALLER THAN SYMBOLS. REPRINTED WITH PERMISSION FROM REF. [76]. COPYRIGHT 2011, TAYLOR & FRANCIS]	76
FIG. 4.15(COLOR ON LINE) PLOT OF AVERAGED $-\log x$ VS. THE COMPRESSIBILITY FACTOR, PV_0/NkT FOR BOTH FCC AND HCP HS CRYSTALS. BENNETT AND ALDER IS REF. [68]. LINES ARE GUIDES TO AN EYE. ERRORS ARE SMALLER THAN SYMBOLS.[REPRINTED WITH PERMISSION FROM REF. [76]. COPYRIGHT 2011, TAYLOR & FRANCIS]	77
FIG. 5.1 THE FLOWCHART OF MOVE SUBROUTINE	84
FIG. 5.2 PSEUDO-CODE FOR SORT SUBROUTINE	86
FIG. 5.3 THE VELOCITY PROFILE.....	91
FIG. 5.4 THE TEMPERATURE PROFILE.....	92
FIG. 5.5 THE NUMBER DENSITY PROFILE.....	92
FIG. 5.6 VELOCITY CORRELATION FUNCTION.....	93
FIG. 5.7 SCHEMATIC DIAGRAM OF THE DISPLACEMENT TERM [56]	94
FIG. 5.8 SHEAR VISCOSITY AS A FUNCTION OF NUMBER DENSITY (WITH $KT = M = \Sigma = 1$, $N^* = N\Sigma^3$).....	96
FIG. 5.9 SELF-DIFFUSION AS A FUNCTION OF NUMBER DENSITY (WITH $KT = M = \Sigma = 1$, $N^* = N\Sigma^3$).....	97
FIG. 5.10 HS SELF-DIFFUSION COEFFICIENT NORMALIZED BY MD SELF-DIFFUSION COEFFICIENT AS A FUNCTION OF NUMBER DENSITY (WITH $KT = M = \Sigma = 1$, $N^* = N\Sigma^3$) .	104
FIG. 5.11 HS SHEAR VISCOSITY NORMALIZED BY MD SHEAR VISCOSITY AS A FUNCTION OF NUMBER DENSITY (WITH $KT = M = \Sigma = 1$, $N^* = N\Sigma^3$)	105
FIG. 5.12 PRESSURE AS A FUNCTION OF NUMBER DENSITY (WITH $KT = M = \Sigma = 1$, $N^* = N\Sigma^3$)	106

FIG. 5.13 SINGLE PARTICLE VELOCITY CORRELATION FUNCTION (SVCF) AS A FUNCTION OF TIME IN LOG-LOG PLOT	106
---	-----

List of Tables

TABLE 2.1 NUMBER OF VACANCY CLUSTER CONFIGURATIONS [67]	16
TABLE 4.1 DATA OF SIMULATION VALUE OF PRESSURE FOR PERFECT CRYSTAL	37
TABLE 4.2 DATA OF PRESSURE FROM SPEEDY'S EQUATION OF STATE FOR MONOVACANCY.....	38
TABLE 4.3 DATA OF CHEMICAL POTENTIAL FOR MONOVACANCY	40
TABLE 4.4 DATA OF FREE ENERGY F_1	42
TABLE 4.5 DATA OF FREE ENERGY F_2	45
TABLE 4.6 CHARACTERISTICS OF THE 9 TYPES OF HCP TRIVACANCIES [65].....	49
TABLE 4.7 DATA OF SUBVOLUME FOR TRIVACANCY SIMULATION.....	50
TABLE 4.8 INSERTION PROBABILITIES OF 4 DIFFERENT FCC TRIVACANCIES AT DIFFERENT DENSITIES	51
TABLE 4.9 INSERTION PROBABILITIES OF 9 DIFFERENT HCP TRIVACANCIES AT DIFFERENT DENSITIES	52
TABLE 4.10 DELETION PROBABILITIES OF 4 DIFFERENT FCC TRIVACANCIES AT DIFFERENT DENSITIES	53
TABLE 4.11 DELETION PROBABILITIES OF 9 DIFFERENT HCP TRIVACANCIES AT DIFFERENT DENSITIES	54
TABLE 4.12 ACTUAL VALUES OF TRIVACANCY CONCENTRATIONS FOR ALL TYPES IN FCC AND HCP HS CRYSTALS AT VARIOUS DENSITIES.....	72
TABLE 5.1 DATA OF SHEAR VISCOSITY	96
TABLE 5.2 DATA OF SELF-DIFFUSION COEFFICIENT	97

Abstract

As a way of solving physical problems with important features at multiple scales, multiscale modeling is widely used in engineering, physics, meteorology, computer science and so forth. Multiscale simulation is to be attempted by using a combination of Molecular Dynamics (MD) and Direct Simulation Monte Carlo (DSMC) and which represent micro- and meso-scale in dimension, respectively.

In this thesis, we introduce two different projects in different scales. In the first project, we study the physical and thermodynamic properties of crystals with defects. In particular, vacancies and their effects to the materials are our interest. Etomica, which is a Java-based open source package, is used to simulate atom-based crystals of interest. A statistical ensemble method, which was developed by Pronk and Frenkel for monovacancy, has been modified to calculate different degree of vacancies. In the second project, we try to solve the problems in a fluid flow with the DSMC algorithm which is widely used in a rarefied gas. A modified DSMC method is attempted to obtain correct transport properties in a gas flow to create a new equation of state for a real gas. In this regard, this work contains two creative points. The first is that one can control to study the degree of vacancy from mono-to-poly so that the extreme case of void in crystal can be studied properly. The second is to break the traditional inconsistency that DSMC yields transport properties for a real gas yet has an ideal gas equation of state, so that it recovers the exact hard sphere (HS) equation of state.

Chapter 1 Introduction

1.1 Molecular Dynamics

Molecular dynamics (MD) is a principal tool for molecular simulation studies. MD requires atoms and molecules to interact for a period of time by approximations of known physical principles or laws, giving view of the motion of the particles.

In general molecular systems contain a huge number of particles, it is often impossible to calculate the properties of such complex systems analytically. When the number of interacting particles is larger than two, the result is a chaotic motion [1-5], which is related to n-body problem. MD simulation circumvents the analytical intractability by using numerical methods. It represents an interface between laboratory experiments and theory, and can be understood as a "virtual experiment". MD probes the relationship between molecular structure, movement and function.

In many aspects, MD simulations are similar to real experiments in terms of actual setting. It follows the same approach as what is used in real experiments. First, we prepare a sample: we select a model system, which consists of N particles and we solve Newton's equations of motion by such as Velocity-Verlet, Predictor-Corrector, etc. for this system until the expected properties of the system do not change much with time. After the system reaches equilibrium, we perform the actual measurement. To measure an observable quantity in a MD simulation, first of all expression of this observable is required as a function of positions and momenta of particles in the system [6].

1.2 Language for Scientific Computing

The FORTRAN language emerges in 1957. Traditionally it has been used for the solution of large scientific problems together with BASIC. They are inappropriate for structured programming and have a low readability to common audience. In recent years Object Oriented Programming (OOP) [7-9] has emerged as the most preferable computational technique in software industry due to its more realistic structure of data storage, portability, and code maintainability and reusability. Thus, C++ [10, 11] and Java [12-14] become the mainstream tools for large program or application development for most programmers.

The choice between Java and C++ is subtle. Java dates back to 1990s. Java as a closest to human language is a pure Object Oriented (OO) language, and everything it does has to be related to objects. It has been chosen by a lot of programmers for its simplicity, excellent usability, applicability for multithreaded programming and fairly powerful capability at Graphical User Interface (GUI). In spite of its virtues, Java program often runs slower compared to C++, and lacks pointers, which makes the programmer difficult to allocate and track the memory effectively. C++, however, is a free-form and general-purpose language. Although it is considered as a middle-level language as it comprises a combination of both high-level and low-level language features (i.e. the extension of C language inherits lots of characteristics from C), C++ is still very good at classes, virtual functions [15], operator overloading [16], multiple inheritance [17], templates and exception handling [18]. Despite its complexity C++ is a more efficient language

compared to Java in term of speed. In a word, the choice of programming language depends on demand. If the programmer needs graphics with a slower simulation run, Java is recommended. Without reason for the graphics, C++ should be a better choice. In our case, we use C++ for DSMC flow project and Java for Polyvacancy one.

1.3 Etomica

Etomica [19-22], developed by Prof. David Kofke and his group at the state university of New York at Buffalo, is a Java API (Application Programming Interface) and development environment for construction and implementation of molecular simulations. It is a GUI designed to provide point-and-click access to the simulation tools contained in the Molecular Simulation API. The API is a set of classes that can be employed to construct a molecular simulation in Java. Simulations may be constructed with them by programming in Java with a text editor, or by using the Etomica development environment. The development environment provides a graphical interface for constructing, running, and saving simulations based on the API. The Etomica environment provides for dynamic modification and interrogation of the simulation, while it is in progress [23]. The Etomica GUI has a lot of benefits. It provides the user with a number of functions. First and foremost, its environment empowers the users with a way to interact with Molecular Simulation API without having too much experience or knowledge of Java programming. Second, by having a GUI that allows real time modification of simulation conditions, users can steer simulations effectively in response to observations made while they are running. Finally, the Etomica supplies the users an

image of the results from the simulation. This gives users visual feedback that complement the usual simple logs of data. This feedback (e.g. clustering of molecules, inflection points in graphs), as well as, outlier in data become visibly apparent almost immediately before lengthy data analysis starts [24].

1.4 Monte Carlo

Monte Carlo (MC) [25-29] is another principal tool for molecular simulation, which uses probabilistic simulation method. Unlike MD, MC requires random numbers to be generated to obtain the simulation results.

After the world's first computer ENIAC [30-32] was invented in 1946, Polish-American mathematician Stan Ulam along with many other scientists was duly impressed by its speed and versatility. However, Stan's extensive mathematical background made him aware that statistical sampling techniques had fallen into desuetude because of the length and tediousness of the calculations. But with this miraculous development of the ENIAC—along with the applications Stan must have been pondering—it occurred to him that statistical techniques should be resuscitated, and he discussed this idea with von Neumann. Thus was triggered the spark that led to the Monte Carlo method [33].

In the early 1950s, electronic computers were no longer predominantly used in military. It was partially used in other fields and this was the beginning of the discipline of computer simulation. When the Los Alamos MANIAC [6] became operational in March 1952, Metropolis was interested in having as broad a spectrum of problems as possible

tried on the machine, in order to evaluate its logical structure and demonstrate the capabilities of the machine.

Metropolis used the idea of underlying matrix of Markov chain [26, 34-37] in his algorithm, which merely contained two stages. This became the world's first Monte Carlo (MC) algorithm and it set up the foundations for subsequent Monte Carlo research. Nowadays, Monte Carlo is widely used in lots of areas, such as physical sciences, designs and visuals, finance and business, telecommunications, games, mathematics and so forth [38-42].

MC methods are a class of computational algorithms, which rely on repeated random samplings to calculate the required results. It is frequently used in simulating physical and mathematical systems. Because of their reliance on repeated computation of random or pseudo-random numbers, the quality of random number is very important. These algorithms are most suitable for calculation by a computer and tend to be used when it is infeasible or impossible to compute an exact result with a deterministic algorithm. In the MC method, it often requires calculate the average number, so sampling technique is very important.

1.5 An Introduction to Direct Simulation Monte Carlo

The primary concern of computational fluid dynamics (CFD) is the development and application of numerical recipes to solve the partial differential equations (PDE) in fluid mechanics [43]. These equations, most notably the Navier-Stokes and Euler equations,

describe Newtonian fluids, for instance, gases and simple liquids, over a wide range of conditions. Although very useful, the continuum description of a fluid has its limit. For example, the flow of a dilute gas requires a kinetic-theory description [44]. Furthermore, Boltzmann equation is a highly nonlinear integro-differential equation and it is quite difficult to obtain an analytical solution. Under such circumstances, various specialized methods for stimulating such flows have been developed.

The direct simulation Monte Carlo (DSMC) method was proposed by Prof. Graeme Bird [45-47], emeritus professor of aeronautics, university of Sydney. DSMC uses probabilistic (i.e. MC) simulation to solve the Boltzmann equation [48-52] for finite Knudsen number fluid flows. It is identical to the Boltzmann equation in nature and usually comprises four subroutines: move, sort, collide and sample. The move subroutine is just to move all the particles to new positions according to their old positions and velocities; the sort subroutine is to sort all the particles in different cells which are already divided before the program starts, and build cross index between particles and cells; the collide subroutine is to make the particles collide within each cell, and it is equivalent to the collision term in the Boltzmann equation; the sample subroutine is to sample all the required data according to statistical methods.

Since its first non-trivial application in 1964, the DSMC method has become one of the standard and reliable approaches to rarefied gas flow problems [53]. Its simulation has been extensively adopted to evaluate the macroscopic characteristics of rarefied gas flows [54].

1.6 Motivation and Objectives

The DSMC algorithm is essentially a particle-based, Lagrangian numerical scheme for solving time-dependent, nonlinear Boltzmann equation for hard sphere (HS) model in dilute gas regime. This scheme resembles MD's particle interaction, but differs in that DSMC generates stochastic collisions within each predetermined cell with rates and probabilities calculated from the kinetic theory. However, DSMC confronts with inconsistency, i.e. it results in transport properties for a dilute HS gas with diameter σ while it obtains ideal gas equation of state (EOS) which implies $\sigma=0$. In its original formulation, DSMC was restricted to dilute gases, yielding only an ideal gas equation of state (EOS) [55]. Some modifications are made to DSMC by Alexander et al [56]. They summarized the DSMC algorithm as two steps of advection and collision. During the advection step, they added a displacement term in order to obtain the exact hard sphere EOS at all densities. This leads to consistent thermodynamic and transport properties in the low density regime. This consistent Boltzmann algorithm (CBA) removes the previous inconsistency of the traditional DSMC and runs faster than MD at low and moderate densities, and recovers HS equation of state to some extent [56]. Nevertheless, the transport properties (i.e. shear viscosity and self-diffusion) obtained at high densities have large discrepancies from the Enskog theory. Therefore, we try to construct an additional term in DSMC to make it consistent with the Enskog theory even at high densities.

In terms of vacancy research, our focus is to build a right passage to study individual high order vacancy in crystal. So far, mono-and divacancy were successfully studied by using

Pronk and Frenkel's method [6], but not for polyvacancy, which is the origin of void development. This is mostly because polyvacancy concentration in crystal is very low. In this thesis, we extend their approach to tri- and higher order of vacancy. We attempt to figure out the thermodynamic properties of the vacancies, develop the method to calculate them easily by knowing their movement and behavior, and grasp the effect of the vacancies to solid properties and phase variation. Finally, determination of the criteria of the smallest void is followed in the course work with study of the role of high order vacancies for creation and propagation crack.

Chapter 2 Fundamentals of Polyvacancy

2.1 Solid defects

A perfect crystal is a crystal with all particles located in their respective sites and orientation throughout the crystalline material without any defects. However, it does not exist in nature. Real crystals always contain defects, which affect their properties. These imperfections, in many situations, may result from deformation and lattice mismatch of the solid, rapid cooling from high temperature, high-energy radiation (i.e. X-ray or neutrons) striking the solid and etc. From a thermodynamic point of view, solid defects are the results of thermodynamic equilibrium contributed by the increase in entropy TS term of the Gibb's free energy:

$$\Delta G = \Delta H - T\Delta S, \quad (2.1)$$

and there are four main types of solid defects: point defects, line defects, plane defects and volume defects. These defects have great influence on the electrical and mechanical properties of crystalline materials.

2.1.1 Point defect

A point defect [57-59] is an empty site, in which an atom does not exist. The point defect has several types, which mainly stand for vacancies, interstitials and impurities. They

may be introduced initially in the crystal growth process, by thermal or mechanical stress into the crystal or as the result of radiation, or electric current, etc. A typical vacancy in 2D plane is shown in Fig. 2.1. The vacancies are observed to move around in solid materials and this phenomenon is known as vacancy diffusion, in which one of the particles surrounding the vacant lattice site exchanges places with a vacancy, resulting in the movement of vacancy observed under microscopic view [60, 61]. By successive jump of atoms into vacant sites, it is possible for a vacancy to move within the crystal, thus it plays an important role in diffusion of atoms through the lattice.

Fig. 2.1 Vacancy point defect

Another type of point defect is interstitial as shown in Fig. 2.2. It is possible for some atoms to fall into the interstitial space between lattice positions, which are not occupied by other atoms. Interstitials usually tend to push surrounding atoms apart and produce distortion of the lattice plane. Interstitial atoms may be caused by severe local distortion during plastic deformation or irradiation. Interstitial particles in solid mixture can easily diffuse through crystal structure via vacant space among the particles.

Fig. 2.2 Interstitial point defect

An impurity is the substitution of a regular atom with another type of atom that does not normally occupy that lattice site as shown in Fig. 2.3. The atom may come from within the crystal, (e.g. a chlorine atom on a sodium site in NaCl crystal) or from the addition of impure chemicals.

Fig. 2.3 Impurity point defect

2.1.2 Other kinds of defects

Line defects, one-dimensional defects, are associated primarily with mechanical deformation [62]. Line defects are also known as the dislocation. It occurs when slips in the crystal plane are caused by an external force. There are various types of plane defects, including twin boundary, grain boundary and tilt boundary. Twin boundary, a highly symmetrical discontinuity in structure, separates two crystalline regions, which are structurally mirror images of each other. Grain boundary is the interface between two adjacent single crystals or grains, which have different orientations. It disrupts the motion of dislocations in material. Tilt boundary is caused when two adjacent grains are tilted. Volume defects are three-dimensional. This defect is much larger than the above-mentioned defects. Most defects usually affect the periodicity of crystal over a certain range of distances from the defect. Nonetheless, local property of materials could be strengthened or weakened by defects.

2.2 Ensemble

In statistical mechanics, the ensemble is an idealization consisting of a large number of replicas of a small system, each of which represents a possible state, that the real system might be in. There are several ensembles, such as canonical, microcanonical, grand canonical, isothermal-isobaric, isoenthalpic-isobaric ensembles and etc. Since vacancies in crystals affect to vary the empty lattice sites and the number of particles in the system, the grand-canonical ensemble is a suitable choice to work along with isothermal-isobaric

ensemble to set up the statistical mechanical description on the thermodynamic properties of crystal with point defects.

2.2.1 Grand canonical ensemble

The grand canonical ensemble is an imaginary collection of identical systems, where each system is in equilibrium with an external reservoir with respect to both particle and energy. It is an extension of canonical ensemble. In this ensemble, chemical potential μ , volume V and temperature T are fixed so that the number of particles, pressure, and energy are fluctuated. It is used in open systems.

2.2.2 Isothermal-isobaric ensemble

The isothermal-isobaric ensemble is a statistical mechanical ensemble that maintains N , temperature T and pressure P as constants. NPT ensemble is of importance generally as in laboratory systems or real life where the mass, temperature and pressure are easily measured and are usually assumed as closed systems with mass, temperature and pressure held constant.

2.3 Structure of crystalline materials

The structure and properties of materials are determined mainly by its constituent chemical elements and the path of manufacturing process. Scientists have made a lot of attempts to understand the basics of materials by relating the particle structures and manufacturing process through thermodynamic laws. Solid materials are made up of atoms, molecules or ions, particles in general, differing in packing arrangements [63]. Solid can be classified as crystalline and non-crystalline materials according to material science. Properties of some materials are directly related to the positioning of particles [64].

In crystalline materials, particles are located at lattice points, which are the regular periodic three dimensional arrays of points in space. A crystal structure is formed when particles pack themselves in repetitive 3D lattice position. The crystal structure can be divided into small repeat units known as the unit cell, which is the most basic pattern of crystal. Types of crystals can be classified by their type of the unit cell, examples being body-centred crystal (BCC), face-centred crystal (FCC) and hexagonal close packed crystals (HCP), as shown from Fig. 2.4 to Fig. 2.6, respectively.

Fig. 2.4 Body-Centred Cubic[65]

Fig. 2.5 Face-Centred Cubic[65]

Fig. 2.6 Hexagonal close packed[65]

The type of unit cell is usually related to the material strength and ductility. For instance, FCC metals like aluminum, copper, and lead have lower strength but higher ductility [65]. HCP metals like magnesium and zinc tend to be less ductile than FCC metals [66].

2.4 Number of vacancies

Most vacancy-related research focuses on monovacancy or divacancy. Note that divacancy is composed of monovacancy and adjacent monovacancy. As the number of vacancy order increases, the number of arrangement of vacant sites in vacancy cluster increase but the concentration decreases. The number of vacancy cluster configurations can be counted through simulations considering all the possible vacancy combinations. Table 2.1 presents the number of configurations in a vacancy cluster for four different crystal structures.

