UNIVERSITY OF CALIFORNIA

Los Angeles

Multiscale Modeling, Simulation and Control of Protein Crystallization Processes

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Chemical Engineering

by

Michael Nayhouse

2015

ABSTRACT OF THE DISSERTATION

Multiscale Modeling, Simulation and Control of Protein Crystallization Processes

by

Michael Nayhouse

Doctor of Philosophy in Chemical Engineering

University of California, Los Angeles, 2015

Professor Panagiotis D. Christofides, Chair

Protein crystallization is a central activity in the pharmaceutical industry which is currently estimated to be over a $1 trillion per year industry. Despite extensive experimental and theoretical work on understanding protein structure and function, there is a lack of a systematic framework that relies on fundamental understanding of the nucleation and growth mechanisms of protein crystals at the microscopic level and utilizes such information to model and operate protein batch crystallization processes at the macroscopic level. Motivated by these considerations, this dissertation is focused on developing a hierarchical and computationally tractable approach to: (a) elucidate the equilibrium fluid-fluid and fluid-solid phase diagrams of globular proteins via coarse-graining techniques, equilibrium Monte Carlo (MC) simulations, and finite-size scaling theory, (b) model crystal growth and morphology via kinetic Monte Carlo (kMC) simulations in order to deduce microscopically

consistent rate laws, and (c) use these microscopic rate laws on the macroscale in order to model and control batch crystallization processes.

The dissertation of Michael Nayhouse is approved.

Paulo Tabuada

Dante Simonetti

James F. Davis

Panagiotis D. Christofides, Committee Chair

University of California, Los Angeles

2015

Contents

List of Figures

ACKNOWLEDGEMENTS

I would first like to express my gratitude to my two co-advisors, Professor Panagiotis D. Christofides and Professor Gerassimos Orkoulas, for their nonstop support throughout my doctoral work. I feel very fortunate I contacted Professor Christofides after being accepted at UCLA. He was very direct which fit well with my personality and I could see the great opportunity I would have if I were to join his laboratory. After my initial contact with Professor Christofides, he told me my background would fit well with a joint research project between him and Professor Orkoulas. I ended up meeting Professor Orkoulas and found an instant connection with him and value all the work we did together. Overall, I could not have asked for two better advisors during my graduate studies.

I would like to thank my parents and sister for their support and encouragement throughout all my life. Their upbringing started me on a path to success that I am extremely grateful to have had. Also, I would like to thank Matt Hartman for always being there for me since we became friends at the age of six and staying in touch over the years. Every time we catch up, it is like we never missed a beat. Moreover, I would like to thank Bill Schleicher for being another very close friend of mine. I am fortunate to have met him and played sports with him throughout my life. Additionally, I would like to thank Huy Nguyen, Henry Leung, and Kevin Ngo for being great friend who I grew closer to during my undergraduate studies and for staying in touch all these years. Furthermore, I would like to express my gratitude to Professor Anna Balazs and Professor Steven Levitan for their continued support and guidance throughout my academic career.

I would like to thank Professor Minhee Yun for the experience working in his laboratory and sending me abroad on an invaluable international research experience during my undergraduate studies. Also, I would like to thank Professor Albert To for his guidance during the end of my undergraduate career, helping me to decide the right path for my graduate studies and continued support over the years.

I would like to thank Joseph Sangil Kwon for being an excellent colleague who worked closely with me on many of our group projects. Our teamwork combined our strengths, thereby accelerating our research progress. Furthermore, I would like to thank all of the undergraduate student researchers I have had the pleasure of advising, along with all the hard work they helped me with over the years. Namely: Ankur Amlani, Vincent Heng, Marquis Crose, and Anh Tran. Also, a special thanks in particular to both Marquis Crose and Anh Tran for their comments in proofreading my dissertation.

I would like to thank all of my lab mates from both my graduate and undergraduate careers. In particular: my first post-doctoral advisor Dr. O. Berk Usta who started me on bash scripting; Dave Perello, Innam Lee, Yushi Hu, and Dave Sanchez who were all great graduate student advisors to me and showed me what to expect as a graduate student; Dr. Meenakshi Dutt and Dr. Olga Kuksenok who reintroduced me to modeling and simulation research before the start of my graduate studies; Sasha Kuksenok who was always there to help when any technical issues came up; Matt Ellis, Larry Gao, Liangfeng Lao, Alex Bartman, Mohsen Heidarinejad, Pek Leosirikul, Xianzhong Chen, David Chilin, and Xinyu Zhang who were great lab mates in addition to being close friends throughout my time at

UCLA.

I was also fortunate to meet and get to know a number of people outside of my research group during my time at UCLA which made the experience much more enjoyable. I would like to thank Daniel Sloan, Phil Jones, and Ethan Schreiber who I had the pleasure of being good friend with, in addition to playing sports with them. Also, I would like to thank Paul (Po-Heng) Lin, Jeni Takasumi, Allison Yorita, Calvin Pham, Daniel Noon, Diana Chien, and Alex Jury for being excellent friends since coming to Los Angeles.

I would like to thank Doc Ellis who was my first programming teacher. His methods were much different than any other teacher I had during high school, which focused on continually writing code in order to gain tools which could then be used together in order to solve an endless number of problems. Without him, I might not have realized my talent in programming which eventually led to my desire to pursue a Ph.D. in modeling, simulation, and control.

I must also thank the UCLA CBE administrative staff, especially John Berger, for their assistance throughout my graduate studies.

Finally, I would like to thank Professor James F. Davis, Professor Dante Simonetti, and Professor Paulo Tabuada for agreeing to serve on my doctoral committee.

Financial support from the National Science Foundation (NSF) CBET-0967291 is gratefully acknowledged and my work could not have been done without it. Additionally, the NSF Graduate Research Fellowship, DGE-1144087, was also a great help in allowing me freedom in my graduate studies. I would also like to thank the Extreme Science and En-

gineering Discovery Environment (XSEDE), which is supported by NSF grant number TG-CCR120003, for allowing me to use computational time on state of the art computing clusters.

Chapter 2 is a version which combines: M. Nayhouse, A. M. Amlani, and G. Orkoulas. A Monte Carlo study of the freezing transition of hard spheres. *J. Phys.: Condens. Matter*, 23:325106, 2011; M. Nayhouse, A. M. Amlani, and G. Orkoulas. Precise simulation of the freezing transition of supercritical Lennard-Jones. *J. Chem. Phys.*, 135:154103, 2011; M. Nayhouse, A. M. Amlani, V. R. Heng, and G. Orkoulas. Simulation of fluid-solid coexistence via thermodynamic integration using a modified cell model. *J. Phys.: Condens. Matter*, 24:155101, 2012; M. Nayhouse, V. R. Heng, A. M. Amlani, and G. Orkoulas. Precise simulation of subcritical freezing using constrained cell models. *Journal of Physics A: Mathematical and Theoretical*, 45:155002, 2012.

Chapter 3 is a version of: M. Nayhouse, J. S. Kwon, and G. Orkoulas. Phase transitions, criticality and three-phase coexistence in constrained cell models. *J. Chem. Phys.*, 136:201101, 2012.

Chapter 4 is a version of: V. R. Heng, M. Nayhouse, M. Crose, A. Tran, and G. Orkoulas. Direct determination of triple-point coexistence through cell model simulation. *J. Chem. Phys.*, 137:141101, 2012.