Table 2.1 Number of vacancy cluster configurations [67]

Crystal Type		SC	BCC	FCC	HCP
N	1	1	1	1	1
	2	1	1	1	2
	3	2	3	4	9
	4	7	12	20	57
	5	23	61	131	458

Note: *N* refers to the number of vacancies in a vacancy cluster. SC, simple cubic; BCC, body-centered cubic; FCC, face-centered cubic; HCP, hexagonal close-packed.

In some simulations, more than one type of vacancy is simulated. For example, Bennett and Alder simulated the system of 5 monovacancies, where each vacancy can diffuse and merge into different types of clusters[68]. Therefore, at any point of time, there might be 5 monovacancies, 2 monovacancies and a trivacancy, or a divacancy and a trivacancy etc.

2.5 Wigner-Seitz cell

The Wigner-Seitz (WS) cell is a geometrical construction used in the study of crystals. As an intrinsic nature, the WS cell around a lattice point is defined as the locus of points in space that are closer to that lattice point than to any other lattice points. In order to construct the WS cell, we have to first pick up a lattice point, then draw the lines to all nearest lattice points, and draw the perpendicular bisector planes of these lines. The volume enclosed by these planes is called the WS cell as shown is Fig. 2.7.

Fig. 2.7 The construction of Wigner-Seitz cells

The WS cell represents the volume, within which a particle is confined, and is related to the probability of the microstate, in which a particle is contained in a certain volume. The larger the volume of the WS cell, the greater the probability, and in turn the greater the entropy of the particle, which results in low free energy.

2.6 Single Occupancy Cell Method

The Single Occupancy Cell (SOC) method is used to confine the movement of particles within a cell around their lattice sites. If the constraints are not imposed, the lattice sites are consistent with the average position of the particles. If the density is high enough, SOC method will not have effect on the result of the predicted crystal properties. The SOC constraint was applied throughout our simulations. This ensures that the lattice retained its structure at all densities, and prevented the diffusion of vacancies for the appropriate track of individual vacancies.

Chapter 3 Kinetic Theory

Gases can be studied by either small scale movement of individual molecules or large scale action of the gas as a whole. We can directly measure the large scale action of the gas. But in order to study microscopic action of each individual molecule, we have to use a theoretical method. The kinetic theory of gases was the first successful theory explaining the macroscopic irreversibility through a statistical point of view. It assumes that the size of molecules is very small relative to the distance between them. The molecules are in frequent motion and constantly collide with each other and with walls of the container they reside in. An individual molecule possesses the basic physical properties like mass, energy and momentum. The density of a gas is just the sum of the mass of the molecules divided by the volume. The pressure is a measure of the rate of change of momentum of the molecules. As the particles collide with the walls of a container, the molecules transfer momentum on the walls, yielding a force, which could be measured. The pressure we need is statistically the total force of all the molecules divided by the area. The temperature is a measurement of the mean kinetic energy of the gas. The molecules are in continuous random motion, and there is an energy, $(1/2)mv^2$, associated with that motion.

3.1 Basic Concepts

We introduce the definitions of some basic concepts in the kinetic theory such as mean free path, collision interval, Knudsen number, Boltzmann equation, etc. These concepts will be use in Section 5.1.

3.1.1 The Hard Sphere Model

The hard sphere (HS) model is widely used as modeling particles in the statistical mechanical theory of fluids. HS is defined as impenetrable sphere that cannot overlap with each other in space. They imitate the extremely strong repulsion that atoms and spherical molecules experienced at a contact distance. Particles of HS of diameter d follow the interaction potential of the following pairwise function:

$$V(\mathbf{r}_1, \mathbf{r}_2) = \begin{cases} 0 & \text{if } |\mathbf{r}_1 - \mathbf{r}_2| \geq d \\ \infty & \text{if } |\mathbf{r}_1 - \mathbf{r}_2| < d \end{cases}, \quad (3.1)$$

where $V(r_1, r_2)$ denotes the potential of the two particles, and $\mathbf{r}_1$ and $\mathbf{r}_2$ denote the position vectors of the two particles. When two particles with diameter d_1 and d_2 collide, the radius of the effective collision area is $r = (d_1 + d_2)/2$, and the total collision cross section is

$$\sigma_T = \pi r^2. \quad (3.2)$$

The collision cross section is shown in Fig. 3.1.

Fig. 3.1 Collision cross section of hard sphere molecules

3.1.2 Mean Free Path

One of the basic concepts in the kinetic theory is the mean free path, which is the smallest length scale for describing molecular motion. It is defined as the average distance traveled by a molecule between two successive collisions. For accurate calculation, we imagine a thought experiment where in a collisional tube a molecule of diameter d and velocity $\bar{v}$ travels through a sea of other molecules. (See Fig. 3.2)

Fig. 3.2 Molecules in a collisional tube

We suppose that molecules are identical. During time t this particle will sweep a volume $\pi d^2 \bar{v} t$. If we ignore collisions involving more than two molecules (it proves to be feasible, see the Grad-Boltzmann limit at section 3.1.6), the rest collisions will be the binary collision between this molecule and the rest molecules that it meets during its travel, i.e., the estimated number of collisions in time t is number of gas molecules inside this volume: $n_v \pi d^2 \bar{v} t$, where n_v is the number density of the particles. The distance traveled by the molecule is $\bar{v} t$. Then we can compute the mean free path λ according to the definition as

$$\lambda = \frac{\bar{v} t}{n_v \pi d^2 \bar{v} t} = \frac{1}{n_v \pi d^2}. \quad (3.3)$$

The problem with this expression is that the average molecular velocity is used, but it ignores that the target molecules are also moving. The frequency of collisions depends upon the average relative velocity of the randomly moving molecules. What is needed is the average relative velocity, and the calculation of that velocity from the molecular speed distribution yields the result, thus the average relative velocity is

$$\overline{v_{rel}} = \sqrt{2} \bar{v}, \quad (3.4)$$

which revises the expression for the effective volume swept out in time t as $\sqrt{2} n_v \pi d^2 \bar{v} t$.

The resulting mean free path is then

$$\lambda = \frac{1}{\sqrt{2} n_v \pi d^2}. \quad (3.5)$$

3.1.3 Mach Number

Mach number is defined as the ratio of the speed of object moving through any substance to the speed of sound traveled in that substance. It is usually used to represent an object's relative speed (i.e. such as space shuttle or missile) to sound's speed. Its formulation is expressed as follows:

$$Ma = \frac{v_o}{v_s}, \quad (3.6)$$

where Ma is the Mach number, v_o is the velocity of object (relative to the medium), and v_s is the velocity of sound in the medium. Especially, the speed of sound is expressed as follows:

$$v_s = \sqrt{\frac{\gamma \cdot k_B \cdot T}{m}}, \quad (3.7)$$

where γ is the adiabatic index, the ratio of specific heat of a gas at constant pressure to specific heat at constant volume (C_p/C_v), k_B is the Boltzmann constant, T is the absolute temperature in Kelvin, and m is the mass of a single molecule.

3.1.4 Collision Interval [69]

The collision interval is defined as the average time spent between two successive collisions, in terms of the mean molecular velocity $\bar{\xi} = \sqrt{2k_B T/m}$ as:

$$\tau \approx \frac{\lambda}{\bar{\xi}} \sim \frac{\nu}{v_s}, \quad (3.8)$$

where ν denotes the kinematic viscosity. The collision interval τ represents the smallest time scale present in the kinetic theory, and v_s represents the isentropic sound speed.

3.1.5 Knudsen Number

Knudsen number is a number used to describe the flow of a low density gas, equal to the ratio of the mean free path to the characteristic length, such as boundary layer thickness, or apparatus dimension. It can be expressed as:

$$Kn = \frac{\lambda}{L}, \quad (3.9)$$

where L denotes the characteristic length. Knudsen number is commonly used to describe the extent to which the gas behave like a collection of independent particles (i.e. free-molecule regime, Knudsen number much larger than unity) or like a viscous fluid (i.e. continuous regime, Knudsen number much smaller than unity). Intermediate regimes are termed as transition region, and slip flow region. The Knudsen number defines the importance of the molecular time scale. The definition in terms of time scales is

$$Kn = \frac{\tau \bar{\xi}}{L} \sim \frac{\tau v_s}{L}. \quad (3.10)$$

Another important expression to relate Knudsen number, Mach number and Reynolds number is

$$Kn = \frac{Ma}{Re} \sqrt{\frac{\gamma\pi}{2}}, \quad (3.11)$$

where Re denotes Reynolds number, $Re = \rho UL/\mu$, the ratio of inertial force and viscous force.

3.1.6 The Grad-Boltzmann Limit [69]

One of the important assumptions made in the kinetic theory of dilute gases is that only pair collisions are important. The mathematical limit, where this is a valid assumption and the Boltzmann equation holds, is known as Grad-Boltzmann limit. This is defined as the limit when N tends to infinity, d_m tends to zero, Nd_m^3 (i.e. the volume occupied by the particles) tends to zero, while Kn (and thus Nd_m^2) remains constant. The limit of the volume occupied per particle going to zero, while the mean free path remains constant, gives a quantitative measure for rarefaction, and thus a well-defined domain of validity for rarefied gas dynamics. In this limit, one aspect of the intermolecular forces, namely departure from the ideal gas law, disappears while the effect of collisions on the time evolution of the gas remains finite and manifests itself at the macroscopic level as transport coefficients, for instance, viscosity, self-diffusion, thermal conductivity and so forth.

3.2 The Boltzmann Equation

The Boltzmann equation, also known as the Boltzmann transport equation, makes a description of the statistical distribution of one particle in a fluid. It is one of the most important equations in non-equilibrium statistical mechanics, the area of statistical mechanics that deals with systems far from the thermodynamic equilibrium. The Boltzmann equation is used to study how a fluid transports physical quantities such as heat and charge, thus help to derive transport properties such as viscosity, self-diffusion and thermal conductivity. Before introducing the Boltzmann equation, we define $f(\mathbf{x}, \mathbf{c}, t) d\mathbf{x} d\mathbf{c}$ is the probability of finding a molecule in the differential volume $d\mathbf{x} d\mathbf{c}$ about $\mathbf{x}$ and $\mathbf{c}$. The Boltzmann equation describes a rarefied gas by considering two processes: convection and collision. The equation reads

$$\frac{\partial f}{\partial t} + \frac{\partial f}{\partial \mathbf{x}} \cdot \mathbf{c} + \frac{\partial f}{\partial \mathbf{c}} \cdot \frac{\mathbf{F}}{m} = \left. \frac{\partial f}{\partial t} \right|_{coll}, \quad (3.12)$$

where f is the one-particle distribution function, $\mathbf{F}$ is the external force exerted on each particle, and $\left. \frac{\partial f}{\partial t} \right|_{coll}$ is the collision term, which describes the effect of binary collisions between particles on the single particle distribution function.

The collision term is the most complicated one in the Boltzmann equation and it mostly treats the binary collision. During the process of a binary collision, independent of the specific intermolecular forces, particles undergo the laws of energy conservation and momentum conservation, which can be expressed in the following equations:

$$m_1 \mathbf{v}_1 + m_2 \mathbf{v}_2 = m_1 \mathbf{v}_1' + m_2 \mathbf{v}_2', \quad (3.13)$$

$$\frac{1}{2} m_1 v_1^2 + \frac{1}{2} m_2 v_2^2 = \frac{1}{2} m_1 v_1'^2 + \frac{1}{2} m_2 v_2'^2, \quad (3.14)$$

where m_1 and m_2 represent the masses of the two particles, $(\mathbf{v}_1, \mathbf{v}_2)$ represent the pre-collisional velocities and $(\mathbf{v}_1', \mathbf{v}_2')$ represent the post-collisional velocities. If we assume they have equal mass, Eqs. (3.13) and (3.14) can be simplified as:

$$\mathbf{v}_1 + \mathbf{v}_2 = \mathbf{v}_1' + \mathbf{v}_2', \quad (3.15)$$

$$v_1^2 + v_2^2 = v_1'^2 + v_2'^2. \quad (3.16)$$

It is obvious that from these two equations there is a complete symmetry between $(\mathbf{v}_1, \mathbf{v}_2)$ and $(\mathbf{v}_1', \mathbf{v}_2')$. If $\mathbf{v}_1$ and $\mathbf{v}_2$ are given with $\mathbf{v}_1'$ and $\mathbf{v}_2'$ to be determined, we have four equations (i.e. three from Eq. (3.15) because $\mathbf{v}_1$ and $\mathbf{v}_2$ are vectors, and one from Eq. (3.16)) with six unknowns. To this end, two more parameters are required to solve these equations. Here we introduce the unit vector $\boldsymbol{\alpha}$ in the direction of the change in velocity $\mathbf{v}_1$ and A as its magnitude

$$\mathbf{v}_1' - \mathbf{v}_1 = A \boldsymbol{\alpha}. \quad (3.17)$$

In order to eliminate A and obtain an explicit solution in terms of $\boldsymbol{\alpha}$, which is the unit vector along the line joining the centers of the particles for spherical particles, substituting Eq. (3.17) into (3.15) gives

$$\mathbf{v}_2' - \mathbf{v}_2 = -A \boldsymbol{\alpha}. \quad (3.18)$$

Hence we have $\mathbf{v}_1' = \mathbf{v}_1 + A$ and $\mathbf{v}_2' = \mathbf{v}_2 - A$. Substitute them into Eq. (3.16),

$$\mathbf{v}_1'^2 - \mathbf{v}_2'^2 = (\mathbf{v}_1 + A)^2 + (\mathbf{v}_2 - A)^2, \quad (3.19)$$

provided that $A \neq 0$, expansion and simplification of Eq. (3.19) gives

$$A = \alpha \mathbf{V}, \quad (3.20)$$

where

$$\mathbf{V} = \mathbf{v}_2 - \mathbf{v}_1. \quad (3.21)$$

By substituting Eq. (3.21) into Eqs. (3.17) and (3.18), we can obtain

$$\left. \begin{aligned} v_1' &= v_1 + \alpha (\alpha \bullet V) \\ v_2' &= v_2 - \alpha (\alpha \bullet V) \end{aligned} \right\}. \quad (3.22)$$

Our desired solution has finally emerged. However, this parametrization of the solution of Eqs. (3.15) and (3.16) is not a one-to-one relation. The great circle $\alpha \mathbf{V} = 0$ corresponds to a single solution (i.e. identity). All other solutions are covered twice, once by each hemisphere $\alpha \mathbf{V} < 0$ and $\alpha \mathbf{V} > 0$, whereas if we insist on the definition of α as having the direction $\mathbf{v}_1' - \mathbf{v}_1$, we must have

$$\alpha \bullet (\mathbf{v}_1' - \mathbf{v}_1) > 0. \quad (3.23)$$

This removes the ambiguity and makes the parametrization a one-to-one relation except for $\alpha \mathbf{V} = 0$. We remark that Eq. (3.22) represents a linear transformation of the six-space $(\mathbf{v}_1, \mathbf{v}_2)$ into itself. The inverse transformation can be easily found, thus we have

$$\mathbf{V}' = \mathbf{v}_2' - \mathbf{v}_1' - 2 \left(\frac{1}{2} (\mathbf{v}_1' + \mathbf{v}_2') \cdot \mathbf{V} \right) \mathbf{V} = -2 \left(\frac{1}{2} (\mathbf{v}_1' + \mathbf{v}_2') \cdot \mathbf{V} \right) \mathbf{V}, \quad (3.24)$$

from which

$$\mathbf{V} \cdot \mathbf{V}' = - \frac{1}{2} (\mathbf{v}_1' + \mathbf{v}_2') \cdot \mathbf{V}. \quad (3.25)$$

By plugging Eq. (3.25) into (3.22) we find

$$\left. \begin{aligned} \mathbf{v}_1' &= \mathbf{v}_1 + \mathbf{V}' \\ \mathbf{v}_2' &= \mathbf{v}_2 - \mathbf{V}' \\ \mathbf{V}' &= \mathbf{v}_2' - \mathbf{v}_1' \end{aligned} \right\}. \quad (3.26)$$

The Boltzmann equation is used in Chapter 5.

3.3 The Chapman-Enskog Theory

In statistical mechanics, the Maxwell–Boltzmann distribution describes particle speeds in gases, where the particles move freely without interacting with one another, except for very brief elastic collisions in which they may exchange momentum and kinetic energy, but do not change their respective states of intramolecular excitation, as a function of the temperature of the system, the mass of the particle, and speed of the particle. It is a probability distribution for the speed of a particle constituting the gas - the magnitude of its velocity vector meaning that for a given temperature, the particle will have a speed selected randomly from the distribution, but is more likely to be within one range of some speeds than others.

The Chapman-Enskog theory [47] provides a type of solution to the Boltzmann equation for a restricted set of problems, in which the distribution function is slightly perturbed from the equilibrium Maxwellian form. It assumes that the distribution function f can be expressed in the form of power series:

$$f = f^{(0)} + \varepsilon_0 f^{(1)} + \varepsilon_0^2 f^{(2)} + \dots, \quad (3.27)$$

where ε_0 is a parameter, which can be regarded as a measure of either the mean collision time or Knudsen number. The first term $f^{(0)}$ is the Maxwellian distribution f_0 for an equilibrium gas, thus an alternative form of the expression is

$$f = f_0(1 + \Phi_1 + \Phi_2 + \dots). \quad (3.28)$$

The equilibrium distribution function constitutes the known first-order solution of this equation and the second order solution requires the determination of the parameter Φ_1 .

Solutions of that particular Boltzmann equation is expressed as follows:

$$f = f_0(1 + \Phi_1), \quad (3.29)$$

which was obtained by Enskog and Chapman independently, and these equations formulate the subject matter of the classical work by Chapman and Cowling [44]. For a simple gas, Φ_1 depends only on the density, stream velocity and temperature so that the resulting solution constitutes a normal solution of the Boltzmann equation. The Chapman-Enskog theory is used in Section 5.3 and Section 5.4.

3.4 Boundary Conditions

Boundary condition is an important aspect in computational physics. When particles move to the edge of the domain, boundary conditions will be applied. There are three types of common boundary conditions: periodic wall boundary condition, specular wall boundary condition and diffuse wall boundary condition. These will be used in Section 5.1.2.

3.4.1 The Periodic Wall Boundary Condition

The periodic wall boundary condition is the simplest case in all the boundary conditions. Suppose we have a domain $a \leq x < b$, if a particle moves a distance σ beyond the lower bound of the domain to the position $x = a - \sigma$, this amounts to the particle moving to the new position $x = b - \sigma$. Similarly, if a particle moves a distance σ beyond the upper bound of the domain to the position $x = b + \sigma$, it amounts to the particle moving to the new position $x = a + \sigma$. All particle velocities remain unchanged.

3.4.2 The Specular Wall Boundary Condition

In the specular wall boundary condition, each particle which reaches the edge of the domain is reflected with perfectly elastic collision dynamics. The kinetic energy of the particle keeps unchanged while the sign of the normal component of the velocity is

reversed. This is the computational representation of the specular wall boundary condition.

3.4.3 The Diffuse Wall Boundary Condition [70]

The diffuse wall boundary condition is the most realistic and meanwhile the most complicated of the three models. When a particle reaches the edge of the domain, it will be given a new velocity magnitude and direction, which results in the change in the macroscopic velocity as well as the internal energy. To implement the diffuse wall boundary condition, the first step is to update the position of the particle according to its previous position and velocity. Then two tests are run to determine whether or not the particle is still in the computational domain. If the particle is found to have left the domain, the particle is moved backward along its previous path onto the appropriate boundary and the amount of time it took the molecule to return to the boundary is computed.

3.5 Probabilistic Simulation Methods

The earliest probabilistic simulation method originated in 1962. It was introduced by Haviland and Lavin, termed as the test particle Monte Carlo Method [47]. In this method an initial estimate has to be made of the distribution function over the whole flow field. With the pre-assumed distribution a great number of test particle trajectories are

calculated. The target distribution is then updated in time from the history of the test or incident molecules. This process is continued until the target distribution is reached.

An alternative approach to the test particle method is the direct simulation Monte Carlo (DSMC) method, which introduces a time variable and tracks the trajectories of a large number of simulated molecules. This probabilistic method is usually applicable to dilute gas flows and was first applied to the homogeneous gas relaxation problem [47]. The essential part of DSMC algorithm is that molecular motion and intermolecular collision are uncoupled over a certain time step. In DSMC, the region of interest is divided into numerous spatial cells, which are then divided further into certain sub-cells. Usually the cell size should be smaller than mean free path, otherwise the error of the computational results will be unacceptable. As long as this condition is satisfied, the computational result doesn't depend on the cell size. When the cell size is close to zero, DSMC becomes the exact solution of Boltzmann equation. In this sense, DSMC is the numerical solution of the time-dependent Boltzmann equation. The time step should be smaller than mean collision time.

The DSMC method employs a large number of simulated molecules with correct physical size to represent the real molecules which contain position and velocity information. They engage in collisions which follow the conservation law of energy and momentum. Unlike the particle interaction routines used in MD that each particle is given the opportunity to interact with every other particle, the collision in DSMC only happens between two molecules within the same cell. This makes the computational efficiency of the DSMC method far higher than that of the MD method. The DSMC method can be applied to weakly ionized gases, micro-electrical-mechanical systems (MEMS), space

shuttle re-entry dynamics and so forth. The probabilistic simulation method is used in Section 5.1.

Chapter 4 Theoretical calculation of polyvacancy

4.1 Theoretical calculation of monovacancy

The Pronk and Frenkel approach [71] provides a way to obtain the concentration of monovacancies in imperfect crystals. They present the formalism in the grand canonical ensemble, which considers a system with a fixed number of lattice sites M and a number of vacancies n contained in a volume V at temperature T . It is convenient to write down the expression for the grand potential of the crystal Ξ'

$$\Xi'_M = \sum_{n=0}^M \exp[(M-n)\beta\mu] Q_{M-n}(V, T), \quad (4.1)$$

where $\beta=1/k_B T$. We denote the free energy of a crystal with no vacancies by $F^{(0)} = -k_B T \ln Q_M$. Since in equilibrium the concentration of vacancies in a crystal is very low, we presume that vacancies do not interact. This assumption makes it easier to calculate the canonical partition function of a crystal with n vacancies

$$Q_{M-n}(V, T) \approx \frac{M!}{n!(M-n)!} Q^{(n)}(V, T) = \frac{M!}{n!(M-n)!} \exp(-\beta F^{(n)}), \quad (4.2)$$

where $F^{(n)}$ denotes the free energy of a crystal with n vacancies. It can be expressed as

$$F^{(n)} = F^{(0)} - nf_1 = Mf_0 - nf_1, \quad (4.3)$$

where f_0 is the free energy per particle in perfect crystal and $-f_1$ is the change in free energy due to the creation of a single vacancy. Combining Eqs. from (4.4) to (4.5), we have

$$\Xi'_M = \sum_{n=0}^M \frac{M!}{n!(M-n)!} \exp[(M-n)\beta\mu] \exp[-\beta(Mf_0 - nf_1)]. \quad (4.4)$$

By applying the binomial expansion, Eq. (4.4) can be simplified as

$$\Xi'_M = Q_M \exp(M\beta\mu) [1 + \exp(-\beta[\mu - f_1])]^M. \quad (4.5)$$

Since $\exp(-\beta[\mu - f_1])$ is usually less than unity, we can write the above equation as

$$\Xi'_M = Q_M \exp(M\beta\mu) \exp[M \exp(-\beta[\mu - f_1])]. \quad (4.6)$$

By applying a thermodynamic relation

$$\langle M - n \rangle = \frac{\partial \ln \Xi'}{\partial \beta\mu}, \quad (4.7)$$

the average number of vacancies can be derived as follows:

$$\langle n \rangle = M \exp[-\beta(\mu - f_1)]. \quad (4.8)$$

The monovacancy concentration x_{MV} is therefore expressed as

$$x_{MV} = \frac{\langle n \rangle}{M} = \exp[-\beta(\mu - f_1)]. \quad (4.9)$$

Hence, in order to compute monovacancy concentration x_{MV} , we have to know the chemical potential μ and calculate the free energy f_1 one by one.

4.1.1 Calculation of chemical potential μ

Before the chemical potential can be obtained, we start by calculating the pressure. In a FCC-structured crystal, we use the MeterPressureHardTensor class in the Etomica simulation package to compute the pressure tensor under MD3D (i.e. Molecular Dynamics in the 3-dimensional space). Then the value of pressure is obtained by averaging the 3 diagonal elements of the pressure tensor matrix formulated as:

$$P = (P_{xx} + P_{yy} + P_{zz}) / 3. \quad (4.10)$$

The simulation is carried out at 4 different densities and the results are shown in Table 4.1.

Table 4.1 Data of simulation value of pressure for perfect crystal

ρ^*	P_{xx}	P_{yy}	P_{zz}	P_{sim}
1.04063	11.64200	11.66600	11.64100	11.64967
1.08786	13.87700	13.86700	13.88400	13.87600
1.13137	16.61800	16.60800	16.59900	16.60833
1.17851	20.83100	20.80700	20.83100	20.82300

Note: ρ^* is the dimensionless number density, $kT=m=\sigma=1$, $\rho^* = \rho\sigma^3$; P_{xx} , P_{yy} , and P_{zz} are the three dimensionless diagonal components of pressure tensor; P_{sim} is the dimensionless average value.

Meanwhile, for the solid phase for the hard-sphere model, one can compute the theoretical value of compressibility factor via Speedy's equation of state [72] for FCC

$$z_{solid} = \frac{3}{1-n^*} - 0.5921 \frac{n^* - 0.7072}{n^* - 0.601}, \quad (4.11)$$

where n^* is the dimensionless density, i.e., $n^* = \rho\sigma^3 / \sqrt{2}$. Since $z = PV/Nk_B T = P/\rho k_B T$, one can get the theoretical value of P by $P = z\rho k_B T$ as shown in Table 4.2.

Table 4.2 Data of pressure from Speedy's equation of state for monovacancy

ρ^*	ρ^* / ρ_0	z_{solid}	P_{Speedy}
1.04063	0.73584	11.23080	11.68709
1.08786	0.76923	12.78168	13.90463
1.13137	0.80000	14.72389	16.65817
1.17851	0.83333	17.67855	20.83437

Note: ρ^* is the dimensionless number density, $kT = m = \sigma = 1$, $\rho^* = \rho\sigma^3$; ρ_0 is the close-packed dimensionless number density; z_{solid} is the compressibility factor; P_{Speedy} is the dimensionless pressure from Speedy's equation of state for monovacancy.

The comparison from the simulation and the equation of state is shown in Fig. 4.1, which shows excellent consistency between them.

Fig. 4.1 Comparison of simulation and theoretical values of pressure for monovacancy

After obtaining the pressure, we start to compute the chemical potential with the following equation

$$\mu = \frac{G}{N} = \frac{F}{N} + \frac{PV}{N} = \frac{F}{N} + \frac{P}{\rho}. \quad (4.12)$$

In order to compute the chemical potential, one must know the value of F first. By applying a basic thermodynamic formula $dF = -SdT - PdV$, $dF = -PdV$ can be obtained at a fixed temperature. The thermodynamic integration from a reference state to an expected state gives rise to

$$F - F_0 = -\int_{V_0}^V PdV. \quad (4.13)$$

The division with the number of particles N gives

$$\begin{aligned} \frac{F}{N} &= \frac{F_0}{N} - \int_{\rho_0}^{\rho} Pd\left(\frac{1}{\rho}\right) = \frac{F_0}{N} + \int_{\rho_0}^{\rho} \frac{P}{\rho^2} d\rho. \\ &= \frac{F_0}{N} + k_B T \int_{\rho_0}^{\rho} \frac{z}{\rho} d\rho. \end{aligned} \quad (4.14)$$

By substituting Eq. (4.11) into (4.14), we have

$$\frac{F}{N} = \frac{F_0}{N} + k_B T \int_{n_0^*}^{n^*} \left[\frac{3}{(1-n^*)n^*} - 0.5921 \frac{n^* - 0.7072}{n^*(n^* - 0.601)} \right] dn^*. \quad (4.15)$$

After the integration of the expression of F is finalized as:

$$\frac{F}{N} = \frac{F_0}{N} + k_B T \left(-3 \ln \frac{1-n^*}{1-n_0^*} + 2.3033 \ln \frac{n^*}{n_0^*} + 0.1046 \ln \frac{n^* - 0.601}{n_0^* - 0.601} \right). \quad (4.16)$$

From combining Eqs. (4.12) and (4.16), the formalism for chemical potential is finally appeared as a function of density under isothermal condition

$$\mu = \frac{F_0}{N} + k_B T \left(-3 \ln \frac{1-n^*}{1-n_0^*} + 2.3033 \ln \frac{n^*}{n_0^*} + 0.1046 \ln \frac{n^* - 0.601}{n_0^* - 0.601} \right) + \frac{P}{\rho}. \quad (4.17)$$

Since $F_{\text{excess}} = F - F_{\text{id}}$, $F = F_{\text{excess}} + F_{\text{id}}$. We choose the reference state at $\rho_0=1.040628$ (i.e. melting point density for FCC HS solid), and the reference value of F_{excess} and F_{id} from Ref. [73]. We obtain the value of F and μ at different densities as shown in Table 4.3.

Table 4.3 Data of chemical potential for monovacancy

ρ^*	ρ^* / ρ_0	P	F	μ
1.04063	0.73584	11.64967	4.96808	16.16293
1.08786	0.76923	13.87600	5.49893	18.25428
1.13137	0.80000	16.60833	6.03614	20.71597
1.17851	0.83333	20.82300	6.69333	24.36223

Note: ρ^* is the dimensionless number density, $kT=m=\sigma=1$, $\rho^* = \rho\sigma^3$; ρ_0 is the close-packed dimensionless number density; P is the dimensionless pressure; F is the dimensionless free energy; μ is the dimensionless chemical potential for monovacancy; F is based on per particle.

4.1.2 Calculation of the monovacancy free energy f_1

It is more convenient to compute f_1 , the change in free energy in a crystal due to the removal of a vacancy, rather than to compute $-f_1$ directly. Here we apply the Widom particle insertion method. Then f_1 is given by

$$f_1 = -k_B T \ln \left(\frac{V_{WS} \langle \exp(-\beta \Delta U) \rangle}{\Lambda^d} \right), \quad (4.18)$$

where V_{WS} is the volume of the Wigner-Seitz cell, ΔU is the change in the interaction potential energy due to the insertion of a trial particle and Λ^d is the de Broglie wavelength in the d dimension. For hard spheres, it can be expressed as

$$f_1 = -k_B T \ln \left(\frac{V_{WS} P_{acc}(V_{WS})}{\Lambda^d} \right), \quad (4.19)$$

where $P_{acc}(V_{WS})$ is the probability that the trial insertion in the WS cell will be accepted. Since most of the WS cell is not accessible, we attempt the insertion in a subvolume. Moreover, the removal of a particle from a subvolume of the WS cell in a perfect crystal (without vacancies) has to be considered as well. The expression for f_1 is then

$$f_1 = -k_B T \ln \left(\frac{v P_{acc}(v)}{P_{rem}(v) \Lambda^d} \right), \quad (4.20)$$

where v is the subvolume of the WS cell and $P_{rem}(v)$ is the probability that a particle happens to be inside this subvolume. Certainly in the final expression for the vacancy concentration, the factor Λ^d is canceled by the same term in the ideal portion of the chemical potential.

The SOC method is used to confine the particles within an inscribed sphere with an appropriate diameter. The insertion-deletion simulation (IDS) approach [74] is applied to compute insertion and deletion probability separately to obtain the final result for f_1 at varied densities.

Usually since the vacancy concentration is very low (for monovacancy less than 10^{-3} , for di- and tri-vacancy even lower), we cannot make a simulation system containing thousands of lattice sites with one vacancy which will consume great computational load. The choice of system size should be careful. Too large size sometimes doesn't decrease the computational error too much and causes huge computational cost. Too small size cannot represent the real system and causes large computational error. Appropriate system size should be chosen based on large number of simulation experiments, experience or literatures. In this simulation we obtain proper system size and choose 108 lattices constructed in a cubic box for FCC structure. The result is shown in Table 4.4.

Table 4.4 Data of free energy f_1

ρ^*	P_{rem}	P_{acc}	ν	f_1
1.04063	0.99995	0.00225	0.18340	7.79189
1.08786	1.00000	0.00195	0.17544	7.98138
1.13137	1.00000	0.00152	0.16869	8.26786
1.17851	1.00000	0.00090	0.16194	8.82973

Note: ρ^* is the dimensionless number density, $kT=m=\sigma=1$, $\rho^* = \rho\sigma^3$; P_{rem} is the removal probability; P_{acc} is the acceptance probability; ν is the dimensionless subvolume of the Wigner-Seitz cell; f_1 is the dimensionless formation free energy of monovacancy.

4.1.3 Calculation of monovacancy concentration

After the results of chemical potential μ and free energy f_1 are obtained, we can compute the monovacancy concentration x_{MV} by Eq. (4.9) and the result is shown in Fig. 4.2.

Fig. 4.2 Plot of monovacancy concentration as a function of number density

The simulation value of the monovacancy concentration agrees very well with the reference value [75] which applies the Widom particle insertion method to obtain the results.

4.2 Theoretical calculation of divacancy

4.2.1 Calculation of adjacent vacancy free energy f_2

By applying Pronk and Frenkel's biased Widom particle insertion methodology to a NN (Nearest Neighbor) vacancy but with different mechanism of the simulation from monovacancy, we formulate the calculation of free energy f_2 as follows:

$$f_2 = -k_B T \ln \left(\frac{v P_{acc}(v)}{P_{rem}(v) \Lambda^3} \right), \quad (4.21)$$

where $P_{acc}(v)$ is the acceptance probability of the trial insertion of the test particle into a biased subvolume v , and $P_{rem}(v)$ is the probability that a particle happens to reside in subvolume v . These contain similar meanings as done in the monovacancy simulation, but the details are different. Fig. 4.3 is an illustration of divacancy simulation. In order to see clearly, the atoms are deleted. Spheres with cyan color represent vacancies, red spheres represent nearest neighbors and blue ones represent shared neighbor particles by each vacancy. To obtain $P_{acc}(v)$, we introduce a divacancy in the system and perform the insertion trial into ‘a’ NN vacancy. To obtain $P_{rem}(v)$, we introduce a monovacancy in a independent simulation and obtain the probability that a NN particle happens to be in subvolume v .

Fig. 4.3 Illustration of divacancy simulation without particles

We use $864(4 \times 6 \times 6 \times 6)$ particles to run the simulation with a total time step of 200,000 with $\Delta t = 0.05 \text{ ps}$, which contain 20,000 as initialization cycles. The GUI (Graphical User Interface) for divacancy simulation is shown in Fig. 4.4.

Fig. 4.4 GUI for divacancy simulation

By doing insertion and deletion simulations independently, we obtain the probabilities and the free energy of adjunct vacancy as shown in Table 4.5.

Table 4.5 Data of free energy f_2

ρ^*	P_{rem}	P_{acc}	v	f_2
1.03760	0.99045	0.00032	0.36539	9.05537
1.05000	0.99438	0.00062	0.36107	8.40692
1.10000	0.99786	0.00027	0.34466	9.27592
1.20000	1.00000	0.00021	0.31594	9.63826

Note: ρ^* is the dimensionless number density, $kT = m = \sigma = 1$, $\rho^* = \rho \sigma^3$; P_{rem} is the removal probability; P_{acc} is the acceptance probability; v is the dimensionless subvolume of the Wigner-Seitz cell; f_2 is the dimensionless adjacent vacancy free energy.

4.2.2 Calculation of divacancy concentration

The total energy of a system containing monovacancies and two NN vacancies (i.e. divacancy) is defined as:

$$F_2 = Mf_0 - nf_1 - mf_2 - m(f_2 - f_1), \quad (4.22)$$

where f_0 is the free energy of a particle in a vacancy-free crystal, $-f_1$ is the free energy of formation of a monovacancy at a lattice site, $-f_2$ is the free energy of formation of a NN vacancy, n and m are the numbers of monovacancies and NN vacancies, respectively, and M is the total number of lattices.

The grand canonical potential is expressed as:

$$\Xi_{M'} \approx \sum_{m=0}^{M'} \exp[(M' - m)\beta\mu] \left(\frac{M'!}{m!(M' - m)!} \times \exp(-\beta F_2) \right), \quad (4.23)$$

where M' is $M-n$, and β is $1/k_B T$. The relation between m and M' can be expressed as:

$$m / M' = \exp[-\beta(\mu + f_1 - 2f_2)]. \quad (4.24)$$

Hence the product of the concentration of monovacancies and m/M' becomes the divacancy concentration x_{DV} which follows as:

$$x_{DV} = \exp[-2\beta(\mu - f_2)]. \quad (4.25)$$

The calculation of chemical potential μ is introduced in Section 4.1.1. Using Eq. (4.25), we obtain the data of divacancy concentration at varied densities, as is shown in

Fig. 4.5. The simulation value agrees quite well with the reference value which applies the Widom particle insertion methodology.

Fig. 4.5 Plot of divacancy concentration as a function of number density

4.3 Theoretical calculation of trivacancy

4.3.1 Trivacancy configurations for FCC and HCP crystals

The most common stackings of crystals are FCC and HCP lattices, which exhibit ABCABC... and ABAB... stackings, where A, B, and C are hexagonal planes of particles and there are stacked by varying 60° , respectively. By analyzing their spatial configurations, different types of trivacancies for FCC and HCP are found. For FCC crystal, there are 4 types of trivacancies, which hold the angles between each vacancy site

of 180° , 120° , 90° and 60° , termed as FCC3.1, FCC3.2, FCC3.3 and FCC3.4, respectively, as shown in Fig. 4.6. Considering all the possibilities, we find out that there is 1 configuration for FCC3.1, 4 configurations for FCC3.2, 2 configurations for FCC3.3 and 4 configurations for FCC3.4.

Fig. 4.6 The 4 types of FCC trivacancies: FCC3.1, FCC3.2, FCC3.3 and FCC3.4, corresponding to angles of 180° , 120° , 90° and 60° , respectively. The dark circles represent the vacancy sites. The lines in FCC3.4 show that the three vacancy sites form an equilateral triangle.

Unlike FCC which has only one divacancy configuration since it is symmetrical in all directions, the study of configuration for HCP trivacancy is a bit more complicated. We start from the two types of divacancies for a HCP crystal, and consider all the possibilities of nearest neighbor sites for each individual divacancy site, which means the total of 11×2 nearest neighbor sites are considered. Finally 9 different types of HCP trivacancies are found, as is shown in Fig. 4.7.

Fig. 4.7 The 9 types of HCP trivacancies: HCP3.1, HCP3.2, HCP3.3, HCP3.4, HCP3.5, HCP3.6, HCP3.7, HCP3.8, HCP3.9. The black and white circles represent the vacancy sites. The three rows, from top to bottom, correspond to trivacancies with 0, 2 and 3 sites in the same plane, respectively.

The characteristics of the nine types of HCP trivacancies are listed in the table below.

Table 4.6 Characteristics of the 9 types of HCP trivacancies [65]

HCP type	N_1	$\theta (^{\circ})$	N_2
HCP3.1	0	110	1
HCP3.2	0	146	2
HCP3.3	2	60	6
HCP3.4	2	90	4
HCP3.5	2	120	4
HCP3.6	3	60	1
HCP3.7	3	60	1
HCP3.8	3	120	2
HCP3.9	3	180	1

Note: N_1 denotes number of spheres in the same plane; θ denotes the angle formed by the trivacancy; N_2 denotes the number of configurations.

4.3.2 Insertion probabilities of trivacancies in FCC and HCP

By doing insertion simulation, we obtain the data of insertion probabilities of each trivacancy site for 4 different FCC trivacancy configurations and 9 different HCP configurations. The subvolumes at different number densities are shown in Table 4.7.

Table 4.7 Data of subvolume for trivacancy simulation

ρ^*	1.03760	1.04086	1.07000	1.10000	1.14000	1.17000	1.20000
v	0.36539	0.36424	0.35432	0.34466	0.33257	0.32404	0.31594

Note: ρ^* is the dimensionless number density, $kT=m=\sigma=1$, $\rho^* = \rho\sigma^3$; v is the dimensionless subvolume of the Wigner-Seitz cell.

The results of insertion probabilities for 4 different FCC trivacancy configurations are presented in Table 4.8. For each insertion, we make 3 runs and then average.