Chapter 5 is a version of: M. Nayhouse, V. R. Heng, A. M. Amlani, and G. Orkoulas. Simulation of phase boundaries using constrained cell models. *J. Phys.: Condens. Matter*, 24:375105, 2012.

Chapter 6 is a version which combines: M. Nayhouse, J. S. Kwon, P. D. Christofides, and G. Orkoulas. Crystal shape modeling and control in protein crystal growth. *Chemical Engineering Science*, 87:216–223, 2013; J. S. Kwon, M. Nayhouse, P. D. Christofides, and G. Orkoulas. Modeling and control of protein crystal shape and size in batch crystallization. *AIChE J.*, 59:2317–2327, 2013.

Chapter 7 is a version of: M. Nayhouse, A. Tran, J. S. Kwon, M. Crose, G. Orkoulas, and P. D. Christofides. Modeling and Control of Ibuprofen Crystal Growth and Size Distribution. *Chemical Engineering Science*, submitted.

VITA

2006–2010	Bachelor of Science, Computer Engineering
	Department of Electrical and Computer Engineering
	University of Pittsburgh
2010	George Washington Prize
	University of Pittsburgh
2010–2014	Graduate Student Researcher / Teaching Assistant
	Department of Chemical and Biomolecular Engineering
	University of California, Los Angeles
2012	Teaching Assistant of the Year
	Department of Chemical and Biomolecular Engineering
	University of California, Los Angeles
2012–2015	National Science Foundation Graduate Research Fellowship

PUBLICATIONS

1. M. Nayhouse, A. Tran, J. S. Kwon, M. Crose, G. Orkoulas, and P. D. Christofides, "Modeling and Control of Ibuprofen Crystal Growth and Size Distribution," *Chemical Engineering Science*, 2015, submitted.

2. M. Nayhouse, J. S. Kwon, V. R. Heng, A. M. Amlani, and G. Orkoulas, "Freezing Transition Studies Through Constrained Cell Model Simulation," *International Journal of Thermophysics*, 2014, **35**, 1661-1676.

3. M. Nayhouse, J. S. Kwon, P. D. Christofides, and G. Orkoulas, "Crystal Shape Modeling and Control in Protein Crystal Growth," *Chemical Engineering Science*, 2013, **87**, 216-223.

4. M. Nayhouse, V. R. Heng, A. M. Amlani, and G. Orkoulas, "Simulation of Phase Boundaries Using Constrained Cell Models," *J. Phys.: Condens. Matter*, 2012, **24**, 375105.

5. M. Nayhouse, J. S. Kwon, and G. Orkoulas, "Phase Transitions, Criticality and Three-Phase Coexistence in Constrained Cell Models," *J. Chem. Phys.*, 2012, **136**, 201101.

6. M. Nayhouse, V. R. Heng, A. M. Amlani, and G. Orkoulas, "Precise Simulation of Subcritical Freezing using Constrained Cell Models," *Journal of Physics A: Mathematical and Theoretical*, 2012, **45**, 155002.

7. M. Nayhouse, A. M. Amlani, V. R. Heng, and G. Orkoulas, "Simulation of Fluid-Solid Coexistence via Thermodynamic Integration Using a Modified Cell Model," *J. Phys.: Condens. Matter*, 2012, **24**, 155101.

8. M. Nayhouse, A. M. Amlani, and G. Orkoulas, "Precise simulation of the fluid-solid transition of supercritical Lennard-Jones," *J. Chem. Phys.*, 2011, **135**, 154103.

9. M. Nayhouse, A. M. Amlani, and G. Orkoulas, "A Monte Carlo study of the freezing transition of hard spheres," *J. Phys.: Condens. Matter*, 2011, **23** (32), 325106.

Chapter 1

Introduction

The pharmaceutical industry is a $1 trillion industry and has been a major contributor to both scientific advancement and economic growth. Proteins themselves play a key role as therapeutics in a number of diseases and protein crystallization is a central activity in the pharmaceutical industry. Specifically, the production of highly-ordered, high-quality protein crystals through batch crystallization processes is vital in devising proteins for therapeutic purposes. In order to gain a strong understanding of the biochemical role of proteins, the three-dimensional structure must be determined. Two experimental techniques used to determine protein structure are nuclear magnetic resonance and X-ray crystallography. For cases of small molar mass (less than 30,000), nuclear magnetic resonance can be used. On the other hand X-ray crystallography can be used for cases of large molar mass as long as the protein crystals are of desired shape and high quality.

Despite extensive experimental and theoretical work on understanding protein structure

and function, there is lack of systematic framework that relies on a fundamental understanding of the nucleation and growth mechanisms of protein crystals at the microscopic level, and utilizes such information to model and operate protein batch crystallization processes at the macroscopic level. Motivated by these considerations, this dissertation does the following: First (1), equilibrium phase diagrams are determined via coarse-graining techniques, Monte Carlo simulations, and finite-size scaling theory. Next (2), crystal growth and morphology is investigated through kinetic Monte Carlo (kMC) simulations which will deduce microscopically consistent rate laws. After that (3), these rate laws are used at the macroscopic level in conjunction with available nucleation rate data to model a batch crystallizer. The results of the batch crystallizer will be connected back to the microscopic kinetic Monte Carlo simulations via model predictive control (MPC), thus optimizing the overall process. These steps are explained in more detail below.

1.1 Equilibrium Monte Carlo Simulation

Based on the McMillan-Mayer description of solutions, coarse graining techniques will be applied to replace the multi-component globular protein solution with a one-component system of particles (proteins) with effective interactions. Specifically, the solvent and salt degrees of freedom will be eliminated and a family of effective pair potentials will be obtained. These effective protein-protein pair potentials will be decomposed and converted into a parametric form that consists of a superposition of simpler short- as well as long-range parts. The phase diagrams will be determined by Monte Carlo simulations of a

reformulated version of the generalized cell model of the Hoover-Ree [55] and the results will be analyzed through finite-size scaling techniques [99]. Specifically, the freezing transition is investigated for systems of hard-spheres and Lennard-Jones particles via Monte Carlo simulations using cell models, thermodynamic integration, and finite size scaling theory. After that, these simulations are then used to elucidate phase boundaries through cell model simulation, including finding triple point coexistence.

1.2 Crystal Growth Rate Models

Since crystal growth is a non-equilibrium process, it will be simulated using kinetic Monte Carlo methods [48]. As is common practice in simulations of crystal growth, the solid-on-solid (SOS) model will be adopted [27]. In the solid-on-solid approximation, particles are deposited on the growing crystal without voids or overhangs and the resulting crystal is highly compacted. The implementation of the kinetic Monte Carlo methodology requires knowledge of the binding energies, the impingement rate, and the crystallization driving force. In previous simulations [27] as well as simulations in this dissertation, a range of values is assigned to the previous parameters until satisfactory agreement with experiments is obtained for the system under consideration. The calculated crystal growth rates will be tested against available experimental data and will be used to deduce molecularly correct microscopic/mesoscopic rate laws.

1.3 Batch Crystallization and MPC of Crystal Shape and Size

The crystal growth rate model is then used along with nucleation rate data to model batch crystallizers through population balance models and conservation equations. These models are used in conjunction with model predictive control methodologies that account for model uncertainty to achieve desired crystal shape and size distributions at the end of the batch crystallization process. The population balance model contains a Gaussian white noise term that describes the stochastic nature of the process owing to the discrete nature of the crystal size distribution and accounts for model uncertainty. The design of the feedback control system to regulate the crystal size distribution employs a novel stochastic model predictive control (MPC) approach [116]. Specifically, due to the infinite dimensional nature of the batch crystallization model, the method of weighted residuals with global basis functions will be first used to derive reduced-order ordinary differential equation models that will be used for controller design.