Table 4.8 Insertion probabilities of 4 different FCC trivacancies at different densities

	ρ^*	1.03760	1.04086	1.07000	1.10000	1.14000	1.17000	1.20000
FCC3.1	IP ₁	2.298×10 ⁻⁴	2.149×10 ⁻⁴	1.723×10 ⁻⁴	1.207×10 ⁻⁴	7.189×10 ⁻⁵	6.369×10 ⁻⁵	3.809×10 ⁻⁵
	IP ₂	9.983×10 ⁻⁵	9.539×10 ⁻⁵	7.407×10 ⁻⁵	5.584×10 ⁻⁵	3.855×10 ⁻⁵	3.761×10 ⁻⁵	2.563×10 ⁻⁵
	IP ₃	2.357×10 ⁻⁴	2.166×10 ⁻⁴	1.704×10 ⁻⁴	1.219×10 ⁻⁴	7.060×10 ⁻⁵	6.252×10 ⁻⁵	3.788×10 ⁻⁵
FCC3.2	IP ₁	2.883×10 ⁻⁴	2.863×10 ⁻⁴	2.199×10 ⁻⁴	1.518×10 ⁻⁴	9.828×10 ⁻⁵	6.700×10 ⁻⁵	4.744×10 ⁻⁵
	IP ₂	1.280×10 ⁻⁴	1.269×10 ⁻⁴	9.729×10 ⁻⁵	6.975×10 ⁻⁵	5.063×10 ⁻⁵	4.157×10 ⁻⁵	3.040×10 ⁻⁵
	IP ₃	2.820×10 ⁻⁴	2.811×10 ⁻⁴	2.122×10 ⁻⁴	1.519×10 ⁻⁴	1.000×10 ⁻⁴	6.834×10 ⁻⁵	4.741×10 ⁻⁵
FCC3.3	IP ₁	1.417×10 ⁻⁴	1.395×10 ⁻⁴	9.956×10 ⁻⁵	6.003×10 ⁻⁵	3.322×10 ⁻⁵	2.540×10 ⁻⁵	1.698×10 ⁻⁵
	IP ₂	7.007×10 ⁻⁵	6.685×10 ⁻⁵	4.766×10 ⁻⁵	3.255×10 ⁻⁵	1.859×10 ⁻⁵	1.551×10 ⁻⁵	1.234×10 ⁻⁵
	IP ₃	1.471×10 ⁻⁴	1.415×10 ⁻⁴	9.818×10 ⁻⁵	6.254×10 ⁻⁵	3.597×10 ⁻⁵	2.352×10 ⁻⁵	1.612×10 ⁻⁵
FCC3.4	IP ₁	9.953×10 ⁻⁵	9.521×10 ⁻⁵	5.314×10 ⁻⁵	2.676×10 ⁻⁵	8.644×10 ⁻⁶	4.613×10 ⁻⁶	1.232×10 ⁻⁶
	IP ₂	1.024×10 ⁻⁴	9.701×10 ⁻⁵	5.825×10 ⁻⁵	3.074×10 ⁻⁵	8.698×10 ⁻⁶	3.907×10 ⁻⁶	1.370×10 ⁻⁶
	IP ₃	1.057×10 ⁻⁴	9.695×10 ⁻⁵	5.221×10 ⁻⁵	2.738×10 ⁻⁵	8.461×10 ⁻⁶	4.127×10 ⁻⁶	1.602×10 ⁻⁶

Note: ρ^* is the dimensionless number density, $kT=m=\sigma=1$, $\rho^* = \rho\sigma^3$; IP_1 , IP_2 , IP_3 are the average insertion probabilities of 3 runs for different vacancy site respectively.

The results of insertion probabilities for 9 different HCP trivacancy configurations are presented in Table 4.9. For each insertion, we make 3 runs and then average.

Table 4.9 Insertion probabilities of 9 different HCP trivacancies at different densities

	ρ^*	1.03760	1.04086	1.07000	1.10000	1.14000	1.17000	1.20000
HCP3.1	IP ₁	3.359×10^{-4}	3.249×10^{-4}	2.323×10^{-4}	1.613×10^{-4}	1.074×10^{-4}	7.914×10^{-5}	4.329×10^{-5}
	IP ₂	1.526×10^{-4}	1.454×10^{-4}	1.036×10^{-4}	7.801×10^{-5}	5.694×10^{-5}	5.089×10^{-5}	3.023×10^{-5}
	IP ₃	3.302×10^{-4}	3.174×10^{-4}	2.374×10^{-4}	1.614×10^{-4}	1.093×10^{-4}	8.186×10^{-5}	4.264×10^{-5}
HCP3.2	IP ₁	2.934×10^{-4}	2.752×10^{-4}	2.137×10^{-4}	1.539×10^{-4}	9.281×10^{-5}	6.359×10^{-5}	3.329×10^{-5}
	IP ₂	1.390×10^{-4}	1.343×10^{-4}	9.712×10^{-5}	7.072×10^{-5}	5.101×10^{-5}	4.057×10^{-5}	2.636×10^{-5}
	IP ₃	2.900×10^{-4}	2.828×10^{-4}	2.150×10^{-4}	1.529×10^{-4}	9.371×10^{-5}	6.516×10^{-5}	3.311×10^{-5}
HCP3.3	IP ₁	1.048×10^{-4}	9.382×10^{-5}	5.468×10^{-5}	2.693×10^{-5}	7.463×10^{-6}	3.465×10^{-6}	1.102×10^{-6}
	IP ₂	1.520×10^{-4}	1.367×10^{-4}	7.818×10^{-5}	4.238×10^{-5}	1.336×10^{-5}	7.019×10^{-6}	2.703×10^{-6}
	IP ₃	1.062×10^{-4}	1.013×10^{-4}	5.645×10^{-5}	2.560×10^{-5}	7.901×10^{-6}	3.269×10^{-6}	1.428×10^{-6}
HCP3.4	IP ₁	1.740×10^{-4}	1.783×10^{-4}	1.131×10^{-4}	6.987×10^{-5}	3.688×10^{-5}	2.029×10^{-5}	1.130×10^{-5}
	IP ₂	6.116×10^{-5}	6.113×10^{-5}	3.838×10^{-5}	2.300×10^{-5}	1.130×10^{-5}	7.246×10^{-6}	4.695×10^{-6}
	IP ₃	1.215×10^{-4}	1.216×10^{-4}	7.719×10^{-5}	4.498×10^{-5}	2.056×10^{-5}	1.126×10^{-5}	5.982×10^{-6}
HCP3.5	IP ₁	2.888×10^{-4}	2.857×10^{-4}	2.095×10^{-4}	1.543×10^{-4}	9.383×10^{-5}	6.419×10^{-5}	4.109×10^{-5}
	IP ₂	1.014×10^{-4}	9.773×10^{-5}	6.867×10^{-5}	4.617×10^{-5}	2.896×10^{-5}	2.212×10^{-5}	1.351×10^{-5}
	IP ₃	2.052×10^{-4}	2.021×10^{-4}	1.456×10^{-4}	9.283×10^{-5}	5.044×10^{-5}	3.406×10^{-5}	1.968×10^{-5}
HCP3.6	IP ₁	2.590×10^{-5}	2.545×10^{-5}	1.078×10^{-5}	4.421×10^{-6}	1.538×10^{-6}	6.807×10^{-7}	3.392×10^{-7}
	IP ₂	2.247×10^{-5}	2.368×10^{-5}	1.096×10^{-5}	5.813×10^{-6}	2.018×10^{-6}	5.969×10^{-7}	2.557×10^{-7}
	IP ₃	2.277×10^{-5}	2.343×10^{-5}	1.032×10^{-5}	5.756×10^{-6}	1.861×10^{-6}	6.301×10^{-7}	2.737×10^{-7}
HCP3.7	IP ₁	7.358×10^{-4}	7.474×10^{-4}	6.890×10^{-4}	5.842×10^{-4}	4.572×10^{-4}	3.779×10^{-4}	2.841×10^{-4}
	IP ₂	7.528×10^{-4}	7.343×10^{-4}	6.827×10^{-4}	5.664×10^{-4}	4.508×10^{-4}	3.759×10^{-4}	2.394×10^{-4}
	IP ₃	7.601×10^{-4}	7.646×10^{-4}	6.696×10^{-4}	6.070×10^{-4}	4.970×10^{-4}	4.069×10^{-4}	2.515×10^{-4}
HCP3.8	IP ₁	1.977×10^{-4}	1.951×10^{-4}	1.294×10^{-4}	7.973×10^{-5}	4.036×10^{-5}	2.298×10^{-5}	1.360×10^{-5}
	IP ₂	6.888×10^{-5}	6.831×10^{-5}	4.237×10^{-5}	2.509×10^{-5}	1.433×10^{-5}	9.622×10^{-6}	6.648×10^{-6}
	IP ₃	2.020×10^{-4}	1.926×10^{-4}	1.388×10^{-4}	8.478×10^{-5}	4.292×10^{-5}	2.433×10^{-5}	1.402×10^{-5}
HCP3.9	IP ₁	1.740×10^{-4}	1.674×10^{-4}	1.067×10^{-4}	6.261×10^{-5}	2.944×10^{-5}	1.699×10^{-5}	7.816×10^{-6}
	IP ₂	5.670×10^{-5}	5.623×10^{-5}	3.742×10^{-5}	2.065×10^{-5}	8.683×10^{-6}	5.103×10^{-6}	3.752×10^{-6}
	IP ₃	1.749×10^{-4}	1.631×10^{-4}	1.110×10^{-4}	6.486×10^{-5}	2.922×10^{-5}	1.513×10^{-5}	7.519×10^{-6}

Note: ρ^* is the dimensionless number density, $kT=m\sigma=1$, $\rho^* = \rho\sigma^3$; IP₁, IP₂, IP₃ are the average insertion probabilities of 3 runs for different vacancy site respectively.

After the insertion probabilities of FCC and HCP at different configurations are obtained, the next step is to get the deletion probabilities correspondingly by doing deletion simulation. Then we can calculate the free energy of each individual vacancy within trivacancies and finally the concentration of trivacancies using the same method as divacancy concentration.

4.3.3 Deletion probabilities of trivacancies in FCC and HCP

By doing deletion simulation, we obtain the data of deletion probabilities of each trivacancy site for 4 different FCC trivacancy configurations and 9 different HCP configurations. The results of deletion probabilities for 4 different FCC trivacancy configurations are presented in Table 4.10. For each deletion, we make 3 runs and then average.

Table 4.10 Deletion probabilities of 4 different FCC trivacancies at different densities

	ρ^*	1.03760	1.04086	1.07000	1.10000	1.14000	1.17000	1.20000
FCC3.1	DP ₁	0.9478	0.9482	0.9494	0.9498	0.9500	0.9500	0.9500
	DP ₂	0.9400	0.9399	0.9475	0.9432	0.9500	0.9500	0.9500
	DP ₃	0.9393	0.9392	0.9405	0.9410	0.9408	0.9380	0.9385
FCC3.2	DP ₁	0.9444	0.9462	0.9488	0.9496	0.9500	0.9500	0.9500
	DP ₂	0.9391	0.9276	0.9373	0.9492	0.9499	0.9500	0.9500
	DP ₃	0.9361	0.9365	0.9379	0.9390	0.9396	0.9393	0.9366
FCC3.3	DP ₁	0.9486	0.9487	0.9496	0.9499	0.9500	0.9500	0.9500
	DP ₂	0.9344	0.9443	0.9478	0.9496	0.9500	0.9500	0.9500
	DP ₃	0.9412	0.9415	0.9424	0.9420	0.9424	0.9405	0.9360
FCC3.4	DP ₁	0.9047	0.9090	0.9161	0.9252	0.9364	0.9468	0.9116
	DP ₂	0.9056	0.9052	0.9139	0.9222	0.9323	0.9468	0.9494
	DP ₃	0.8886	0.8902	0.8996	0.9050	0.9154	0.9212	0.9221

Note: ρ^* is the dimensionless number density, $kT=m=\sigma=1$, $\rho^* = \rho\sigma^3$; DP_1, DP_2, DP_3 are the average deletion probabilities of 3 runs for different vacancy site respectively.

The results of deletion probabilities for 9 different HCP trivacancy configurations are presented in Table 4.11. For each deletion, we make 3 runs and then average.

Table 4.11 Deletion probabilities of 9 different HCP trivacancies at different densities

	ρ^*	1.03760	1.04086	1.07000	1.10000	1.14000	1.17000	1.20000
HCP3.1	DP ₁	0.9974	0.9979	0.9992	0.9998	1.0000	1.0000	1.0000
	DP ₂	0.9939	0.9918	0.9983	0.9996	1.0000	1.0000	1.0000
	DP ₃	0.9877	0.9878	0.9897	0.9895	0.9902	0.9888	0.9870
HCP3.2	DP ₁	0.9969	0.9970	0.9990	0.9997	1.0000	1.0000	1.0000
	DP ₂	0.9893	0.9864	0.9967	0.9992	1.0000	1.0000	1.0000
	DP ₃	0.9875	0.9873	0.9897	0.9903	0.9906	0.9884	0.9845
HCP3.3	DP ₁	0.9494	0.9466	0.9543	0.9750	0.9813	0.9873	0.9973
	DP ₂	0.9566	0.9556	0.9698	0.9807	0.9887	0.9890	0.9993
	DP ₃	0.9359	0.9375	0.9451	0.9523	0.9601	0.9717	0.9682
HCP3.4	DP ₁	0.9990	0.9991	0.9997	1.0000	1.0000	1.0000	1.0000
	DP ₂	0.9905	0.9863	0.9981	0.9995	1.0000	1.0000	1.0000
	DP ₃	0.9898	0.9908	0.9914	0.9915	0.9915	0.9905	0.9890
HCP3.5	DP ₁	0.9966	0.9970	0.9989	0.9997	1.0000	1.0000	1.0000
	DP ₂	0.9829	0.9856	0.9964	0.9992	1.0000	1.0000	1.0000
	DP ₃	0.9860	0.9858	0.9880	0.9896	0.9880	0.9877	0.9860
HCP3.6	DP ₁	0.9871	0.9861	0.9920	0.9983	0.9998	1.0000	1.0000
	DP ₂	0.9872	0.9850	0.9944	0.9826	0.9868	1.0000	1.0000
	DP ₃	0.9633	0.9640	0.9676	0.9694	0.9687	0.9657	0.9614
HCP3.7	DP ₁	0.8584	0.8582	0.8513	0.8358	0.8178	0.7626	0.7433
	DP ₂	0.8555	0.8574	0.8518	0.8248	0.8041	0.7703	0.7303
	DP ₃	0.8558	0.8510	0.8458	0.8385	0.8248	0.8160	0.8078
HCP3.8	DP ₁	0.9966	0.9970	0.9989	0.9997	1.0000	1.0000	1.0000
	DP ₂	0.9863	0.9871	0.9961	0.9990	1.0000	1.0000	1.0000
	DP ₃	0.9841	0.9839	0.9868	0.9869	0.9862	0.9842	0.9805
HCP3.9	DP ₁	0.9970	0.9970	0.9990	0.9998	1.0000	1.0000	1.0000
	DP ₂	0.9810	0.9880	0.9819	0.9990	0.9999	1.0000	1.0000
	DP ₃	0.9865	0.9871	0.9885	0.9890	0.9898	0.9875	0.9860

Note: ρ^* is the dimensionless number density, $kT=m=\sigma=1$, $\rho^* = \rho\sigma^3$; DP_1, DP_2, DP_3 are the average deletion probabilities of 3 runs for different vacancy site respectively.

4.3.4 Trivacancy concentrations of FCC and HCP

We formulate the equilibrium concentration of polyvacancy in crystal by modifying the grand canonical ensemble. The generalized equation can be applied to obtain the concentration of vacancy in any order and shape in crystal. In this work, we adopt a Widom-like particle insertion method implemented with molecular dynamics simulation to obtain individual formation free energies for constituent vacancies in polyvacancy. As a case study, we obtained equilibrium concentrations and formation free energies of four and nine trivacancies identified in face-centered-cubic (FCC) and hexagonal-close-packed (HCP) hard-sphere crystals, respectively. Our result is in excellent agreement with literature data available for the FCC. Further, we elucidate stabilities of equilateral triangular trivacancies and report their relative stabilities in FCC and HCP hard-sphere crystals. Overall trivacancy concentration is higher in the HCP HS crystal [76].

Defects in solids are common in nature. The formation of various types of defects may result from deformation, quenching, radiation (i.e. X-ray or neutrons), and etc [77]. It is known that some in quantity are vital to their hosting crystals, but some are fatal and harmful by excess amount [78]. Among these defects, vacancy, a type of point defect, occurs most commonly. It is created when atoms or molecules are missing from their respective lattice sites or diffused to other spaces in crystal [62, 79]. Vacancies usually exist in low concentration in crystals, yet they are continuously highlighted in solid state fields because of their disproportionate effects on material properties; consequently, there have been many studies depending on solid types, where vacancies make significant impact on solid properties related to fields of superconductivity, electronic property, ferromagnetism, electrical conductivity, and etc [80-83] [76].

It is possible but generally difficult to study vacancies by experiment since multi-monitoring experiments must be done to extract uncorrelated information of vacancies residing within crystals. To make the research of vacancy easy and produce great details in every aspect of it, we resort to the statistical mechanical approach to develop a formalism to handle vacancy and utilize the molecular modeling and simulation method to implement the formalism. As one of classical and theoretical studies, Bennett and Alder [68] reported equilibrium concentrations of mono-to tetravacancies in the face-centered-cubic (FCC) hard-sphere (HS) crystal in the vicinity of melting. They produced concentrations of vacancies from transition probabilities among vacancies, yet failed to do so at high density where motilities of vacancies significantly are reduced. From Pronk and Frenkel's alternative way to consider vacancy [71], retrospective studies on low-order vacancies (i.e. mono-and divacancies) have been done on thermodynamic properties of vacancies, considering their stabilities by determining formation free energies of vacancies [74, 75, 84]. Those studies share interesting advantages in the study of vacancy since, first, the constituent vacancies treated as members in the high-order vacancy (i.e. polyvacancy hereafter) are independently handled in a statistical mechanical derivation, and second, the high density is not an obstacle for obtaining detailed properties of the target vacancy such as formation free energy, concentration, and relative stability [76].

Applicable potentiality of their work to polyvacancy is theoretically established in this work. We derive a theoretical formalism that allows us to obtain the equilibrium concentration of vacancy in any order. To assess the developed equation, trivacancy is chosen and we use the collision-based molecular dynamics (MD) method complemented

with a biased insertion method to obtain its thermodynamic properties. To test sensitivity of the method of the approach, hexagonal-close-packed (HCP) HS crystal was chosen for the comparison against the FCC HS crystal. Our approach can readily distinguish different trivacancies (i.e. 4 types of trivacancies for the FCC HS crystal and 9 types [85] for the HCP) and obtain stabilities of constituent vacancies within those trivacancies while Bennett and Alder's approach could not do so [76].

Equilibrium concentration refers to the ratio of number of vacancies to number of lattice sites. To obtain the equilibrium concentration of the n^{th} order vacancy (i.e. hereafter n^{th} -vacancy), we first define how the system contains the equilibrium number of n^{th} -vacancies. It embraces a same meaning that how many of lattice sites (or particles) are required to accommodate single n^{th} -vacancy. This is understandable because if the number of lattice sites is obtained, the equilibrium concentration of n^{th} -vacancy can be retrieved. For the existence of single n^{th} -vacancy, we postulate that the system is required to have the equilibrium number of vacancies with one order below since higher order vacancies do not form without presence of one-order lower vacancies. Note that when n is 1 (i.e. monovacancy) there are no lower order vacancies, thus the equilibrium number of particles is only considered. Based on this postulation, one can set up the basis and superstructure relation between the system of monovacancy and the system at the next level, which contains single divacancy (e.g. figures 4.8a) and b) depicts this example). The one plays a role of subsystem in the latter since the divacancy-system is composed of the equilibrium number of sets of the monovacancy-system. This scheme ensures the satisfaction of the genuine meaning of equilibrium referred in this work such that interactions between vacancies are negligible. The abovementioned argument is

methodically redefined with a system of single n th-vacancy, which is determined by the number of lattice sites that is equivalent to the number of subsystems $M^{(n-1)}$, the volume V , and the temperature T . A subsystem in $M^{(n-1)}$ consists of cascade of subsystems of mono-, di-, tri-, ... , and $(n-2)$ th-vacancies, and finally a subsystem containing single $(n-1)$ th -vacancy. Figure 4.8 elaborates on the set-up of this system in two dimension. For example, when n is 1, $M^{(0)}$ indicates that there exists single monovacancy in the system (i.e. particles are subsystems in this particular case). When n is 2, $M^{(1)}$ indicates that there exist the equilibrium number of subsystems containing monovacancy and a subsystem containing single divacancy. When n is 3, $M^{(2)}$ indicates that there exist the equilibrium number of subsystems containing divacancy, which separately contain the equilibrium number of subsystems containing monovacancy, and divacancies, and a subsystem containing single trivacancy, and so on [76].

Fig. 4.8 (Color on line) the systems a) (i.e. $M^{(0)}$), b) (i.e. $M^{(1)}$), c) (i.e. $M^{(2)}$), and d) (i.e. $M^{(3)}$) represent the systems of mono-, di-, tri-, and tetra-vacancies, respectively. The system a) is the subsystem of the system b), which becomes the subsystem of the system c), and so on. Lattice sites are not explicitly shown, but it is assumed that vacancies and particles (i.e. gray background) are located at lattice sites. [Reprinted with permission from Ref. [76]. Copyright 2011, Taylor & Francis]

In solving this problem mathematically, we introduce the concept of the nearest neighbor (NN) vacancy, which turns a $(n-1)$ th-vacancy into a n th-vacancy, since it simplifies complex expressions involved in the statistical mechanical derivation. This physical view can be readily understood to observe the development of $M^{(0)}$ to $M^{(1)}$ (i.e. See Fig. 4.8). If single NN vacancy is introduced into one of adjacent sites of monovacancy by removing a particle, the system $M^{(0)}$ increases while monovacancies are accordingly introduced until the system $M^{(1)}$, which has the equilibrium number of

subsystems (i.e. $M^{(0)}$ s), appeared. It is important to note that the system $M^{(1)}$ is not composed of duplicates of $M^{(0)}$ but subsystems $M^{(0)}$ s with different configurations. Finally, our problem is concentrated on capturing the equilibrium number of NN vacancies using statistical mechanics approach. First, we set the number of particles $N = M^{(n-1)} - v_n$, where v_n is the number of NN vacancies. Then a grand potential can be written as

$$\begin{aligned}\Xi_{M^{(n-1)}} &= \sum_{v_n=0}^{M^{(n-1)}} \exp(\beta\mu N) Q_{M^{(n-1)}-v_n}(V, T) \\ &= \sum_{v_n=0}^{M^{(n-1)}} \exp(\beta\mu M^{(n-1)}) \exp(-\beta\mu v_n) Q_{M^{(n-1)}-v_n}(V, T).\end{aligned}\tag{4.26}$$

The partition function $Q_{M^{(n-1)}-v_n}$ is defined as

$$Q_{M^{(n-1)}-v_n}(V, T) \approx \exp(-\beta F_n) \frac{M^{(n-1)}!}{v_n! (M^{(n-1)} - v_n)!},\tag{4.27}$$

where F_n is the total free energy of the system, β is the inverse temperature, $1/kT$ where k is the Boltzmann's constant, and μ is the chemical potential. There may be several types in a polyvacancy due to different orientations of constituent vacancies. Eq. (4.27) is not formulated to handle several types of the polyvacancy altogether. Instead, it can be assumed that NN vacancies are introduced only to particular adjacent sites of $(n-1)$ th vacancies, thus n th vacancies have same shape and orientation. In other words, Eq. (4.27) handles the system containing a particular shape of n th vacancy. With the identified NN vacancies, F_n can be expressed as

$$\begin{aligned}
F_n &= Mf_p - v_1 f_1^* - v_2 f_2^* - v_2 (f_2^* - f_1^*) - v_3 f_3^* - 2v_3 (f_3^* - f_2^*) \cdots - v_n f_n^* - (n-1)v_n (f_n^* - f_{n-1}^*) \\
&= F_{n-1} - v_n (nf_n^* - (n-1)f_{n-1}^*) \\
&= F_{n-1} - v_n (f_n - f_{n-1}),
\end{aligned} \tag{4.28}$$

where M is the total number of lattice sites, f_p is the free energy of the particle, f_i^* s are the free energies of the NN vacancies adjacent to $(i-1)$ th-vacancy or constituent vacancies in i th-vacancy (e.g. f_1^* is the free energy of monovacancy, f_2^* is the free energy of the NN vacancy adjacent to monovacancy or a constituent vacancy in divacancy, f_3^* is the one adjacent to divacancy or a constituent vacancy in trivacancy, and etc), v_n s are the numbers of n th-vacancy, F_{n-1} is the total free energy of the system containing mono to $(n-1)$ th-vacancies, $f_n = n \times f_n^*$ (or $= \sum_{m=1}^{m=n} f_{n,m}^*$), and $f_{n-1} = (n-1) \times f_{n-1}^*$ (or $= \sum_{m=1}^{m=n-1} f_{n-1,m}^*$).

To help understanding of the construction of Eq.(4.28), F_1 , F_2 and F_3 are shown in Fig. 4.9 as examples. It is straightforward to obtain F_1 in Fig. 4.9b) from F in 4.9a). When the NN vacancies are introduced to make divacancies as shown in Fig. 4.9c), monovacancies adjacent to the NN vacancies become one of two constituent vacancies in divacancies. The free energy difference in this transformation must be accounted for, thus the quantity $(f_2^* - f_1^*)$, which are multiplied by the number of the introduced NN vacancies (i.e. $v_2 = 4$ in Fig. 4.9c), is subtracted from F in Fig. 4.9c) to obtain F_2 in Fig. 4.9d). For the trivacancy system, we arbitrarily made single trivacancy. When a NN vacancy is introduced as shown in Fig. 4.9e), divacancy is turned into constituent vacancies in trivacancy. The free energy difference in this transformation is obtained by the quantity $(f_3^* - f_2^*)$ for each constituent vacancy [76].

Fig. 4.9 (Color on line) Schematic diagrams to develop the monovacancy system (a→b), the divacancy system, (c→d), and the trivacancy system (e→f), which are presented with each total free energy F_n , respectively. The lattice sites are occupied by particles unless emptied by vacancies.

[Reprinted with permission from Ref. [76]. Copyright 2011, Taylor & Francis]

There are two constituent vacancies involved in this transformation, thus the factor 2 is multiplied. Note that the quantity $2 \times (f_3^* - f_2^*)$ must be multiplied by the number of introduced NN vacancies v_3 but v_3 is 1 in Fig. 4.9e). To this end, F_3 is obtained as shown in Fig. 4.9f). For convenience, it is assumed that the quantity $(f_i^* - f_{i-1}^*)$ is same in transformations of constituent vacancies in $(i-1)$ th-vacancy to constituent vacancies in i^{th} -vacancy when the NN vacancy is introduced. This is reason why n and $n-1$ are appeared as multipliers in the middle expression in Eq.(4.28). However, free energies of constituent vacancies depend on their positions, more accurately locations of particles surrounding those constituent vacancies. Thus, it is right to express the quantity $(nf_n^* - (n-1)f_{n-1}^*)$ in Eq.(4.28) as $(\sum_{m=1}^{m=n} f_{n,m}^* - \sum_{m=1}^{m=n-1} f_{n-1,m}^*)$ where the second subscript m represents different constituent vacancies. Throughout the derivation, the quantity $(f_n - f_{n-1})$ in Eq.(4.28) is intact, thus both notations do not affect to change the final equation. Finally, we can write the partition function in Eq.(4.27) by using the last expression in Eq.(4.28) as

$$Q_{M^{(n-1)}-v_n}(V, T) \approx \exp\left(-\beta(F_{n-1} - v_n(f_n - f_{n-1}))\right) \frac{M^{(n-1)}!}{v_n!(M^{(n-1)} - v_n)!}. \quad (4.29)$$

The mathematical derivation can be found in the work of Pronk and Frenkel for monovacancy. For completing our argument, we take a similar derivation process. Substituting Eq.(4.29) into Eq.(4.26) results in

$$\Xi_{M^{(n-1)}} \approx \exp\left[\beta(\mu M^{(n-1)} - F_{n-1})\right] \sum_{v_n=0}^{M^{(n-1)}-v_n} \exp(-\beta v_n(\mu + f_{n-1} - f_n)) \frac{M^{(n-1)}!}{v_n!(M^{(n-1)} - v_n)!}. \quad (4.30)$$

By applying the binomial theorem, the summation term in Eq.(4.30) is simplified to

$$\Xi_{M^{(n-1)}} \approx \exp\left[\beta\left(\mu M^{(n-1)} - F_{n-1}\right)\right] \left[1 + \exp\left(-\beta\left(\mu + f_{n-1} - f_n\right)\right)\right]^{M^{(n-1)}}. \quad (4.31)$$

Equation (4.31) is further simplified due to $\exp\left[-\beta\left(\mu + f_{n-1} - f_n\right)\right] \ll 1$,

$$\Xi_{M^{(n-1)}} \approx \exp\left[\beta\left(\mu M^{(n-1)} - F_{n-1}\right)\right] \exp\left[M^{(n-1)} \exp\left(-\beta\left(\mu + f_{n-1} - f_n\right)\right)\right]. \quad (4.32)$$

By taking the natural logarithm of both sides, equation (4.32) becomes

$$\ln \Xi_{M^{(n-1)}} \approx \beta\left(\mu M^{(n-1)} - F_{n-1}\right) + M^{(n-1)} \exp\left(-\beta\left(\mu + f_{n-1} - f_n\right)\right). \quad (4.33)$$

The maximum of $\Xi_{M^{(n-1)}}$ is obtained from summing over the states varied by the change of $M^{(n-1)}$. In this process as similarly shown by others [71, 74] the pressure deviation as a thermodynamic quantity due to the introduction of the NN vacancy is examined to have the maximum variation of $M^{(n-1)}$. To this end, the average number of particles $\langle N \rangle$ in the system is found by differentiating Eq.(4.33) with respect to the chemical potential term $\beta\mu$

$$\left\langle M^{(n-1)} - v_n \right\rangle = \frac{\partial \ln \Xi_{M^{(n-1)}}}{\partial \beta\mu} = M^{(n-1)} - M^{(n-1)} \exp\left(-\beta\left(f_{n-1} - f_n\right)\right) \exp(\beta\mu). \quad (4.34)$$

Hence, we finally derive the equilibrium number of NN vacancies as follows:

$$\frac{\langle v_n \rangle}{M^{(n-1)}} = \exp\left(-\beta\left(\mu + f_{n-1} - f_n\right)\right). \quad (4.35)$$

Equation (4.35) reveals how many sub-systems must be existed to accommodate one NN vacancy to make n^{th} -vacancy. This information is readily given by inversion of Eq.(4.35). One must pay attention to the case when Eq.(4.35) is applied to obtain the concentration of monovacancy by reminding the above-introduced equation $f_{n-1} = (n-1) \times f_{n-1}^*$. When n is 1, f_{n-1} becomes zero. To know actual number of lattice sites in $M^{(n-1)}$, one must utilize equilibrium numbers of subsystems of lower order vacancies. This involves the product calculation, which is illustrated in the following expression:

$$M_n = \left(\frac{M^{(0)}}{\langle v_1 \rangle} \right) \left(\frac{M^{(1)}}{\langle v_2 \rangle} \right) \left(\frac{M^{(2)}}{\langle v_3 \rangle} \right) \cdots \left(\frac{M^{(n-2)}}{\langle v_{n-1} \rangle} \right) \left(\frac{M^{(n-1)}}{\langle v_n \rangle} \right). \quad (4.36)$$

M_n is the total number of lattices to accommodate one n^{th} -vacancy. The first term represents the number of lattices to accommodate a monovacancy; the second term represents the number of subsystems having the number of sites, $M^{(0)} / \langle v_1 \rangle$, to accommodate a NN vacancy introduced to monovacancy to make a divacancy; the third term represents sub-systems having the number of sites $(M^{(0)} / \langle v_1 \rangle) \times (M^{(1)} / \langle v_2 \rangle)$ to accommodate a NN vacancy introduced to a divacancy to make a trivacancy, and so on. From this product calculation, the equilibrium concentration of n^{th} -vacancy, x_n , is found as follows:

$$x_n = \prod_{m=1}^{m=n} \exp(-\beta(\mu + f_{m-1} - f_m)). \quad (4.37)$$

Equation (4.37) is further simplified in terms of the vacancy-order n as follows:

$$x_n = \exp(-\beta(n\mu - f_n)). \quad (4.38)$$

It is important to notice that equation (4.38) does not require any information of lower order vacancies to calculate the concentration of the target order vacancy. Note that f_n is as $\sum_{m=1}^{m=n} f_{n,m}^*$, which is the sum of formation free energies of constituent vacancies [76].

For the case study to show the validity of Eq.(4.38), we chose to study trivacancies in FCC and HCP HS crystals. The concentration form is then as follows:

$$x_3 = \exp(-\beta(3\mu - (f_{3,1} + f_{3,2} + f_{3,3}))). \quad (4.39)$$

The chemical potentials of vacancy-free FCC and HCP HS crystals are obtained from the thermodynamic relation [6] $\beta\mu = \beta f_0 + Z$, where Z is the compressibility factor, PV_0/NkT , where P is the pressure and V_0 is the vacancy-free volume. The Einstein crystal method developed by Frenkel and Ladd allows us to have the free energy of the particle f_0 [73]. By the thermodynamic integration with Speedy's equation of state [72], chemical potentials at different densities other than melting were obtained. Equation(4.39) requires formation free energies of three NN vacancies in trivacancy, therefore three connected vacancies in any shapes of interest are introduced into the simulation system. To evaluate individual formation free energies of NN vacancies, we use a particle insertion method modified from the Widom method [86]. The volume of the vacancy is directly related to the formation free energy of the vacancy by the following equation:

$$-\beta f_{NN} = \ln(v_{vac}/\Lambda^3), \quad (4.40)$$

where λ^3 is the de Broglie thermal wavelength in three-dimension and v_{vac} is the vacancy volume. When a particle is removed from a crystal lattice, the empty volume can be calculated by making a connection with Wigner Seitz (WS) cell with v_{WS} as the volume of it. Mathematically, it is difficult to formulate the empty volume surrounded by moving NN particles in the system, but with the insertion simulation, it can be readily obtained. The change of the potential energy (i.e. ΔU) of the system by inserting a particle into a vacancy is averaged after weighting by the Boltzmann factor. Then, v_{vac} can be predicted by $v_{WS} \times \langle e^{-\beta \Delta U} \rangle$ [76].

A considerable drawback of the insertion method comes from vastness of the original insertion volume of the WS cell. Most area of the WS cell is not accessible due to hindrance of surrounding NN particles against insertion trials. To resolve this issue, Pronk and Frenkel introduced a biased particle insertion method [71]. The core idea is to introduce a sub-volume v smaller than the WS cell and subsequently, the relationship between v_{vac} and v has been found by a simple integration method. A statistical ratio, $v \times P_{in}(v) / v_{vac}$, represents the accessible volume over the volume of vacancy v_{vac} because $P_{in}(v)$ is the accepted-insertion probability of a particle into the sub-volume v . Alternative description of the statistical ratio is a probability that a particle happens to be inside of the sub-volume v . In MD simulation, one can express this argument as follows:

$$P_{del}(v) = \frac{v_{acc}}{v_{vac}} = \frac{v P_{in}(v)}{v_{vac}}, \quad (4.41)$$

where v_{acc} is the accessible volume and $P_{del}(v)$ is the probability of a particle's spent time in the sub-volume v . Therefore, if one performs two independent simulations to obtain $P_{in}(v)$ and $P_{del}(v)$, v_{vac} can be obtained [76].

We elucidated the biased insertion method only for monovacancy, but this approach is easily applicable to the polyvacancy simulation. For $P_{in}(v)$ s in trivacancy, we perform the insertion trials into each constituent vacancy, thus three values of $P_{in}(v)$ s are collected. For $P_{del}(v)$ s, we perform three separate simulations with each containing two NN vacancies. This is to collect $P_{del}(v)$ s of each constituent particle around or between two NN vacancies. We used the spherical shape for the sub-volume v , of which radius is 80% of the radius of a sphere inscribed into the WS cell. For HS MD simulations, a molecular simulator Etomica was used [23]. The MD time step was set to 0.04ps. We carried out three independent simulations to average the results and each was run for 2×10^6 time steps for data production while varying densities ranging from $\rho^* = 1.0376$ (melting point) to 1.2, where ρ^* is $M\sigma^3/V$ with the diameter of HS, σ taken as unity. To avoid the size effect for the introduction of trivacancy, we use 384 ($8 \times 8 \times 6$ unit cells) lattices constructed in a rhombic box. Note that the total volume must not be varied by trivacancy introduced into the system during simulation [76].

Figure 4.10 shows FCC and HCP crystal lattices, which spread out to exhibit $ABCABC...$ and $ABABAB...$ stacking planes, respectively, where A , B , and C are hexagonal planes. From their different spatial configurations, different types of trivacancies for FCC and HCP are obtained. For the FCC crystal, there are 4 types of

trivacancies, which hold the angles between separate vacancy sites as 180° , 120° , 90° and 60° , which are denoted as FCC3.1, 3.2, 3.3, and 3.4, respectively [76].

Fig. 4.10 (Color on line) Schematic of trivacancies in FCC with spread *ABC* planes and HCP with spread *ABA* planes. Filled and open circles represent possible vacant sites in trivacancy. The far-left plane (with three filled circles) is at the top of the stack, the plane B (with six filled circles) is in the middle, and the right plane (with three open circles) from it is the bottom one. On the right hand side, trivacancies are shown with the middle plane as a reference plane. If filled circles are not on vertices of hexagon, those are in the top plane. Open circles, which are on vertices of hexagon, are in the bottom plane. [Reprinted with permission from Ref. [76]. Copyright 2011, Taylor & Francis]

Unlike the FCC crystal that has one type of divacancy configuration due to full symmetry of the lattice, the HCP has two different types of divacancies, which are aligned in one plane (i.e. *A* or *B* plane) or aligned across two planes (i.e. *AB* planes) [74]. From permuting all possible locations for a *NN* vacancy, which can be introduced to

surrounding NN sites of each divacancy, we obtained 9 different types of trivacancies [85], which are denoted from HCP3.1 to HCP3.9, respectively. Due to the actual number of trivacancy is extremely small (See Table 4.12), we compare a negative logarithmic concentration $-\log x$ presented by Bennett and Alder in Figs. 4.11, 4.13, and 4.15. They collected the concentration of trivacancy of the equilateral triangular shape that is same as FCC3.4 (See Fig. 4.10). Our results agree well with theirs over the compressibility factors (i.e. PV_0/NkT) around 8 to 9, but their simulation method could not predict concentrations at high densities due to immobilities of vacancies. Their prediction for the concentration is based on the transition probability between orders of vacancies (i.e. mono-to-di, di-to-tri, tri-to-tetra, etc). It implies complete randomness in the creation of target polyvacancy. In this case, the simulation might have taken enormous time to collect reliable transition probabilities [76].

Fig. 4.11 (Color on line) Plot of $-\log x$ vs. the compressibility factor, PV_0/NkT for the FCC HS crystal. Bennett and Alder is Ref. [68]. Lines are guides to an eye. Errors are smaller than symbols.

[Reprinted with permission from Ref. [76]. Copyright 2011, Taylor & Francis]

From Fig.4.11, different concentrations are observed for different arrangements of trivacancies in the FCC HS crystal. Table 4.12 shows actual figures of trivacancies in both FCC and HCP HS crystals [76].