1.4 Outline of the Dissertation

Chapter 2 focuses on the precise simulation of freezing transitions. Thermodynamic integration techniques are the simplest methods to study fluid-solid coexistence. These methods are based on the calculation of the free energy of the fluid and of the solid phase,

starting from a state of known free energy which is usually an ideal-gas state. Despite their simplicity, the main drawback of thermodynamic integration techniques is the large number of states that must be simulated. In this chapter, the fluid-solid transition is analyzed for a system of hard spheres, in addition to Lennard-Jones particles on both supercritical and subcritical isotherms. The analysis is implemented via a simulation method which is based on a modification of the constrained cell model of Hoover and Ree, which greatly reduces the number of simulated states needed in thermodynamic integration. In the context of hard-sphere freezing, Hoover and Ree simulated the solid phase using a constrained cell model in which each particle is confined within its own Wigner-Seitz cell. Hoover and Ree also proposed a modified cell model by considering the effect of an external field of variable strength. High-field values favor configurations with a single particle per Wigner-Seitz cell and thus stabilize the solid phase. As in the case of hard spheres, constant-pressure simulations of the fully occupied constrained cell model of a system of Lennard-Jones particles at a reduced temperature of two indicate a point of mechanical instability at a density which is approximately 70% of the density at close packing. Furthermore, constant-pressure simulations of the modified cell model indicate that as the strength of the field is reduced, the transition from the solid to the fluid is continuous below the mechanical instability point and discontinuous above. The fluid-solid transition of these systems is obtained by analyzing the field-induced fluid-solid transition of the modified cell model in the high-pressure, zero-field limit. The simulations are implemented under constant pressure using tempering and histogram reweighting techniques. The coexistence pressure and densities

are determined through finite-size scaling techniques for first-order phase transitions which are based on analyzing the size-dependent behavior of susceptibilities and dimensionless moment ratios of the order parameter.

Chapter 3 focuses on phase transitions, criticality, and three phase coexistence using constrained cell models. In simulation studies of fluid-solid transitions, the solid phase is usually modeled as a constrained system in which each particle is confined to move in a single Wigner-Seitz cell. The constrained cell model has been used in the determination of fluid-solid coexistence via thermodynamic integration and other techniques. In this chapter, the phase diagram of such a constrained system of Lennard-Jones particles is determined from constant-pressure simulations. The pressure-density isotherms exhibit inflection points which are interpreted as the mechanical stability limit of the solid phase. The phase diagram of the constrained system contains a critical and a triple point. The temperature and pressure at the critical and the triple point are both higher than those of the unconstrained system due to the reduction in the entropy caused by the single occupancy constraint.

Chapter 4 focuses on direct determination of triple-point coexistence through cell model simulation. In this chapter, the phase diagram of the generalized model is investigated through multicanonical simulations at constant pressure and histogram reweighting techniques for a system of 256 Lennard-Jones particles. The simulation data is used to obtain an estimate of the triple point of the Lennard-Jones system.

Chapter 5 focuses on simulations of phase boundaries using constrained cell models.

More specifically, this chapter focuses on the determination of the phase diagram of a system of particles that interact through a pair potential, $\phi(r)$, which is of the form $\phi(r) = 4\varepsilon\left[(\sigma/r)^{2n} - (\sigma/r)^{n}\right]$ with $n = 12$. The vapor-liquid phase diagram of this model is established from constant-pressure simulations and flat-histogram techniques. The properties of the solid phase are obtained from constant-pressure simulations using constrained cell models. Fluid-solid coexistence at a reduced temperature of 2 is established from constant-pressure simulations of the generalized cell model. The previous fluid-solid coexistence point is used as a reference point in the determination of the fluid-solid phase boundary through a thermodynamic integration type of technique based on histogram reweighting. Since the attractive interaction is of short range, the vapor-liquid transition is metastable against crystallization. In this chapter, the phase diagram of the corresponding constrained cell model is also determined. The latter is found to contain a stable vapor-liquid critical point and a triple point.

Chapter 6 focuses on protein crystal shape modeling and control in batch crystal growth processes. Proteins play a key role as therapeutics in a number of diseases and protein crystallization is a central activity in the pharmaceutical industry. Protein crystals, usually produced through a batch crystallization process, are desired to be of high quality, of desired shape, and within a narrow size and shape distribution range. Motivated by the above considerations, this chapter focuses on the modeling and control of protein crystal shape. The model protein used for this chapter is tetragonal hen egg white lysozyme. The growth of an individual lysozyme crystal is modeled via kinetic Monte Carlo (kMC) sim-

ulations comprising adsorption, desorption, and migration events on the (110) and (101) faces, which are assumed to be independent. The expressions for the rate equations for each event type are similar to those of Durbin and Feher [27]. Extensive testing of the system parameters indicates crossover behavior between the growth rates of the two faces [i.e., (110) and (101)], a fact that has also been observed experimentally. A nonlinear algebraic equation that relates the steady-state growth rate ratios between the (110) and (101) faces, the temperature and concentration, is derived from the kMC simulation data. This nonlinear equation is then utilized by a model predictive controller which regulates the protein crystal to desired shapes subject to manipulated input constraints. The proposed method is shown to successfully regulate protein crystal shape, ranging from equidimensional to more elongated type of structures, in the presence of arbitrary variations of the protein concentration. Furthermore, nucleation and population balance models will be added in order to account for the macroscopic phenomena of a batch crystallization process.

Chapter 7 focuses on modeling and control of ibuprofen crystal growth and size distribution. More specifically, this chapter focuses on multiscale modeling and control of a seeded batch crystallization process used to produce ibuprofen crystals. For the modeling of the crystal growth process, we consider kinetic Monte Carlo (kMC) simulations comprising of molecule adsorption, desorption, and migration type microscopic surface events, similar to the previous chapter. To account for growth rate variability, a model for growth rate dispersion (GRD) is proposed, based on available experimental data, which will be applied at the individual crystal growth level in the kMC simulations. Lastly, a model

predictive controller (MPC) is developed in order to control the crystal size distribution of ibuprofen in the batch crystallization process and the MPC closed-loop performance is compared against constant temperature control (CTC) and constant supersaturation control (CSC) policies. The proposed MPC is able to deal with the constraints of the control problem, in addition to minimizing the spread of the crystal size distribution in a superior fashion compared to the other control methodologies, which improves the crystal product quality at the end of the batch.

Finally, Chapter 8 summarizes the contributions of this dissertation.

Chapter 2

Precise Simulation of Freezing Transitions

2.1 Introduction

Despite its importance in many systems such as protein solutions and colloidal suspensions [35, 93], precise simulation of freezing transitions still remains a challenging task. Early attempts to determine fluid-solid coexistence comprised a separate calculation of the free energies of the two phases via thermodynamic integration types of techniques [55, 56, 52, 57, 36]. In these simulations, the solid phase was modeled as a constrained cell system or as an Einstein crystal. The constrained cell model method of Hoover and Ree [55, 56], in which each particle is confined to move in its own Wigner-Seitz cell, is used less often since it is believed that the integration path encounters anomalies associated

with phase transitions [36]. Direct simulation techniques comprise constrained fluid λ-integration [51, 28] and phase switch methods [127]. In the constrained fluid λ-integration method [51, 28], the solid phase is transformed to the fluid phase through a series of steps that are associated with: (a) switching on and off particle interactions, and (b) activation and deactivation of Gaussian potential wells centered at the sites of the solid phase under consideration. In the phase switch Monte Carlo method [127], the passage from the fluid to the solid phase is obtained through a series of "smart" moves that facilitates the formation of "gateway" states in each phase from which the transition to the other phase can be achieved. This method has been applied to hard-sphere [127] and Lennard-Jones systems [29, 78].