Table 4.12 Actual values of trivacancy concentrations for all types in FCC and HCP HS crystals at various densities

ρ^*	1.0376	1.04086	1.07000	1.10000	1.14000	1.17000	1.20000
FCC3.1	5.92×10^{-11}	4.03×10^{-11}	1.32×10^{-12}	2.08×10^{-14}	2.55×10^{-17}	4.37×10^{-20}	1.84×10^{-23}
FCC3.2	3.09×10^{-11}	1.93×10^{-11}	6.29×10^{-13}	1.05×10^{-14}	1.26×10^{-17}	2.36×10^{-20}	1.07×10^{-23}
FCC3.3	2.14×10^{-10}	1.47×10^{-10}	6.41×10^{-12}	1.38×10^{-13}	2.24×10^{-16}	4.37×10^{-19}	1.69×10^{-22}
FCC3.4	2.54×10^{-10}	1.84×10^{-10}	1.69×10^{-11}	1.04×10^{-12}	8.63×10^{-15}	9.36×10^{-17}	2.24×10^{-19}
HCP3.1	3.63×10^{-10}	2.39×10^{-10}	7.98×10^{-12}	1.39×10^{-13}	1.89×10^{-16}	3.67×10^{-19}	1.75×10^{-22}
HCP3.2	5.06×10^{-10}	3.46×10^{-10}	1.23×10^{-11}	2.19×10^{-13}	2.76×10^{-16}	5.27×10^{-19}	2.43×10^{-22}
HCP3.3	4.03×10^{-9}	2.77×10^{-9}	2.46×10^{-10}	1.52×10^{-11}	1.69×10^{-13}	2.31×10^{-15}	5.48×10^{-18}
HCP3.4	4.29×10^{-9}	3.09×10^{-9}	1.56×10^{-10}	4.70×10^{-12}	1.41×10^{-14}	5.19×10^{-17}	4.35×10^{-20}
HCP3.5	9.66×10^{-10}	6.53×10^{-10}	2.57×10^{-11}	4.87×10^{-13}	8.15×10^{-16}	1.75×10^{-18}	8.45×10^{-22}
HCP3.6	4.49×10^{-7}	3.11×10^{-7}	3.58×10^{-8}	3.19×10^{-9}	3.29×10^{-11}	2.23×10^{-13}	1.01×10^{-15}
HCP3.7	1.62×10^{-11}	1.06×10^{-11}	2.96×10^{-13}	4.40×10^{-15}	6.08×10^{-18}	1.26×10^{-20}	6.53×10^{-24}
HCP3.8	2.06×10^{-9}	1.49×10^{-9}	7.66×10^{-11}	1.90×10^{-12}	4.99×10^{-15}	1.53×10^{-17}	1.01×10^{-20}
HCP3.9	3.53×10^{-9}	2.67×10^{-9}	1.44×10^{-10}	4.19×10^{-12}	1.62×10^{-14}	5.12×10^{-17}	4.44×10^{-20}

Note: ρ^* is the dimensionless number density, $kT=m=\sigma=1$, $\rho^* = \rho\sigma^3$;

We found $P_{in}(v)$ s of three constituent vacancies in trivacancy are same for FCC3.4, but for other shapes of trivacancies, $P_{in}(v)$ of the center vacancy (See Fig. 4.9) in trivacancy smaller than those in the other two constituent vacancies. From this information, formation free energies of constituent vacancies are different even within the polyvacancy itself as well as stabilities of those vacancies. The equilateral triangular trivacancy exhibits the highest concentration among all trivacancy configurations, and it illustrates that the configuration is the most stable structure, also proven by the lowest formation free energy shown in Fig. 4.12 [76].

Fig. 4.12 (Color on line) Plot of the formation free energy of trivacancy $-f_3$ vs. the dimensionless density ρ^* for the FCC HS crystal. Lines are guides to an eye. Errors are smaller than symbols.[Reprinted with permission from Ref. [76]. Copyright 2011, Taylor & Francis]

More specifically, we can use the concept of driving force, which is to remove a vacant site by inserting a particle and $\mu - f$ (e.g. the exponent of Eq. (4.38) for monovacancy) is the identity of the force. Since it is always positive, the force becomes weak as the formation free energy decreases. In short, it is difficult to get rid of vacancies if the driving force is small and consequently, trivacancy with lower formation free energy becomes more stable. To have low formation free energy, constituent vacancies in trivacancy should have low $P_{in}(v)$ and high $P_{del}(v)$, which cause high equilibrium concentration. This observation is consistent with structural information of trivacancies.

Trivacancy structures of FCC3.1 and FCC3.2 are quite open, thus the insertion trial of a particle to fill a constituent vacancy is more easily acceptable. However, those of FCC3.3 and FCC3.4 have coagulated empty volume, which seems to be feasible to accept insertion trials but becomes more susceptible of embracing surrounding NN particles. Thus, insertions are hindered by moved-in NN particles so that low $P_{in}(v)$ s are produced [76].

We apply the same analysis to study trivacancies in the HCP HS crystal. Bennett and Alder's $-\log x$ (x is the trivacancy concentration) values for the FCC HS crystal were plotted in Fig. 4.13 for convenient comparison. In the HCP crystal, there are three types (i.e. HCP 3.3, HCP 3.6, and HCP 3.7) of the equilateral triangular trivacancies depending on arrangements of surrounding neighbor particles, but only one has the lowest formation free energy, which is HCP3.6. $P_{in}(v)$ s of constituent vacancies in HCP3.6 are much smaller than those in FCC3.4. The effect of this difference comes from different bottom planes (i.e. C or A , see Fig. 4.10), with which trivacancies are differently surrounded. It is clearly shown from Figs. 4.12 and 4.14 that trivacancy of HCP3.6 is more stable than that of FCC3.4. Therefore, the most stable trivacancy is found in the HCP HS crystal as HCP3.6 and its concentration becomes the largest [76].

Fig. 4.13 (Color on line) Plot of $-\log x$ vs. the compressibility factor, PV_0/NkT for the HCP HS crystal. Bennett and Alder is Ref. [68]. Lines are guides to an eye. Errors are smaller than symbols.[Reprinted with permission from Ref. [76]. Copyright 2011, Taylor & Francis]

Further, one of the deceptions is that $P_{in}(v)$ s of constituent vacancies in the equilateral triangular trivacancy are supposed to be same. However, we found one exception in the case of HCP3.3, where $P_{in}(v)$ of one of constituent vacancies is larger than those of the other two. It is interesting to observe that even with different $P_{in}(v)$ s of HCP3.3, its stability is greater than FCC3.4. The other equilateral triangular trivacancy (i.e. HCP3.7) contains constituent vacancies with same $P_{in}(v)$ s, but it has the highest formation free energy (i.e. less stable) so that the concentration becomes the lowest among other

trivacancies in the HCP HS crystal. It is worthy of note that overall concentrations averaged from all trivacancies for each crystal reveal higher trend in the HCP HS crystal as shown in Fig. 4.15 [76].

Fig. 4.14 (Color on line) Plot of the formation free energy of trivacancy $-f_3$ vs. the dimensionless density ρ^* for the HCP HS crystal. Lines are guides to an eye. Errors are smaller than symbols. Reprinted with permission from Ref. [76]. Copyright 2011, Taylor & Francis]

Fig. 4.15(Color on line) Plot of averaged $-\log x$ vs. the compressibility factor, PV_0/NkT for both FCC and HCP HS crystals. Bennett and Alder is Ref. [68]. Lines are guides to an eye. Errors are smaller than symbols.[Reprinted with permission from Ref. [76]. Copyright 2011, Taylor & Francis]

There have been few studies, which can provide a formulated framework to handle polyvacancy, but this work may offer a reliable approach to study it. With the information of structure and formation free energy of the vacancy, the developed formula reveals the equilibrium concentration of polyvacancy. The concept of the NN vacancy introduced here made the complex derivation in the statistical mechanical calculation much simpler. Applications of Eq.(4.38) were consistent with equations derived by Pronk and Frenkel [71] and Kwak *et al.* [74] for monovacancy and divacancy, respectively. To validate Eq.(4.38), we estimated the concentrations of trivacancies in FCC and HCP HS crystals as a case study with a help of MD simulation

incorporated with a modified Widom-like insertion method. In the course of predicting the equilibrium concentration, distinct structures of trivacancies were identified - 4 types in the FCC crystal and 9 types in the HCP- by permutation calculation, and their stabilities have been studied in details by investigating formation free energies. It was possible to identify the stability of trivacancy; angled form is more stable than straight one but exceptional cases exist due to different arrangements of surrounding neighbor particles. Quantification of relative stability is also possible and revealed that the most stable trivacancy is HCP3.6, which has the equilateral triangular shape. We have shown HS crystals as a study case, but it can be applied to any potential models. Finally, with the same approach taken in this work and the derived formula, one can actualize the specific study of vacancy so that it can serve to relate the smallest order of vacancy related to a void creation, theoretically. Then, fundamental questions risen in studies of voids such as interactions, closures, growth and coalescence, and migration [87-91], may not be difficult to answer after all [76].

Chapter 5 Computational Experiment of DSMC

Traditionally, in MD each particle in the domain of interest is modeled as being given the opportunity to interact with every other particle [70]. When particle collisions are coupled with their motions, the problem becomes very time-consuming. Not until the invention of DSMC method, proposed by Graham Bird [47], it becomes computationally less expensive and much more efficient so that DSMC shows two or three orders of magnitude faster than MD. In DSMC, the computational domain is divided into discrete cells, each of which consists of a certain number of particles. Upon realization of the DSMC algorithm, simulated particles are used to represent real molecules. They do not need to be a one-to-one relationship to real molecules. In a dimensioned computation, real molecules are often of the order of Avogadro's number $O(10^{23})$ while simulated particles are often of the order of $O(10^4)$ to $O(10^6)$, depending on number of cells used. Therefore each simulated particle represents about $O(10^{17})$ to $O(10^{19})$ molecules. In some cases the number of real molecules represented by a simulated molecule is not necessarily greater than one.

The DSMC algorithm can be either viewed as a simplified MD model, or a MC method for the integration of the time dependent Boltzmann equation. It uses a new sorting and collision sampling routine, which makes the algorithm very efficient. The

collision from DSMC is featured by stochastic simulation and converges to kinetic theory, which is the representation of dilute gases by averaging over a large number of samples.

5.1 General Procedure of DSMC

In general, DSMC contains an initialization step and a time loop, which are described as follows:

5.1.1 DSMC initialization

Before the DSMC flow solver begins, a number of initialization steps have to be taken. First, all the basic properties, such as density, initial temperature, system size, number of cells, number of pseudo-particles, mass and diameter of gas, should be initialized. Second, all the particles are distributed across the domain and their positions are allocated according to Eq. (5.1)

$$\begin{cases} x_i = (X_{\max} - X_{\min})R_f + X_{\min} \\ y_i = (Y_{\max} - Y_{\min})R_f + Y_{\min} \\ z_i = (Z_{\max} - Z_{\min})R_f + Z_{\min} \end{cases}, \quad (5.1)$$

where x_i , y_i , and z_i are the three Cartesian coordinates of particle i , R_f is a random number uniformly distributed in $[0,1]$, and $X_{\max}$ and $X_{\min}$, $Y_{\max}$ and $Y_{\min}$, $Z_{\max}$ and $Z_{\min}$ are upper and lower boundary of x , y , z direction respectively. In this case, a uniform grid is used. However, particles need not be distributed uniformly among the cells. Here in order to

generate the random number R_f , $\text{srand48}()$ and $\text{drand48}()$ functions are used under Linux G++ compiler. Third, after the positions of all the particles are set, the initial velocities of them are given. The DSMC method is essentially equivalent to the Boltzmann equation, which for any unsteady system, is dependent on a set of initial conditions. Initially the flow is taken to be in thermal equilibrium. The particles are set to follow the Boltzmann-Maxwell distribution. Since the velocity of each particle is the sum of the mean velocity of the flow and the thermal velocity, and the initial mean flow velocity is already known from the beginning, sampling from the Maxwellian velocity distribution is only left.

The Maxwell-Boltzmann distribution describes how the motions of the microscopic molecules are related to macroscopic, measurable properties like temperature T . One of the important assumptions of this distribution is that the components of velocity are distributed independently of each other. Based on the conservation laws of mass, energy and momentum, given conditions

$$\begin{aligned}\int f d\mathbf{v} &= n \\ \int f \mathbf{v} d\mathbf{v} &= n\mathbf{u} \\ \int f m v^2 d\mathbf{v} &= n m u^2 + 3 n k_B T,\end{aligned}\tag{5.2}$$

where n is the number density, $\mathbf{v}$ is the velocity vector, $\mathbf{u}$ is the macroscopic velocity vector of the flow, m is the mass of particle, and k_B is the Boltzmann constant, the velocity distribution function f can be derived and the final result is as follows:

$$f(\mathbf{v}) = 4\pi n v^2 \left(\frac{m}{2\pi k_B T}\right)^{\frac{3}{2}} \exp\left(-\frac{m(\mathbf{v} - \mathbf{u})^2}{2k_B T}\right).\tag{5.3}$$

In the Maxwell-Boltzmann distribution there are three important properties. Those are most probable speed (v_p), average speed ($\langle v \rangle$) and root-mean-square speed (v_{rms}). The most probable speed is associated with the highest point in the distribution. It can be derived by taking the derivative of Eq. (5.3) with respect to $\mathbf{v}$ to be 0. The result is shown as follows:

$$v_p = \sqrt{\frac{2k_B T}{m}}. \quad (5.4)$$

The average speed is the sum of all the speeds of molecules divided by the number of particles. It can be derived by taking the integration:

$$\langle v \rangle = \int v f(v) dv = \sqrt{\frac{8k_B T}{\pi m}}. \quad (5.5)$$

The root-mean-square speed is expressed as follows:

$$v_{rms} = \sqrt{\int v^2 f(v) dv} = \sqrt{\frac{3k_B T}{m}}. \quad (5.6)$$

From Eq. (5.4) we can see that, the Maxwell-Boltzmann distribution is the Gaussian distribution [92] with the mean $\mu = \mathbf{u}$ and standard deviation $\sigma = \sqrt{\frac{k_B T}{m}}$, then the Box-Muller method [93, 94] is applied to generate the Gaussian distribution, which satisfies the prescribed mean and standard deviation.

$$z = \mu + \sigma \sqrt{-2 \log(1 - R_f)} \cos(2\pi R_f). \quad (5.7)$$

5.1.2 DSMC time loop

After the initialization is finished, the DSMC time loop starts. It is a four-step process, which comprises move, sort, collide and sample. These four subroutines are carried out repeatedly in each time loop. The procedure will be discussed one by one in detail. In the move subroutine, particles are moved to the new positions according to their old positions and velocities as follows,

$$\begin{cases} x_i' = x_i + v_x \Delta t \\ y_i' = y_i + v_y \Delta t, \\ z_i' = z_i + v_z \Delta t \end{cases} \quad (5.8)$$

where x_i , y_i , and z_i represent the three old Cartesian coordinates, x_i' , y_i' , and z_i' represent the three new Cartesian coordinates, v_x , v_y , and v_z represent three velocity components, and Δt represents the time interval. For example of taking a single direction of the movement of the particle, the flowchart for the move subroutine is shown in Fig. 5.1.

Fig. 5.1 The flowchart of move subroutine

As seen from the above flowchart, when a particle reaches the edge of the domain, the boundary condition is applied to generate a new velocity and a new position for the particle. Typically there are three types of boundary conditions: periodic boundary, specular boundary and diffuse boundary. The periodic boundary is a representation of an infinitely large periodic domain. In this boundary condition, particles moving out of one boundary can be regarded as moving in from another boundary. The position is displaced by the length of the corresponding direction of the domain, i.e. if new position $x' > X_{\max}$, then the particle will be placed to the position $x' - (X_{\max} - X_{\min})$; if $x' < X_{\min}$, then the particle will be placed to the position $x' + (X_{\max} - X_{\min})$. The velocity remains unchanged. In specular boundary condition, particles reaching the edge will be reflected to a new position, which is the specular image of the position where the particle should have arrived without

boundary, i.e. if $x' > X_{max}$, particle will be placed to the position $2X_{max} - x'$; if $x' < X_{min}$, particle will be placed to the position $2X_{min} - x'$. If the wall is stagnant, the direction of the velocity becomes its opposite. The diffuse boundary condition is the most realistic yet complicated model. Each particle that reached the edge will be given a velocity with new magnitude and direction.

After all particles are moved to new positions and boundary conditions have been applied, the next step is to sort them into different cells. The pseudo-code for the sort subroutine is shown in Fig. 5.2.

```

// Find the cell address for each particle
for i ∈ [0,NPART)
{
    cell_x ← xi/Δx
    cell_y ← yi/Δy
    cell_z ← zi/Δz
    jx[i] ← NX*NY*cell_z + NX*cell_y + cell_x
}

// Count the number of particles in each cell
for j ∈ [0,NX*NY*NZ)
{ cell_n[j] = 0
}
for i ∈ [0,NPART)
{ cell_n[ jx[i] ]++
}

// Build index list as cumulative sum of the
// number of particles in each cell
m=0
for j ∈ [0,NX*NY*NZ)
{
    index[j] ← m
    m ← m + cell_n[j]
}

// Build cross-reference list
for j ∈ [0,NX*NY*NZ)
{ temp[j] = 0
}
for i ∈ [0,NPART)
{
    jcell ← jx[i] // Cell address of ipart
    k ← index[jcell] + temp[jcell]
    Xref[k] ← i
    temp[jcell] ← temp[jcell] + 1
}

```

Fig. 5.2 Pseudo-code for sort subroutine

First, the cell address for each particle is found. *Cell_x*, *cell_y*, and *cell_z*, these three temporary variables represent the cell index in three Cartesian coordinates. Then the cell address $NX*NY*cell_z + NX*cell_y + cell_x$ (i.e. *NX*, *NY*, and *NZ* represent total number of cells in each coordinate directions respectively) is found and put in the array *jx*.

Second, the number of particles in each cell is initialized to 0 for the consequent action that looping over the particles makes sure the particle number is accumulated in each cell. Third, an index list has to be built. The variable m is initialized to 0 first and added to the j th element of array $cell_n$, which represents the number of particles in each cell. When the value of m is given to the j th element of array $index$, the $index$ is pointed to the first particle in each cell. So far the index list building has been finished. Fourth, array $temp$ represents relative index of each particle in each cell. Its elements are initialized to 0 and incremented by 1 after every loop over the particles. Then let variable $k = index[jcell] + temp[jcell]$, where $jcell$ is the cell index of particle i . Afterwards, another array $Xref$ is created. Let the k th element of $Xref$ point to particle i . The cross-reference index is built. It will be used in the collide subroutine.

The sort subroutine provides an alternative way of processing collisions within each cell rather than dealing with successive collisions in MD. This process reduces the time complexity from $O(N^2)$ to $O(N)$. In the collide subroutine, numerous collisions take place between particles in the same cell regardless of where each individual particle locates within the cell. The probability of a collision between particle i and j is proportional to their relative speed [95]

$$(P_{col})_{i,j} = \frac{|\mathbf{u}_i - \mathbf{u}_j|}{\sum_{m=1}^{N_p} \sum_{n=1}^{m-1} |\mathbf{u}_m - \mathbf{u}_n|}, \quad (5.9)$$

where N_p is the number of particles in the cell. The double summation represents the sum of relative speed of all the particle pairs in the cell. Due to the high computational expense of this double summation, an alternative strategy is used. In order to perform the

actual collisions, a number of attempted collisions are computed. The number of collision pairs to be selected for possible collisions in each time step is expressed as:

$$n_{cp} = \frac{N_p^2 \pi d^2 v_{rm} \Delta t}{2 \Delta x \Delta y \Delta z}, \quad (5.10)$$

where v_{rm} denotes the maximum relative speed inside a cell and is updated in each time step; d denotes the diameter of molecule; Δt is the time step; Δx , Δy , and Δz are the cell length in each direction, respectively.

After the number of candidate collision pairs is determined, one can loop over the total number of possible collision pairs and pick out two particles for collision out of the cell stochastically. Here an acceptance-rejection mode, which represents the MC characteristic, is used. If the collision probability between the collision pairs is expressed as follows:

$$P_{col} = \frac{v_r}{v_{rm}}. \quad (5.11)$$

If the quantity P_{col} is large enough (i.e. greater than the random number generated by the computer), then the collision is accepted; otherwise, it is rejected. If the collision is accepted, post-collision velocities are to be determined according to the conservation of momentum and energy expressed in Eqs. (3.13) and (3.14). Here we introduce the centre of mass velocity $\mathbf{v}_m$ and pre- and post- relative velocities $\mathbf{v}_r$ and $\mathbf{v}_r'$:

$$\mathbf{v}_m = \frac{m_1 \mathbf{v}_1 + m_2 \mathbf{v}_2}{m_1 + m_2}, \quad (5.12)$$

$$\mathbf{v}_r = \mathbf{v}_1 - \mathbf{v}_2, \quad (5.13)$$

$$\mathbf{v}_r' = \mathbf{v}_1' - \mathbf{v}_2'. \quad (5.14)$$

Then Eqs. (3.13) and (3.14) give the following solution

$$\mathbf{v}_1' = \mathbf{v}_m + \frac{m_2}{m_1 + m_2} \mathbf{v}_r', \quad (5.15)$$

$$\mathbf{v}_2' = \mathbf{v}_m - \frac{m_1}{m_1 + m_2} \mathbf{v}_r'. \quad (5.16)$$

In Eqs. (3.13) and (3.14), there are four equations with six unknowns (i.e. three components of $\mathbf{v}_1'$ and $\mathbf{v}_2'$), thus two other parameters are needed. Since the magnitude of relative velocity is unchanged by the collision,

$$|\mathbf{v}_r'| = |\mathbf{v}_r|. \quad (5.17)$$

In order to solve $\mathbf{v}_r'$ we generate two random numbers R_f1 and R_f2 in the range $[0,1]$ to obtain two azimuth angles θ and φ in three-dimensional space.

$$\begin{aligned} \cos \theta &= 1 - 2R_f1 \\ \sin \theta &= \sqrt{1 - \cos^2 \theta} \\ \varphi &= 2\pi R_f2. \end{aligned} \quad (5.18)$$

After two azimuths are determined, the three components of $\mathbf{v}_r'$, denoted as v_{rx}' , v_{ry}' and v_{rz}' , are finalized.

$$\begin{aligned}
v'_{rx} &= |\mathbf{v}_r| \cos \theta \\
v'_{ry} &= |\mathbf{v}_r| \sin \theta \cos \varphi \\
v'_{rz} &= |\mathbf{v}_r| \sin \theta \sin \varphi.
\end{aligned} \tag{5.19}$$

Once the particle positions and velocities are obtained from the previous three subroutines, the last step is to sample macroscopic properties at each cell location. The sample subroutine is not called every time step but at some user-specified steps, e.g. after initial transient transition to equilibrium. Usually in the simulation 200,000 steps are used to ensure the equilibrium is achieved. The number density is calculated as follows:

$$n = \frac{N_p}{\Delta x \Delta y \Delta z}. \tag{5.20}$$

The three velocity components are computed as:

$$\begin{aligned}
v_x &= \frac{1}{N_p} \sum_{i=1}^{N_p} v_{xi} \\
v_y &= \frac{1}{N_p} \sum_{i=1}^{N_p} v_{yi} \\
v_z &= \frac{1}{N_p} \sum_{i=1}^{N_p} v_{zi}.
\end{aligned} \tag{5.21}$$

Since in the Maxwell-Boltzmann distribution the variance of velocity equals $\frac{3k_B T}{m}$, then

the temperature can be calculated from the variance, which is formulated as:

$$T = \frac{m}{3k_B} \left(\frac{1}{N_p} \sum_{i=1}^{N_p} (v_{xi}^2 + v_{yi}^2 + v_{zi}^2) - \left(\frac{1}{N_p} \sum_{i=1}^{N_p} v_{xi} \right)^2 - \left(\frac{1}{N_p} \sum_{i=1}^{N_p} v_{yi} \right)^2 - \left(\frac{1}{N_p} \sum_{i=1}^{N_p} v_{zi} \right)^2 \right). \tag{5.22}$$

5.2 Couette flow

In order to validate the DSMC algorithm, we choose Couette flow as a test case. We set up our model as the laminar flow of a viscous fluid between two parallel plates moving in opposite direction in y axis. As is known to all, in Couette flow the velocity component u_y is a linear function of x . The velocity profile is shown in Fig. 5.3.