In their original work associated with the determination of the hard-sphere freezing transition via thermodynamic integration techniques [55, 56], Hoover and Ree modeled the solid phase as a constrained system in which each particle is confined to move in its own Wigner-Seitz cell. Constant-volume simulations of the constrained cell model detected a point of mechanical instability at a density which is approximately 64% of the density at close packing [56]. This mechanical instability point appeared in the form of a cusp or a kink in the pressure–density curve. For densities that are lower than the densities associated with the cusp, the solid phase cannot survive without the presence of the cell walls. It should be noted that one usually extends the simulations from the high-density, nearly-incompressible limit to the low-density, ideal-gas region for which the free energy can be obtained analytically when using thermodynamic integration techniques based on

constrained cell models. In the same work, Hoover and Ree proposed a more general cell model [55] in order to reduce the number of simulations needed in thermodynamic integration by adding a homogenous external field that controls the relative stability of the solid versus the fluid phase. High field values force single occupancy configurations with one particle per Wigner-Seitz cell and thus favor the solid phase. Normal (unconstrained) system behavior is restored in the limit of vanishing field. Hence, the constrained cell model is a special case of the more general or modified cell model. Hoover and Ree thought that the modified cell model could be used to link the fluid with the solid phase on a constant-density (or pressure) path by gradually increasing the strength of the field, thus reducing the number of simulated states in thermodynamic integration techniques. It is important to emphasize that Hoover and Ree did not exclude the possibility of a terminal (i.e., critical/Curie) type of point separating continuous from discontinuous behavior. In their original work associated with hard-sphere freezing, Hoover and Ree investigated the properties of this model at low densities through cluster expansion techniques. This model has received very little attention in simulations of freezing transitions.

2.2 The constrained and modified cell models

Consider the determination of the fluid-solid transition of a model system at temperature T or inverse temperature $\beta = 1/(k_B T)$, where k_B is Boltzmann's constant. For a system of N particles, simulations in the constrained cell model are implemented by dividing the simulation volume, V, into N Wigner-Seitz cells representative of the solid phase under

consideration. For the case for which the solid phase is of the face-centered cubic (fcc) type, these N cells are dodecahedra with rhombic faces and $N = 4n^3$ $(n = 2, 3, \dots)$ for a cubic simulation box with periodic boundary conditions. Each particle is assigned and confined to move in a single cell and particle displacements violating the single occupancy constraint are rejected by the Metropolis acceptance criteria. For the fluid phase, no such single occupancy constraint exists and the particles are free to move within the entire simulation volume V. For N particles at temperature T and pressure p, the isothermal-isobaric partition functions of the fluid and the solid, Δ_f and Δ_s, are given by

$$\Delta_f(N, p, T) = e^{-\widetilde{\mu}_f N} = \sum_i e^{-\beta E_i - \widetilde{p} V_i}, \tag{2.1}$$

$$\Delta_s(N, p, T) = e^{-\widetilde{\mu}_s N} = \sum_i{}^{*} \, e^{-\beta E_i - \widetilde{p} V_i}, \tag{2.2}$$

where index i enumerates the states and E_i and V_i are the energy and volume in state i. In Eqs. (2.1) and (2.2), $\widetilde{p} = \beta p$, $\widetilde{\mu}_f = \beta \mu_f$, and $\widetilde{\mu}_s = \beta \mu_s$, where μ_f and μ_s are the chemical potentials of the fluid and the solid phase at T and p. The star $(*)$ in Eq. (2.2) signifies the presence of the single occupancy constraint. In Monte Carlo simulations for both cases, the elementary updates comprise particle displacements and volume changes that are accepted with standard Metropolis acceptance criteria [37].

Furthermore, in isothermal-isobaric simulations the volume, V, varies and thus the density, $\rho = N/V$, and the volume per particle, $v = 1/\rho = V/N$, fluctuate about mean values, $\langle \rho \rangle$ and $\langle v \rangle$ respectively. Reduced variables are defined as $\langle \rho^* \rangle = \langle \rho \rangle \sigma^3$ (average reduced density), $\langle v^* \rangle = \langle v \rangle / \sigma^3$ (average reduced volume per particle), $T^* = k_B T / \varepsilon$ (reduced temperature), and $p^* = \widetilde{p} \sigma^3 = \beta p \sigma^3 = p \sigma^3 / k_B T$ (reduced pressure). The definition of the

reduced pressure used in this work ($p^* = p\sigma^3/k_BT$) is different from that used in previous work ($p^* = p\sigma^3/\varepsilon$) for similar systems [57, 52, 59]. As stated previously, in order to reduce the number of simulated states in thermodynamic integration techniques, Hoover and Ree considered a generalized (or modified) cell model. Specifically, they introduced a homogenous external field variable that interacts with each of the N Wigner-Seitz cells. The magnitude of the interaction depends on the specific value of the field and on the number of particles in a given cell. High values of the field variable favor single occupancy configurations and thus stabilize the solid phase. Hence, the constrained cell model is a limiting case of the generalized cell model. The unconstrained system, for which there is no restriction as to the occupancy of a given Wigner-Seitz cell, is recovered in the limit of vanishing field. Recently, Nayhouse *et al.* [98, 85, 86, 84, 87, 89, 88, 54] formulated this model in the isothermal-isobaric ensemble and investigated its behavior via Monte Carlo simulations. The results of these simulations are found in this, as well as the next three chapters.

Consider arbitrary assignments of the N particles into the N Wigner-Seitz cells. The variable s_j is used to describe the occupation status (number of particles) of cell j ($j = 1,2,\ldots,N$). Variable s_j only takes two values: -1 if cell j contains one particle and zero otherwise. The number of singly occupied cells of a configuration is given by the variable $|M|$, where $M = \sum_j s_j$. Consider the external field variable B, and suppose that the interaction of the field with cell j is Bs_j. If $b = \beta B$, the isothermal-isobaric partition function, Δ_m,

for N particles at p and T, is

$$\Delta_m(N,p,T,b) = e^{-\widetilde{\mu}_m N} = \sum_i e^{-\beta E_i - \widetilde{p} V_i - b M_i}. \tag{2.3}$$

In Eq. (2.3), $|M_i|$ is the number of singly occupied cells in state i and $\widetilde{\mu}_m = \beta\mu_m$, where μ_m is the chemical potential at p, T, and b. In Monte Carlo simulations of a system with partition function given by Eq. (2.3), the elementary steps comprise particle displacements and volume changes that are accepted with standard Metropolis acceptance criteria [37]. For particle displacements, the acceptance probability from state i to state j, α_{ij}, must account for changes in the variable M when the center of a particle moves out of its current cell and enters a neighboring one. In such a case

$$\alpha_{ij} = \min\left[1, \exp\left(-\beta\Delta E - b\Delta M\right)\right], \tag{2.4}$$

where $\Delta E = E_j - E_i$ and $\Delta M = M_j - M_i$.

2.3 Hard spheres

Nayhouse *et al.* [98, 85] investigated the properties of the modified and the constrained cell model via constant-pressure simulations for a system of hard spheres. As anticipated, we found that the constrained cell model (i.e., $b = \infty$) contains a point of mechanical instability (see e.g., Fig. 2.1 for the case of hard spheres). The mechanical instability appeared in the form of an inflection point in the pressure-density isotherms.

The effect of the external field, b, for a system of hard spheres is shown in Fig. 2.2. In that figure, the variation of the external field, b, with the average fraction of singly occupied

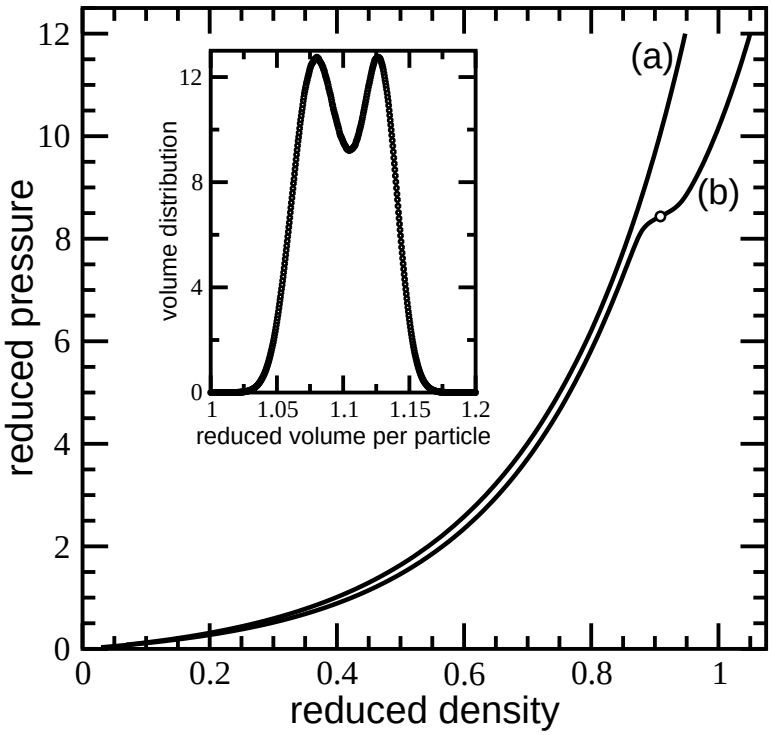


Figure 2.1: Reduced pressure, $\beta p\sigma^3$, vs reduced density, $\rho\sigma^3$, for $N = 256$ hard-spheres. (a): fluid (b): constrained cell model. The inset shows the distribution of the reduced volume for the constrained cell model at $\beta p\sigma^3 = 8.4358$. This point corresponds to the open circle, ($\circ$), on the curve (b) in the main figure.

cells, $-\langle M\rangle/N$, is investigated for $N = 256$ hard spheres for four isobars that correspond to reduced pressures $\beta p\sigma^3 = 7$, 9, 10, and 11 respectively. The notation $\langle x\rangle$ stands for the ensemble average of the fluctuating variable x under constant pressure. The behavior shown in Fig. 2.2 is reminiscent to that of a system undergoing a phase transition. As b is reduced at constant pressure, the passage from the solid to the fluid phase is continuous at low and intermediate pressures, see, e.g., isobar (a) in Fig 2.2. However, for pressures $\beta p\sigma^3 \gtrsim 8.5$ there is a range of values of the external field b, for which the distribution of volume V, and of the number of singly occupied cells $|M|$, both exhibit a two-peak structure separated by a minimum. The inset of Fig. 2.2 shows such a bimodal volume distribution for $\beta p\sigma^3 = 10$ and $b = 0.626$. The particular value of b has been selected through histogram reweighting [31, 32] to yield an average fraction of singly occupied cells of 95%. The low-volume peak, centered at $v^* \cong 1$, corresponds to a solid phase for which the fraction of singly occupied cells is unity. The high-volume peak, centered at $v^* \cong 1.1$, corresponds to a fluid-like phase for which the fraction of singly occupied cells is $\sim 82\%$. Hence, at high pressures ($\beta p\sigma^3 \gtrsim 8.5$), the passage from the solid to the fluid phase as the field decreases occurs via a phase transition.

From the previous considerations it follows that the behavior of the isobars shown in Fig. 2.2 seems to be consistent with the existence of a Curie type of point at $8 \lesssim \beta p\sigma^3 \lesssim 9$ in which the high-pressure, field-induced fluid-solid transition appears to terminate. The field-induced fluid-solid transition for hard spheres has been studied for system sizes $N =$ 108 to 1372 and the resulting coexisting densities are shown in Fig. 2.3 together with the

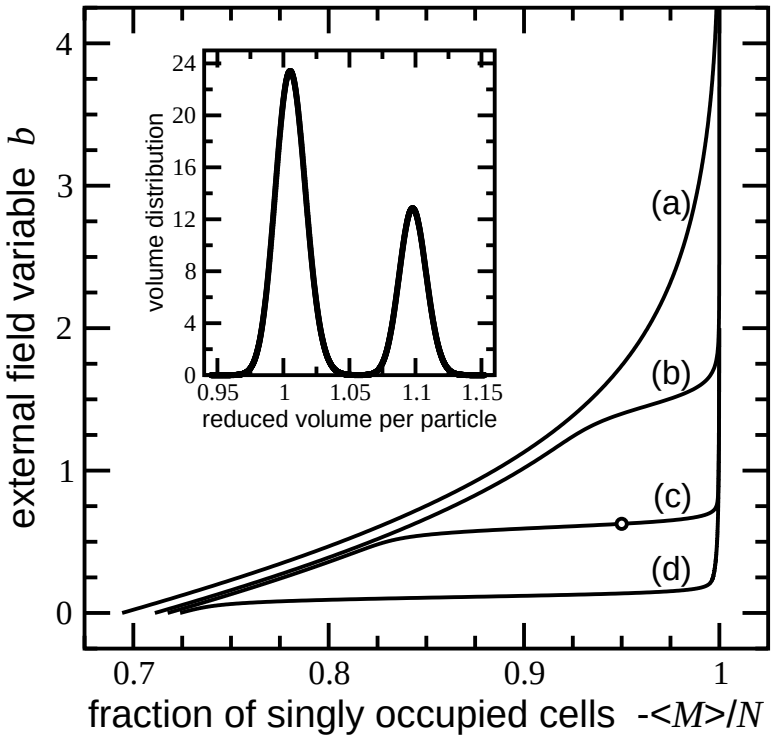


Figure 2.2: External field variable b, vs average fraction of singly occupied cells $-\langle M\rangle/N$, for $N = 256$ hard spheres. The curves correspond to reduced pressures (a) $\beta p\sigma^3 = 7$, (b): $\beta p\sigma^3 = 9$, (c): $\beta p\sigma^3 = 10$, and (d): $\beta p\sigma^3 = 11$. The inset shows the volume distribution at $\beta p\sigma^3 = 10$ and $b = 0.626$. For this state, the average fraction of singly occupied cells is 95%. This state corresponds to the open circle, (○), on the curve (c) of the main figure.

p–ρ curves of the fluid and the solid phase of the hard-sphere system from Fig. 2.1. Phase coexistence for a given b was determined by locating the value of the pressure for which the resulting volume distribution (histogram) comprised two peaks of equal area. This type of calculation is known as the equal weight criterion [12]. The results shown in Fig. 2.3 indicate that this field-induced fluid-solid coexistence lies in between the two p–ρ curves of the fluid and solid phases. In addition, as the pressure increases, the corresponding value of the field, b, decreases. Furthermore, as $b \to 0$, the coexisting fluid and solid densities of the modified cell model approach those of the fluid and the constrained cell model respectively. This fact will be used in the next subsection to determine fluid-solid coexistence of the hard-sphere model.