Fig. 5.3 The velocity profile

As we can see from this profile, u_y is a linear function of x , which is consistent with the derivation from the Navier-Stokes equation. Also, the temperature profile and the number density profile are shown in Figs. 5.4 and 5.5 as well.

Fig. 5.4 The temperature profile

Fig. 5.5 The number density profile

The velocity correlation function [96] (VCF) describes how strongly the particle velocity at some time t is correlated to its initial value at time $t=0$. The normalized velocity correlation function can be formulated as:

$$VCF = \frac{\langle \mathbf{v}(t) \bullet \mathbf{v}(0) \rangle}{\langle v^2(0) \rangle}. \quad (5.23)$$

At time $t=0$, mathematically $\mathbf{v}(t)$ equals $\mathbf{v}(0)$ then the numerator in Eq. (5.23) equals the denominator and the VCF equals 1, as shown in Fig. 5.6. This means the velocity is perfectly correlated with the initial one. As time elapses, the particle experiences collisions, which change the particle's speed and direction. The instantaneous velocity vector becomes less and less correlated with its original value. After enough collisions the particle velocity becomes uncorrelated with its original value and the average dot-product between the current and original value becomes zero.

Fig. 5.6 Velocity correlation function

5.3 Consistent Boltzmann Algorithm

The DSMC algorithm leads to transport properties for dilute HS gas while it has an ideal gas equation of state. Modifications were made by Alexander, Garcia and Alder [56] and as a result, the consistent Boltzmann Algorithm (CBA) comes out. The CBA introduces a

displacement term in the advection process to remove the inconsistency that traditional DSMC encounters and recovers equation of state for HS gas.

In standard DSMC, once a pair is selected for a collision, their post-collision velocities are determined stochastically, consistent with the conservation of energy and momentum. The collision is executed with particles remaining in place. In the Boltzmann equation the advection process corresponds to point particles, so the virial is zero, which corresponds to the ideal gas. To obtain the correct HS virial, the CBA includes the extra displacement in the advection step that the particles would have experienced if they had collided as hard spheres [56],

$$\mathbf{d} = \frac{\mathbf{v}_r' - \mathbf{v}_r}{|\mathbf{v}_r' - \mathbf{v}_r|} \sigma, \quad (5.24)$$

where σ is the molecular diameter, and $\mathbf{v}_r'$ and $\mathbf{v}_r$ are the post-collision and pre-collision relative velocities, respectively. Particle 1 is displaced by vector distance $\mathbf{d}$ and particle 2 by $-\mathbf{d}$, as shown in Fig. 5.7.

Fig. 5.7 Schematic diagram of the displacement term [56]

Then after collision particles are displaced by following relations [55],

$$\begin{aligned}\mathbf{r}_1(t + \Delta t) &= \mathbf{r}_1(t) + \mathbf{v}_1'(t)\Delta t + \mathbf{d} \\ \mathbf{r}_2(t + \Delta t) &= \mathbf{r}_2(t) + \mathbf{v}_2'(t)\Delta t - \mathbf{d},\end{aligned}\tag{5.25}$$

where Δt is the time step, $\mathbf{r}_1$ and $\mathbf{r}_2$ are the position vectors of particle 1 and 2 respectively, and $\mathbf{v}_1'$ and $\mathbf{v}_2'$ are the post-collision velocities of particle 1 and 2 respectively. The shear viscosity η is evaluated in terms of an autocorrelation function [97-99]

$$\eta = \frac{1}{Vk_B T} \int_0^\infty \langle J_{xy}(t) J_{xy}(t+s) \rangle dt, \tag{5.26}$$

where

$$J_{xy} = m \left[\sum_{i=1}^N u_i v_i + \sum_{c=1}^\infty (y_i - y_j) \Delta^c u_i \delta(t - t_c) \right] \tag{5.27}$$

is the xy -component of the microscopic stress tensor. In Eq. (5.27), y_i is the y -coordinate of the particle i at time t ; u_i and v_i are the x , y -component of its velocity. The change in the x -component of velocity for particle i is $\Delta^c u_i$, where the index c represents a collision between particle i and j , which happens at time t_c . Applying the CBA combined with Eqs. (5.26) and (5.27), we calculate the shear viscosity as shown in Table 5.1.

Table 5.1 Data of shear viscosity

n^*	η_{HS}	$\eta_{\text{vdw}}/\eta_{\text{HS}}$	η_{vdw}	η_{sim}	$\eta_{\text{E}}/\eta_{\text{HS}}$	η_{E}
0.0141	0.1820	0.9630	0.1753	0.1760	0.9900	0.1802
0.0707	0.1870	0.9300	0.1739	0.1750	1.0000	0.1870
0.2828	0.2730	0.8970	0.2449	0.2420	1.0100	0.2757
0.4714	0.5120	1.0440	0.5345	0.5360	0.9800	0.5018
0.7071	1.3710	1.2740	1.7467	1.7480	0.9090	1.2462
0.7857	1.8980	1.3740	2.6079	2.6060	0.9090	1.7253
0.8839	3.7780	1.1360	4.2918	4.2930	0.6940	2.6219
0.9428	7.3460	0.7890	5.7960	5.7970	0.4630	3.4012

Note: $kT=m=\sigma=1$, $n^*=n\sigma^3$, η_{HS} is from MD HS simulations [98]; η_{VDW} is from van der Waals CBA viscosity [55]; η_{sim} is our computed value; η_{E} is from Enskog theory [55].

The plot of these data is shown as follows:

Fig. 5.8 Shear viscosity as a function of number density (with $kT = m = \sigma = 1$, $n^* = n\sigma^3$)

Also, we compute the self-diffusion coefficient, which is started from the Green-Kubo relation [100],

$$D = \int_0^\infty \langle v_{xi}(t) \bullet v_{xi}(0) \rangle dt. \quad (5.28)$$

By using a simple numerical quadrature, we obtain the results of the self-diffusion coefficient at different number densities as shown in Table 5.2.

Table 5.2 Data of self-diffusion coefficient

n^*	D_{HS}	D_{vdw}/D_{HS}	D_{vdw}	D_{sim}	D_{sim}/D_{HS}	D_E/D_{HS}	D_E
0.0141	14.9810	0.9990	14.9660	14.9880	1.0005	0.9990	14.9660
0.0707	2.8210	0.9840	2.7759	2.7780	0.9848	0.9830	2.7730
0.2828	0.5990	0.8880	0.5319	0.5340	0.8915	0.8490	0.5086
0.4714	0.2930	0.9880	0.2895	0.2870	0.9795	0.7580	0.2221
0.7071	0.1160	2.5150	0.2917	0.2930	2.5259	0.8000	0.0928

Note: $kT=m=\sigma=1$, $n^*=n\sigma^3$, D_{HS} is from MD HS simulations [101]; D_{VDW} is from van der Waals CBA self-diffusion coefficient [55]; D_{sim} is our computed value; D_E is from Enskog theory [55].

The plot of these data is shown as follows:

Fig. 5.9 Self-diffusion as a function of number density (with $kT = m = \sigma = 1$, $n^* = n\sigma^3$)

From the calculations of these two examples, we can see that our computational results are consistent with Garcia's CBA algorithm, which verifies the validity of our code.

5.4 Modified Enskog Method

We propose a modified direct simulation Monte Carlo (DSMC) method, called modified Enskog method, which lifts the low-density limit existed in the application of the original DSMC. To assess this adapted method, transport properties of hard-sphere systems have been predicted both at dense states as well as dilute, and we observed the excellent accuracy over existing DSMC-based algorithms including the Enskog theory. More interestingly, its background principle appears very comparable to the acceleration of the molecular dynamics in mesoscale, thus the modified DSMC may hold a key role to bridge the gap between stochastic and deterministic methods in large scale.

An overwhelming majority of fluid flow problems of physical and engineering interest cannot be solved using atomistic simulation methods, such as molecular dynamics (MD), due to the enormous number of degrees of freedom constituting the macroscopic systems [6]. While high performance and progress in large-scale distributed computing has made MD as an acceptable tool for engineering design based on equilibrium properties [102, 103], for hydrodynamics usage of MD is not so obvious because of computational cost in probing long time and large length-scale dynamics. Thus, for most of engineering flows simulation algorithms based on continuum Navier-Stokes equation were used as modeling tools [43]. On the other hand, the role of molecular simulation algorithm in fluid dynamics was either limited to determine equilibrium and transport properties or better constitutive models for complex fluid [104-106].

However, in recent decades, tremendous effort has been put into the development and implementation of fluid simulation algorithm based on real or fictitious MD. Broadly

classified as mesoscale simulation tools, Lattice Boltzmann method (LBM), Dissipative Particle Dynamics (DPD), stochastic rotation dynamics are a few successful simulation methodologies developed out of these endeavors [107-109]. These methods are quite successful in emerging fields such as micro and nanofluidics, where very often continuum hypothesis breaks down and pure microscopic simulations are too expensive to be useful. The basic underlying principle behind these mesoscale simulation methodologies is that for many body dynamics often precise details of underlying microscopic dynamics is not so important and same macroscale dynamics can emerge out of collective motion of widely different microscopic rules. Thus, the basic goal in mesoscale modeling is to find out computationally convenient microscopic equation of motion, sufficient to describe macroscale dynamics of interest. In this regards, particle-based methods such as DPD has taken the viewpoint that a particle in mesoscale system is a coarse-grained description of pure microscopic system and in reality each mesoscale particle describes an ensemble of microscopic particles. Thus, the equation of motion should be obtained via proper coarse-graining of microscopic dynamics. Once we accept this viewpoint, systematic coarse-grained tools developed in kinetic theory and non-equilibrium statistical physics provide mathematical framework to develop mesoscale tools. A series of work in this direction has established that a self-consistent formulation of discrete version of non-equilibrium statistical mechanics is possible with well defined microscopic link. For example, in case of DPD microscopic basis, the method in terms of the underlying Fokker-Planck model is clarified [110]. Similarly, in case of LBM, the link with Boltzmann equation is established and the role of H-theorem in case of discrete kinetic theory is emphasized.

It can be safely said that developing simulation tools based on kinetic theory in general and Boltzmann equation in particular is one of the important goals of mesoscale simulation methods. Probably, one of the first such mesoscale simulation methods was developed by Bird in the context of aerospace problems and this tool was applied with remarkable success in highly non-equilibrium situation such as supersonic flows and rarefied gas dynamics in micro-channels [111]. Thus for rarefied flow physics, which is predicted extremely well by Boltzmann equation even in states very far from equilibrium [112], we have wide spectrum of mesoscale tools ranging from Direct Simulation Monte Carlo (DSMC) for high Mach number (i.e. ratio of flow speed to the thermal speed) and DPD for moderately low Mach number to LBM for extreme sub-sonic flows.

Unfortunately, in dense gas regime this connection between particle-based mesoscale methods and kinetic theory via Enskog theory (or revised Enskog theory (RET)[113]) is not so obvious. Similar to Boltzmann's model of dilute gas, motions of particles in these models are decomposed into two parts: propagation at constant velocity followed by collisions in which exchange of momentum between particles occurs. However, unlike Boltzmann model, collisions are understood to be non-local events due to the presence of finite size particles. The non-local nature of collision is difficult to handle both for Monte Carlo (MC) method as well as for kinetic modeling via simplified phenomenological theories of BGK type.

A first attempt to create particle-based mesoscale method for dense gas was made by consistent Boltzmann algorithm (CBA) [56], where the crucial ingredient of non-local collision was introduced via additional displacement in the advection process and

enhanced collision rate due to denseness. In its application to hard sphere (HS) fluid flows with finite Knudsen numbers, CBA removes usual inconsistency, which the traditional DSMC method bears: it yields transport properties for a dilute HS gas yet has an ideal gas equation of state and recovers the HS equation of state with virtually no additional computational cost. This model was able to predict dense gas property in a qualitatively correct fashion. However, the quantitative agreement between this modified DSMC and Enskog or MD results is not possible. It was argued that the reason for the quantitative disagreement is absence of backscattering events connected with the structural effects in the model, and this becomes severe at high density regime.

In this work, we show that the basic idea that non-local collision or modified displacement in the advection process will have same effect for collective dynamics can be made quantitatively correct. We show that the basic ingredient missing in CBA is a proper modification of molecular velocity in advection step due to denseness. Once this effect is taken care correctly, the model is able to reproduce result of MD fairly accurate.

In our proposed algorithm, the advection force that particles experience after collision is reevaluated by a dimensionless factor χ , which takes the density-expanded terms excluding the ideal gas term in the compressibility factor Z (i.e. $\chi = Z - 1$) that is formulated to provide the post-collision velocity through $\chi = (V - V_0)/(V_0 - U)$, where V is the post-collision velocity, V_0 is the pre-collision velocity, and U is the average velocity. In actual simulation, particles whose velocities are larger than the average velocity will be accelerated but whose velocities are smaller than the average velocity will be decelerated. This will disperse the particles according to the velocity distribution.

The pre-and post-collision relative velocity is proportional to the pre-and average relative one, and that amount is chosen to be same as the non-ideal (i.e. excess) portion deviated from the ideal gas state. This algorithm, when applied to the HS system, yields correct HS virial beyond the 2nd order consistency, which is what Alexander *et al* [56] claimed in the work of CBA. Thus, we can state that the proposed algorithm advances the dispersion of particles towards the right nature of the physical system even with non-deterministic DSMC simulation.

To verify our argument, we obtained the pressure tensor of HS system from virial theorem in collision within DSMC simulation and evaluated transport properties, which are self-diffusion coefficient and shear viscosity, as well as single particle velocity autocorrelation over wide range of densities in the fluid region. Figure 5.10 shows results of relative self-diffusion coefficients obtained from methods of CBA, Enskog and our algorithm.

Fig. 5.10 HS self-diffusion coefficient normalized by MD self-diffusion coefficient as a function of number density (with $kT = m = \sigma = 1$, $n^* = n\sigma^3$)

The self-diffusion coefficient of our algorithm shows best consistency with that of MD and with much improvement against the prediction from CBA, especially at high density. The shear viscosity also confirms the same trend, which is drawn from Figure 5.11.

Fig. 5.11 HS shear viscosity normalized by MD shear viscosity as a function of number density
(with $kT = m = \sigma = 1, n^* = n\sigma^3$)

However, when the density of the HS fluid closes to the vicinity of freezing density, we observed severe deviations from shear viscosity values of MD. The main factor of this discrepancy may come from a deficient set-up of the collision frequency, which is based on Alexander *et al*'s collision algorithm [15]. At higher densities the collision rate in a HS gas is controlled by $\Gamma = \Lambda Y$, where Λ is Boltzmann collision rate and Y is known from MC simulation. Note that the Boltzmann collision rate's origin is from low density. Our algorithm does not contain self-modulated controller for the collision rate, but gradual adaptation of actual collision rate of HS MD is also possible so that the discrepancy might be finally suppressed. By calculating the pressure tensor based on the virial theorem, the pressure values of CBA, MD and our algorithm were plotted in Figure 5.12, which shows overall better agreement of our algorithm.

Fig. 5.12 Pressure as a function of number density (with $kT = m = \sigma = 1$, $n^* = n\sigma^3$)

Furthermore, the single particle velocity autocorrelation function (SVCF) decays as $t^{-3/2}$ (i.e. as it should) which is shown in Figure 5.13.

Fig. 5.13 Single particle velocity correlation function (SVCF) as a function of time in log-log plot

In conclusion, we have proposed a modified DSMC algorithm for the Boltzmann equation, which makes DSMC valid for not only rarefied HS system but also dense HS one. This new algorithm gains an obvious advantage over existing DSMC-based algorithms as well as Enskog theory in prediction of transport coefficients.

Chapter 6 Conclusions and Outlook

6.1 Conclusions

In this thesis, we have introduced two different projects: calculation of polyvacancy concentration in crystals and a modified DSMC algorithm in micro- and meso-scale in dimension, which use molecular dynamics and direct simulation Monte Carlo technique, respectively. For the DSMC project, it can be applied in rarefied gas dynamics. For example, space re-entry of vehicles, gaseous flow in micro-devices. A few possible emerging uses are: High Mach turbulence, micro-macro coupling. For the polyvacancy project, this is a fundamental work based on creation of knowledge about how to predict thermodynamic properties of vacancies. It is worth in thermodynamics. For actual applications, we can study the relation of melting of solids with effects of vacancy concentration. Vacancy concentration affects strain of solid, fractures, other defective material formations. We can also predict the properties of new materials, such conductivity and ductility.

In the first project, the grand canonical ensemble is used to derive the monovacancy concentration. The insertion-deletion biased method developed by Pronk and Frenkel is applied to obtain the formation free energy of vacancy. Based on this, we derive the generalized equation to calculate the vacancy concentration in crystals with any number of vacancies. With the collision-based Molecular Dynamics (MD), using the insertion-

deletion bias method, we apply this equation to 4 types of FCC and 9 types of HCP, carry out the simulation and provide more comprehensive results than predecessors.

In the second project, we try to solve the problems in a fluid flow with the DSMC algorithm which is widely used in a rarefied gas. We use DSMC method to model fluid and construct an additional term in traditional DSMC, which breaks the inconsistency that exists in traditional DSMC algorithm and maintains high accuracy even at high densities which cannot be achieved by all previous algorithms even including consistent Boltzmann algorithm. We build up the new relationship of transport properties and density, and create a new thermodynamic equation.

6.2 Outlook

In our current work we have used two major molecular simulation techniques: Monte Carlo and Molecular Dynamics whereas they are applied in different fields. In the future these two techniques can be combined together. For instance, we can develop hybrid MD-DSMC simulation which so far has been used in microfluidics modeling [114] and development of kinetic nucleation model [115], and apply it to single to multiphase systems, interfacial problems, shock wave problems and so on.

References

- [1] M.N. Čířier, Chaotic motion in classical fluids with scale relativistic methods, *Journal of Mathematical Physics*, 50 (2009).
- [2] M. Guzzo, E. Lega, C. Froeschlé, On the numerical detection of the effective stability of chaotic motions in quasi-integrable systems, *Physica D: Nonlinear Phenomena*, 163 (2002) 1-25.
- [3] P.C. Müller, J. Bajkowski, D. Söffker, Chaotic motions and fault detection in a cracked rotor, *Nonlinear Dynamics*, 5 (1994) 233-254.
- [4] L. Palkovics, P.J.T. Venhovens, Investigation on stability and possible chaotic motions in the controlled wheel suspension system, *Vehicle System Dynamics*, 21 (1992) 269-296.
- [5] F. Peterka, J. Vacík, Transition to chaotic motion in mechanical systems with impacts, *Journal of Sound and Vibration*, 154 (1992) 95-115.
- [6] D. Frenkel, B. Smit, *Understanding Molecular Simulation From Algorithms To Applications*, academic press, San Diego, 1996.
- [7] R. Brun, F. Rademakers, ROOT - An object oriented data analysis framework, *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 389 (1997) 81-86.
- [8] S.R. Chidamber, C.F. Kemerer, Metrics suite for object oriented design, *IEEE Transactions on Software Engineering*, 20 (1994) 476-493.
- [9] Y. Shoham, Agent-oriented programming, *Artificial Intelligence*, 60 (1993) 51-92.

- [10] D. Gregor, J. Järvi, J. Siek, B. Stroustrup, G. Dos Reis, A. Lumsdaine, Concepts: Linguistic support for generic programming in C++, 2006, pp. 291-310.
- [11] A. Singleton, Genetic programming with C++, Byte, 19 (1994).
- [12] H.M. Deitel, P.J. Deitel, Java How to Program, 7th ed., Prentice Hall Inc., New Jersey, 2007.
- [13] J. Hannemann, G. Kiczales, Design pattern implementation in Java and AspectJ, 2002, pp. 161-173.
- [14] A. Naressi, C. Couturier, J.M. Devos, M. Janssen, C. Mangeat, R. De Beer, D. Graveron-Demilly, Java-based graphical user interface for the MRUI quantitation package, Magnetic Resonance Materials in Physics, Biology and Medicine, 12 (2001) 141-152.
- [15] C. Heinlein, Global and local virtual functions in C++, Journal of Object Technology, 4 (2005) 71-93.
- [16] B. Eckel, thinking in C++ second edition, Prentice Hall Inc., New Jersey, 2000.
p511-574.
- [17] R.S. Weiner, L.J. Pinson, Practical example of multiple inheritance in C++, SIGPLAN Notices (ACM Special Interest Group on Programming Languages), 24 (1989) 112-115.
- [18] C. De Dinechin, C++ exception handling, IEEE Concurrency, 8 (2000) 72-79.
- [19] H.F. Bucher, A.J. Schultz, D.A. Kofke, An Eclipse-based environment for molecular simulation, 2005, pp. 130-134.
- [20] D.A. Kofke, B.C. Mihalick, Web-based technologies for teaching and using molecular simulation, Fluid Phase Equilib., 194-197 (2002) 327-335.

- [21] E.J. Maginn, J.R. Elliott, Historical perspective and current outlook for molecular dynamics as a chemical engineering tool, *Industrial and Engineering Chemistry Research*, 49 (2010) 3059-3078.
- [22] A.J. Schultz, D.A. Kofke, Molecular simulation using a graphical user interface, 2005, pp. 3642.
- [23] D.A. Kofke, B.C. Mihalick, Web-based technologies for teaching and using molecular simulation, *Fluid Phase Equilibria*, 194-197 (2002) 327-335.
- [24] B.C. Mihalick, Development of a standalone Java-based molecular simulation environment, State University of New York at Buffalo, 2001, Master Thesis.
- [25] W.K. Hastings, Monte carlo sampling methods using Markov chains and their applications, *Biometrika*, 57 (1970) 97-109.
- [26] P.J. Green, Reversible jump Markov chain monte carlo computation and Bayesian model determination, *Biometrika*, 82 (1995) 711-732.
- [27] A. Doucet, S. Godsill, C. Andrieu, On sequential Monte Carlo sampling methods for Bayesian filtering, *Statistics and Computing*, 10 (2000) 197-208.
- [28] A.M. Ferrenberg, R.H. Swendsen, New Monte Carlo technique for studying phase transitions, *Physical Review Letters*, 61 (1988) 2635-2638.
- [29] A.M. Ferrenberg, R.H. Swendsen, Optimized Monte Carlo data analysis, *Physical Review Letters*, 63 (1989) 1195-1198.
- [30] M. Bullynck, L. de Mol, Setting-up early computer programs: D. H. Lehmer's ENIAC computation, *Archive for Mathematical Logic*, 49 (2010) 123-146.
- [31] W.B. Fritz, ENIAC - a problem solver, *IEEE Annals of the History of Computing*, 16 (1994) 25-45.

- [32] C. Rope, ENIAC as a stored-program computer: A new look at the old records, *IEEE Annals of the History of Computing*, 29 (2007) 82-87.
- [33] N. Metropolis, the beginning of the Monte Carlo method, *Los Alamos science special issue*, (1987).
- [34] A. Doucet, A.M. Johansen, V.B. Tadić, On solving integral equations using Markov chain Monte Carlo methods, *Applied Mathematics and Computation*, 216 (2010) 2869-2880.
- [35] M. Hahn, S. Frühwirth-Schnatter, J. Sass, Markov chain Monte Carlo Methods for parameter estimation in multidimensional continuous time Markov switching models, *Journal of Financial Econometrics*, 8 (2010) 88-121.
- [36] B. Larget, D.L. Simon, Markov chain Monte Carlo algorithms for the Bayesian analysis of phylogenetic trees, *Molecular Biology and Evolution*, 16 (1999) 750-759.
- [37] R. Nielsen, J. Wakeley, Distinguishing migration from isolation: A Markov chain Monte Carlo approach, *Genetics*, 158 (2001) 885-896.
- [38] G.E. Evans, B. Jones, The application of Monte Carlo simulation in finance, economics and operations management, 2009, pp. 379-383.
- [39] O.M. Belotserkovskii, Y.I. Khlopkov, Monte Carlo methods in applied mathematics and computational aerodynamics, *Computational Mathematics and Mathematical Physics*, 46 (2006) 1418-1441.
- [40] B. Lautrup, MONTE CARLO METHODS IN THEORETICAL HIGH-ENERGY PHYSICS, *Communications of the ACM*, 28 (1985) 358-362.

- [41] Y.J. Yang, S.Q. Li, W. Zhang, Monte Carlo simulation for bits error rate of mobile telecommunication systems, Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition), 37 (2007) 18-21.
- [42] W. Leckie, M. Greenspan, Monte-Carlo methods in pool strategy game trees, 5th Computers and Games Conference, 2007, pp. 244-255.
- [43] J.H. Ferziger, M. Peric, Computational Methods for Fluid Dynamics, Springer, Berlin, 1996.
- [44] S. Chapman, T.G. Cowling, The Mathematical Theory of Non-uniform Gases, Cambridge university press, Cambridge, 1970.
- [45] G.A. Bird, Approach to translational equilibrium in a rigid sphere gas, Physics of Fluids, 6 (1963) 1518-1519.
- [46] G.A. Bird, Molecular Gas Dynamics, Clarendon, Oxford, 1976.
- [47] G.A. Bird, Molecular Gas Dynamics and The Direct Simulation of Gas Flows, Clarendon, Oxford, 1994.
- [48] J.P. Blaizot, E. Iancu, A Boltzmann equation for the QCD plasma, Nuclear Physics B, 557 (1999) 183-236.
- [49] S. Fukui, R. Kaneko, Anlysis of ultra-thin gas film lubrication based on linearized Boltzmann equation: first report-derivation of a generalized lubrication equation including thermal creep flow, Journal of Tribology, 110 (1988) 253-262.
- [50] A.J.C. Ladd, Numerical simulations of particulate suspensions via a discretized Boltzmann equation. Part 1. Theoretical foundation, Journal of Fluid Mechanics, 271 (1994) 285-309.

- [51] L. Mieussens, Discrete-Velocity Models and Numerical Schemes for the Boltzmann-BGK Equation in Plane and Axisymmetric Geometries, *Journal of Computational Physics*, 162 (2000) 429-466.
- [52] T. Platkowski, R. Illner, Discrete velocity models of the Boltzmann equation: A survey on the mathematical aspects of the theory, *SIAM Review*, 30 (1988) 213-255.
- [53] L. Marino, The velocity distribution function in direct simulation Monte Carlo method with an application to extended thermodynamics, *Physics of Fluids*, 15 (2003) 662-667.
- [54] G.A. Bird, Recent advances and current challenges for DSMC, *Computers and Mathematics with Applications*, 35 (1998) 1-14.
- [55] A.L. Garcia, F.J. Alexander, B.J. Alder, A particle method with adjustable transport properties - The generalized consistent boltzmann algorithm, *Journal of Statistical Physics*, 89 (1997) 403-409.
- [56] F.J. Alexander, A.L. Garcia, B.J. Alder, A consistent Boltzmann algorithm, *Physical Review Letters*, 74 (1995) 5212-5215.
- [57] J. Schneider, K. Maier, Point defects in silicon carbide, *Physica B: Physics of Condensed Matter*, 185 (1993) 199-206.
- [58] T.Y. Tan, U. Gösele, Point defects, diffusion processes, and swirl defect formation in silicon, *Applied Physics A Solids and Surfaces*, 37 (1985) 1-17.
- [59] V.V. Voronkova, R. Falsterb, Intrinsic point defects and impurities in silicon crystal growth, *Journal of the Electrochemical Society*, 149 (2002) G167-G174.
- [60] K.A. Jackson, *Kinetic Processes: Crystal Growth, Diffusion, and Phase Transitions in Materials*, Wiley-VCH, Weinheim, 2004.

- [61] L.L.Y. Adeline, Studies of monovacancy in hard-sphere solids by molecular simulation: simple cubic and body centered cubic, in, Nanyang Technological University, 2008.
- [62] J.F. Shackelford, Introduction to Materials Science for Engineers, 6th ed., Pearson education international, Davis, 2005.
- [63] W. Soboyejo, Mechanical Properties of Engineered Materials, Marcel dekker Inc., 2003.
- [64] D. William, J. Callister, Material Science and Engineering An Introduction, 6th ed., John Wiley & Sons, 2003.
- [65] T.S.L. Grace, Studies of polyvacancies in model solids by molecular simulation, in, Nanyang Technological University, 2009, Bachelor Thesis.
- [66] D. US, DOE Fundamentals Handbook Material Science: Service NTI, Editor., Washington, 1993.
- [67] A. SA, C. AG, F. BAS, The classification of close-packed clusters of substitutional point defects in crystals, Philo mag. A., 48 (1983) 31-39.
- [68] C.H. Bennett, B.J. Alder, Studies in molecular dynamics. IX. Vacancies in hard sphere crystals, The Journal of Chemical Physics, 54 (1971) 4788-4796.
- [69] S. Ansumali, Minimal kinetic modeling of hydrodynamics, in, ETH Zurich, 2004, PhD thesis.
- [70] D. Genovesi, A Java-Based Direct Monte Carlo Simulation of A Nano-Scale Pulse Detonation Engine, Pennsylvania State University, 2002, Master Thesis.
- [71] S. Pronk, D. Frenkel, Point defects in hard-sphere crystals, Journal of Physical Chemistry B, 105 (2001) 6722-6727.

- [72] R.J. Speedy, Pressure and entropy of hard-sphere crystals, *Journal of Physics Condensed Matter*, 10 (1998) 4387-4391.
- [73] D. Frenkel, A.J.C. Ladd, New Monte Carlo method to compute the free energy of arbitrary solids. Application to the fcc and hcp phases of hard spheres, *The Journal of Chemical Physics*, 81 (1984) 3188-3193.
- [74] S.K. Kwak, Y. Cahyana, J.K. Singh, Characterization of mono- and divacancy in fcc and hcp hard-sphere crystals, *Journal of Chemical Physics*, 128 (2008).
- [75] S.K. Kwak, D.A. Kofke, Elastic constants and the effect of strain on monovacancy concentration in fcc. hard-sphere crystals, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 70 (2004) 214113/214111-214113/214116.
- [76] C. Liu, P.C. Su, S.K. Kwak, Generalized equilibrium concentration of polyvacancy: Case study for trivacancy in hard-sphere crystals, *Molecular Physics*, 109 (2011) 2461-2470.
- [77] M.L. Swanson, Stage I recovery of deformed metals. I, *Phys. Status Solidi A*, 3 (1970) 287-298.
- [78] H.J. Queisser, E.E. Haller, Defects in semiconductors: some fatal, some vital, *Science* (Washington, D. C.), 281 (1998) 945-950.
- [79] R.L. Rowley, *Statistical Mechanics for Thermophysical Property Calculations*, Prentice Hall, Englewood Cliffs, 1994.
- [80] W.H. Lee, H.H. Sung, K.J. Syu, J.Y. Chen, Vacancy effects on T_c in superconducting $\text{LaPt}_{1-x}\text{Si}$ ($0 \leq x \leq 0.2$), *Phys. C (Amsterdam, Neth.)*, 470 S774-S775.

- [81] D.H. Galvan, A. Posada Amarillas, R. Nunez-Gonzalez, S. Mejia, M. Jose-Yacaman, Study of Vacancies and Pd Atom Decoration on the Electronic Properties of Bilayer Graphene, *J. Supercond. Novel Magn.*, 23 1543-1550.
- [82] F. Gao, J. Hu, J. Wang, C. Yang, H. Qin, M. Jiang, Ferromagnetism driven by vacancies and C/N substitution at SrO (100) surface, *Solid State Sci.*, 13 126-130.
- [83] J.-x. Ding, J.-q. Zhang, S.-l. Wang, X.-m. Mu, X.-g. Xu, Q.-t. Gu, Y. Sun, W.-j. Liu, G.-x. Liu, Effect of typical anionic impurity on electrical conductivity of KDP crystal, *Rengong Jingti Xuebao*, 39 (2010) 1089-1093.
- [84] S.K. Kwak, D.A. Kofke, Effect of monovacancies on the relative stability of fcc. and hcp. hard-sphere crystals, *J. Chem. Phys.*, 122 (2005) 176101/176101-176101/176102.
- [85] P.H. Bisio, A.M. Monti, Small vacancy clusters in hcp. metals, *Phys. Status Solidi B*, 135 (1986) 545-553.
- [86] B. Widom, Some topics in the theory of fluids, *J. Chem. Phys.*, 39 (1963) 2808-2812.
- [87] M. Syha, D. Weygand, A generalized vertex dynamics model for grain growth in three dimensions, *Modell. Simul. Mater. Sci. Eng.*, 18 015010/015011-015010/015019.
- [88] Y.S. Lee, S.U. Lee, C.J. Van Tyne, B.D. Joo, Y.H. Moon, Internal void closure during the forging of large cast ingots using a simulation approach, *J. Mater. Process. Tech.*, 211 1136-1145.
- [89] A. Simar, H.-J.L. Voigt, B.D. Wirth, Molecular dynamics simulations of dislocation interaction with voids in nickel, *Comput. Mater. Sci.*, 50 1811-1817.
- [90] T. Tang, S. Kim, M.F. Horstemeyer, Molecular dynamics simulations of void growth and coalescence in single crystal magnesium, *Acta Mater.*, 58 4742-4759.

- [91] S.Y. Hu, C.H. Henager, Phase-field simulation of void migration in a temperature gradient, *Acta Mater.*, 58 3230-3237.
- [92] R.L. Liboff, *Kinetic Theory: Classical, Quantum, and Relativistic Descriptions*, 3rd ed., Springer, New York, 2003.
- [93] W.J. Thistleton, J.A. Marsh, K. Nelson, C. Tsallis, Generalized box-Müller method for generating q-Gaussian random deviates, *IEEE Transactions on Information Theory*, 53 (2007) 4805-4810.
- [94] H. Tanizaki, A simple gamma random number generator for arbitrary shape parameters, *Economics Bulletin*, 3 (2008).
- [95] F.J. Alexander, A.L. Garcia, the direct simulation Monte Carlo method, *Computer Simulations*, (1997) 588-593.
- [96] A. Shalchi, H. Döring, Velocity correlation functions of charged test particles, *Journal of Physics G: Nuclear and Particle Physics*, 34 (2007) 859-870.
- [97] F.J. Alexander, A.L. Garcia, B.J. Alder, Cell size dependence of transport coefficients in stochastic particle algorithms, *Physics of Fluids*, 10 (1998) 1540-1542.
- [98] B.J. Alder, D.M. Gass, T.E. Wainwright, Studies in molecular dynamics. VIII. The transport coefficients for a hard-sphere fluid, *The Journal of Chemical Physics*, 53 (1970) 3813-3826.
- [99] T. Wainwright, Calculation of hard-sphere viscosity by means of correlation functions, *The Journal of Chemical Physics*, 40 (1964) 2932-2937.
- [100] D.J.E.a.G.P. Morriss, *Statistical Mechanics of Non-equilibrium Liquids*, Australian national university press, Australia, 2007.

- [101] J.J. Erpenbeck, W.W. Wood, Self-diffusion coefficient for the hard-sphere fluid, *Physical Review A*, 43 (1991) 4254-4261.
- [102] K.Y. Sanbonmatsu, C.S. Tung, High performance computing in biology: Multimillion atom simulations of nanoscale systems, *Journal of Structural Biology*, 157 (2007) 470-480.
- [103] F. Cappello, S. Djilali, G. Fedak, T. Herault, F. Magniette, V. Néri, O. Lodygensky, Computing on large-scale distributed systems: XtremWeb architecture, programming models, security, tests and convergence with grid, *Future Generation Computer Systems*, 21 (2005) 417-437.
- [104] Z.-H. Li, L. Bi, H.-X. Zhang, L. Li, Gas-kinetic numerical study of complex flow problems covering various flow regimes, *Computers & Mathematics with Applications*, 61 (2011) 3653-3667.
- [105] R.G. Winkler, M. Ripoll, K. Mussawisade, G. Gompper, Simulation of complex fluids by multi-particle-collision dynamics, *Computer Physics Communications*, 169 (2005) 326-330.
- [106] Z. Li, H. Zhang, Study on the unified algorithm for three-dimensional complex problems covering various flow regimes using boltzmann model equation, *Science in China, Series G: Physics, Mechanics and Astronomy*, 52 (2009) 124-138.
- [107] S. Chen, G.D. Doolen, Lattice Boltzmann method for fluid flows, in, 1998, pp. 329-364.
- [108] T. Soddemann, B. Dünweg, K. Kremer, Dissipative particle dynamics: A useful thermostat for equilibrium and nonequilibrium molecular dynamics simulations, *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 68 (2003) 467021-467028.

- [109] T. Ihle, D.M. Kroll, Stochastic rotation dynamics: A Galilean-invariant mesoscopic model for fluid flow, *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 63 (2001) 0202011-0202014.
- [110] E. Barkai, Fractional Fokker-Planck equation, solution, and application, *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 63 (2001) 461181-461187.
- [111] M.W. Lee, J.J. Park, D.Y. Kim, S.S. Yoon, H.Y. Kim, S.C. James, S. Chandra, T. Coyle, Numerical studies on the effects of stagnation pressure and temperature on supersonic flow characteristics in cold spray applications, *Journal of Thermal Spray Technology*, 20 (2011) 1085-1097.
- [112] R.B. Bond, M.A. Gallis, J.R. Torczynski, DSMC predictions of non-equilibrium reaction rates, 2011, pp. 1203-1208.
- [113] H. Van Beijeren, Equilibrium distribution of hard-sphere systems and revised enskog theory, *Physical Review Letters*, 51 (1983) 1503-1505.
- [114] M. Gad-El-Hak, *Liquids: The holy grail of microfluidics modeling*, 2004, pp. 332-337.
- [115] J. Zhong, D.A. Levin, Development of a kinetic nucleation model for a free-expanding argon condensation flow, *AIAA Journal*, 45 (2007) 902-911.

Publication List

1 C. Liu, P.C. Su, S.K. Kwak, Generalized equilibrium concentration of polyvacancy: Case study for trivacancy in hard-sphere crystals, Mol. Phys., 109 (2011) 2461-2470.

2 C. Liu, S. Ansumali and S. K. Kwak^{*}, A modified DSMC algorithm and application, 14th Asia Pacific Confederation of Chemical Engineering (2012)

3 C. Liu, S. K. Kwak and Lim Kok Hwa^{*}, Generalized polyvacancy equilibrium concentration, the international symposium on exploring the frontiers of chemical and biochemical engineering (2012)

4 C. Liu, S. Ansumali and S. K. Kwak^{*}, Minimal molecular dynamics for dense flows of hard sphere fluid, To appear in Phys. Rev. Lett. (2012)

5 C. Liu, S. Ansumali, S. K. Kwak and Lim Kok Hwa^{*}, A hybrid MD-DSMC simulation methodology for multiphase system. (In preparation)

6 C. Liu, S. Ansumali, S. K. Kwak and Lim Kok Hwa^{*}, A novel DSMC algorithm for shock wave. (In preparation)