Referring to Fig. 2.3, the location of the terminal point and the resulting type of criticality could be determined by finite-size scaling techniques for second-order phase transitions [97] for off-lattice, continuum model systems. Such an analysis was not implemented in the present work since the modified cell model is an artificial one and the thermodynamic states associated with intermediate values of b are not real, physical states. The sole purpose of this model is to facilitate transitions from a fluid to a solid phase and thus provide a direct simulation approach to equilibrium freezing transition of the system under study. Fig. 2.3 suggests that the implied terminal point associated with the the field-induced phase transition of the modified cell model lies very close to the mechanical instability point of the constrained cell model which appears as an inflection point in p–ρ curve at $\beta p\sigma^3 \cong 8.5$, see Fig. 2.1. This issue merits separate investigation however. In this context, it is im-

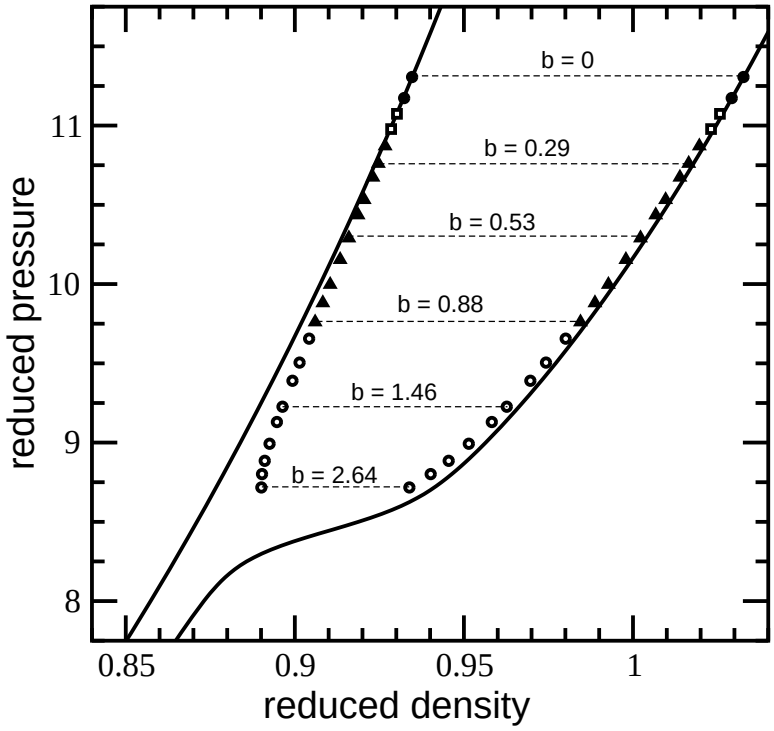


Figure 2.3: Phase diagram of the modified cell model for hard spheres in the pressure-density phase. The solid lines correspond to the fluid and the constrained solid phases [i.e., curves (a) and (b) in Fig. 2.1]. The coexisting solid and fluid phases for finite values of b are shown as points and they correspond to system sizes, N of (○): 1372, (▲): 864, (□): 500, and (●): 256 particles. For reasons of clarity, the results for $N = 108$ have been omitted from the figure. The dashed horizontal lines are the tie-lines and they are drawn for clarity.

portant to emphasize that the underlying mechanism for the field-induced phase transition of the modified cell model might be related to the mechanical stability of the constrained cell model. In other words, a solid that is mechanically stable (i.e., it can exist without the confinement imposed by artificial cell walls) will always be transformed discontinuously (i.e., through a first-order phase transition) to a fluid phase.

2.3.1 Simulation of Freezing Transitions

In the present subsection, the freezing transition of a system of hard spheres is analyzed from simulations of the modified cell model. The simplest way to implement such a task is via thermodynamic integration. As already commented, one calculates the free energy of both the fluid and the solid phase for a series of states starting from a state of known free energy which is usually an ideal-gas state. The main drawback of thermodynamic integration is the large number of states that must be simulated. Since there is no path that can link the two phases, the simulations must be extended towards a region in which the two free energies can be determined independently by some other means, e.g., analytically in the ideal-gas region. However, as Hoover and Ree pointed out [55], the modified cell model can provide a path of constant pressure (or density) that links the fluid and the solid phase, thus reducing the number of states that must be simulated. Hence, calculation of the complete p–ρ isotherms of the two phases is unnecessary. In the present section, the freezing transition of the hard sphere system is studied by analyzing the field-induced phase transition of the modified cell model in the limit of vanishing external field via tempering

[77, 76] and histogram reweighting techniques [31, 32].

According to simulated tempering [77, 76], several replicas of the system are combined to form an expanded ensemble. In the content of this work, replica m corresponds to a system of $N(=4n^3$ with $n = 3$, 4, 5, and 6) particles at temperature T, pressure p_m, and external field b_m. If $\Delta_m = \Delta(N, p_m, T, b_m)$ is the isothermal-isobaric partition function of the m-th replica, the partition function of the expanded ensemble that consists of K replicas is defined as follows:

$$\Phi = \sum_{m=1}^{K} \Delta_m \, e^{\eta_m}, \tag{2.5}$$

where η_m is a weighting factor associated with the m-th replica.

Inspection of the partition function Φ in Eq. (2.5) indicates that three types of elementary updates are possible: particle displacements and volume changes that are accepted with standard Metropolis acceptance criteria [37] and transitions between the K replicas. The switches between different replicas are accepted via a heat-bath type of algorithm [96, 92]. In such a case, the transition probability, P_{mn}, from replica m to replica n is given by

$$P_{mn} = \frac{e^{-\beta P_n V - b_n M + \eta_n}}{\sum_{r=1}^{K} e^{-\beta P_r V - b_r M + \eta_r}}. \tag{2.6}$$

The weighting factor η_m controls the frequency with which replica m occurs in the course of a simulation. From Eq. (2.5), it follows that the probability, π_m, of observing replica m, irrespective of the particular microscopic configuration is

$$\pi_m = \frac{1}{\Phi} \Delta_m \, e^{\eta_m}. \tag{2.7}$$

The weighting factor η_m must be selected such that the probabilities π_m are of the same order of magnitude (i.e., roughly equal). If $G_m = -k_B T \ln \Delta_m$ is the Gibbs free energy at T, P_m, and b_m, the choice $\eta_m = \beta G_m$ ensures that each replica is visited with the same frequency. Since the free energies, G_m, are unknown, an iterative procedure is required during which the weights are adjusted and refined until all the π_m are roughly equal. After a suitable set of η_m has been so obtained, a long simulation is executed to collect good-quality statistics. The main outcome of these simulations are histograms of the fluctuating variables (i.e., V and M) that are manipulated and analyzed through histogram reweighting [31, 32].

In order to study the field-induced phase transition of the modified cell model [cf. Fig. 2.3], a locus of points must be constructed that approximates the phase boundary in the p–b plane and from which a set of K replicas can be selected. A simple choice comprises a line on which the corresponding volume distribution consists of two peaks of equal height. Such a line may be established from simulations of small systems. As data for larger systems become available, this line may be adjusted and refined through histogram reweighting. The choice of this locus is not unique and other possibilities also exist; see, e.g., Ref. [97] and discussion therein.

The previous methodology was used on a system of hard spheres of $N = 108$, 256, 500, 864 particles. A set of $k = 30$–60 replicas (depending on N) was selected for pressures in the range $8.5 \lesssim \beta p \sigma^3 \lesssim 11.6$, cf. Fig. 2.3. Although Fig. 2.3 contains results for $N = 1372$ particles, these simulations were not extended towards $b \to 0$ since such a task

would require a large number of replicas with a substantial increase of the computational time. The weights η_m were estimated from simulations of small systems ($N = 108$–256) and used as initial estimates for larger systems ($N \geq 500$) via rescaling. Once a satisfactory set of weights was determined for a given N, a long simulation was executed to obtain good-quality statistics. Each simulation was executed in sweeps, where one sweep corresponds to $N = 4n^3$ elementary updates: (40% displacements, 40% volume changes, and 20% replica switches). The total length of each simulation in terms of sweeps depends on the number of simulated replicas and the number of entries required to obtain histograms free of spurious structure in order to implement a reliable analysis via histogram reweighting. Typically, a single histogram requires $(5$–$20)\times 10^6$ entries, depending on system size (i.e., N). Errors were estimated by dividing each simulation into 10 blocks. The reported uncertainties correspond to two standard deviations of the respective block averages.

For solid-like configurations associated with low values of the external field (i.e., $b < 0.1$) and low volumes (i.e., $v^* < 1.02$), the particles may drift away from their respective Wigner-Seitz cells. This happens because the low value of the field cannot stabilize the particles in their Wigner-Seitz cells. In previous simulations of solid phases, either through the constrained cell model or the Einstein crystal, the previous problem was alleviated by maintaining the center of mass of the crystal fixed and applying appropriate corrections to recover the properties of the unconstrained system [36, 102, 57]. In the present work, instead of maintaining a fixed center of mass, a different type of constraint is applied in the simulations. Specifically, for low values of the external field ($b < 0.1$) and for low volumes

($v^* < 1.02$), configurations other than those associated with single occupancy are prohibited. Transitions that violate this constraint are automatically rejected by the acceptance criteria and the current state is counted again. For higher values of b ($b \gtrsim 0.1$) and for low volumes, such a constraint is unnecessary since the magnitude of the field is sufficient to prevent drifting.

2.3.2 Data Analysis and Phase Coexistence

As already commented, the freezing transition of the system under consideration can be studied by analyzing the field-induced phase transition of the modified cell model in the limit of vanishing field. The volume distributions at fluid-solid (fcc) coexistence for a system of hard spheres, obtained using the procedure outlined in the previous section, are shown in Fig. 2.4 for $N = 256$, 500, and 864 particles. The distribution for $N = 108$ is not shown for clarity. In the context of this work, phase coexistence is determined by finding (through histogram reweighting) the value of the pressure for which the corresponding volume distribution consists of two peaks of equal area [12]. This procedure yields N-dependent pressures that can be extrapolated towards $N \to \infty$ using finite-size scaling techniques for first-order phase transitions [9, 8, 6, 7]. In the following paragraphs, additional estimators for the pressure at coexistence are developed in order to achieve better precision.

The variation of the reduced volume per particle, v^*, with reduced pressure, $\beta p \sigma^3$, at the transition, is shown in Fig. 2.5. The rounding effects associated with phase transitions in the finite systems are clearly visible in this figure. Whereas in the thermodynamic limit

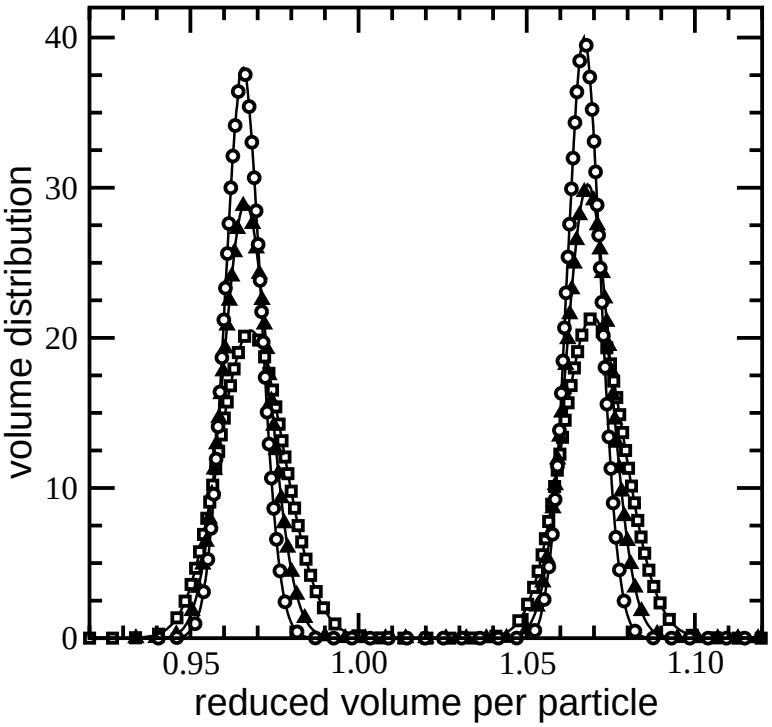


Figure 2.4: Distribution of reduced volume per particle, v^*, for the hard sphere system at fluid-solid (fcc) coexistence ($b = 0$). (□): $N = 256$, $\beta p\sigma^3 = 11.3063 \pm 0.0070$; (▲): $N = 500$, $\beta p\sigma^3 = 11.4075 \pm 0.0078$; (○): $N = 864$, $\beta p\sigma^3 = 11.4630 \pm 0.0085$; The distribution for $N = 108$ (not shown in the figure for clarity) corresponds to $\beta p\sigma^3 = 11.0851 \pm 0.0030$.

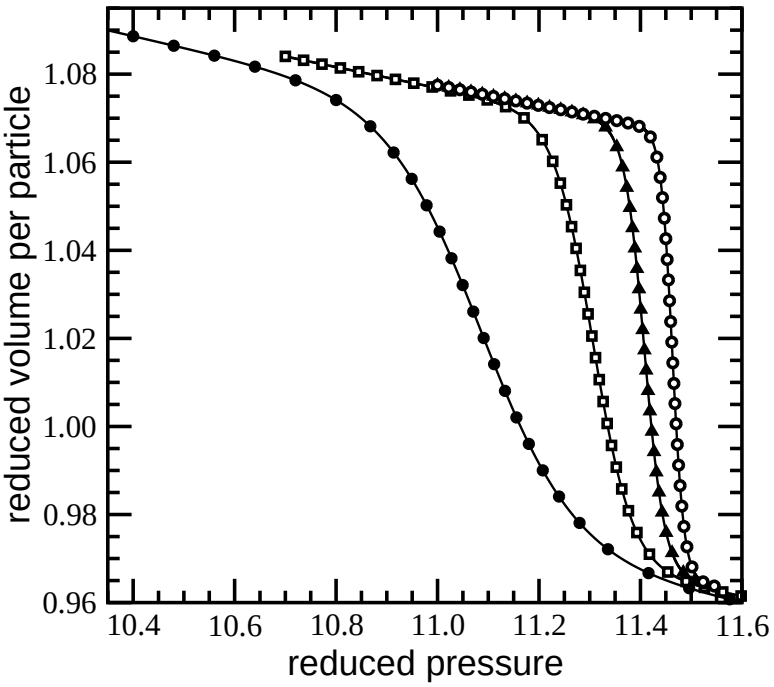


Figure 2.5: Variation of the reduced volume per particle, v^*, with reduced pressure, $\beta p\sigma^3$, at fluid-solid (fcc) coexistence for the hard sphere system. From left to right, the curves correspond to number of particles (•): $N = 108$, (□): $N = 256$, (▲): $N = 500$, and (○): $N = 864$.

($N \to \infty$) the volume is discontinuous at the transition, for finite systems it varies gradually and continuously from the low–p fluid phase to the high–p solid phase. This gradual variation becomes sharper and sharper as the system size, N, increases. The same rounding effects are also observed in the variation of the fraction of singly occupied cells, $-\langle M\rangle/N$, with pressure, shown in Fig. 2.6. In that case, the fraction of singly occupied cells rises smoothly from $\sim$72% (fluid) to complete single-cell occupancy (solid).

In addition to the N-dependent coexistence pressure determined through the equal area

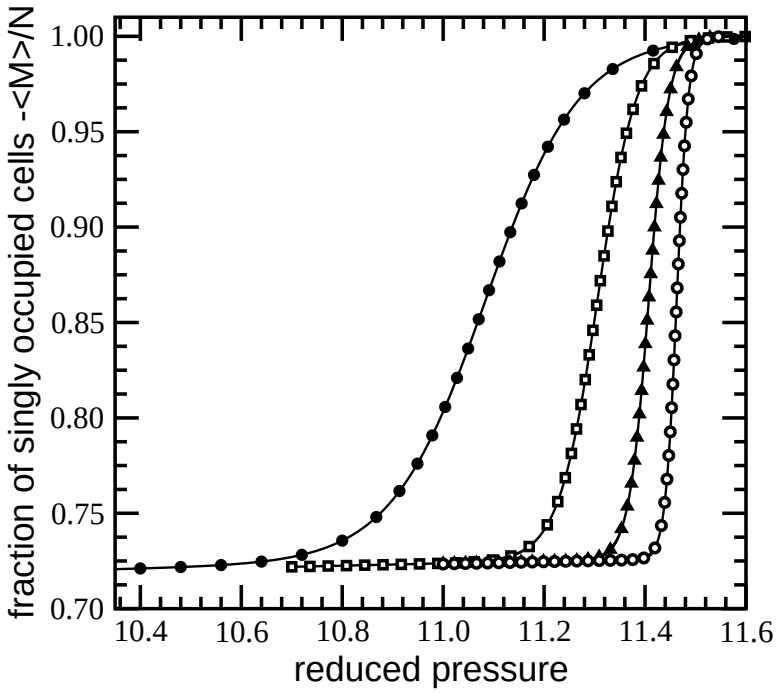


Figure 2.6: Variation of the fraction of singly occupied cells, $-\langle M\rangle/N$, with reduced pressure, $\beta p\sigma^3$, at fluid-solid (fcc) coexistence for the hard sphere model. The labeling of the curves is the same as that of Fig. 2.5.

criterion [cf. Fig. 2.4], other estimators can also be determined by considering second- and higher-order derivatives of the Gibbs free energy. The susceptibility, χ_2, is defined as

$$\chi_2 = \frac{1}{N}\frac{\partial^2 \ln \Delta}{\partial(\beta p)^2} = -\frac{\partial\langle v\rangle}{\partial\beta p} = \frac{1}{N}\langle(\delta V)^2\rangle, \tag{2.8}$$

where $\delta V = V - \langle V\rangle$. This function, which is essentially the slope of the v–p curves in Fig. 2.5, is plotted in Fig. 2.7 in terms of the pressure. In the limit of infinite size, the v–p curve is discontinuous at the transition, and hence χ_2 is a δ-function. For finite systems, χ_2 is a "rounded" δ-function as it is evident in Fig. 2.7. Finite-size scaling theory for first-order phase transitions [9, 8, 6, 7] asserts that the maximum, $\chi_2^{(\max)}$, and the pressure that corresponds to the maximum, $p^{(\max)}$, have the following scaling behavior

$$\chi_2^{(\max)} \sim N, \tag{2.9}$$

$$p^{(\max)} - p^{(\infty)} \sim \frac{1}{N}, \tag{2.10}$$

in the limit of large N. In Eq. (2.10), $p^{(\infty)}$ is the coexistence pressure in the $N \to \infty$ limit. The scaling behavior of the susceptibility maximum, $\chi_2^{(\max)}$, is shown in the inset of Fig. 2.7. It is observed that $\chi_2^{(\max)}$ is linear in N in accord with the scaling ansatz, Eq. (2.9). The N-dependent pressures, $p^{(\max)}$, that correspond to the maxima are nearly indistinguishable from those obtained through the equal area criterion. Consideration of the behavior of the slopes of the curves in Fig. 2.6 leads to similar conclusions.

To this end, consider the third-order susceptibility, χ_3, defined as

$$\chi_3 = \frac{\partial\chi_2}{\partial\beta p} = \frac{1}{N}\frac{\partial^3 \ln \Delta}{\partial(\beta p)^3} = -\frac{1}{N}\langle(\delta V)^3\rangle. \tag{2.11}$$

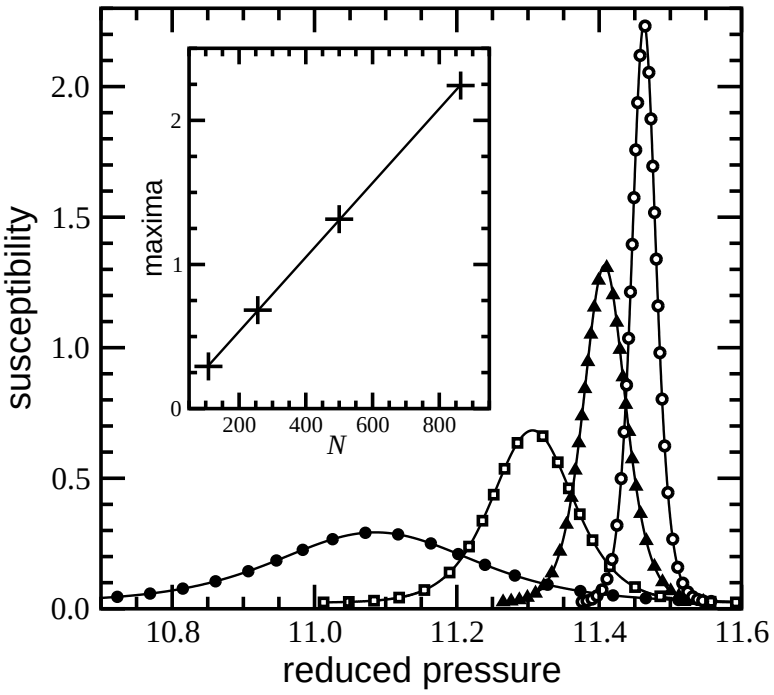


Figure 2.7: Variation of the susceptibility, χ_2, with reduced pressure, $\beta p\sigma^3$, at the fluid-solid coexistence for the hard sphere system. The labeling of the curves is the same as that of Fig. 2.5. The inset shows the behavior of the susceptibility maxima with system size, N.

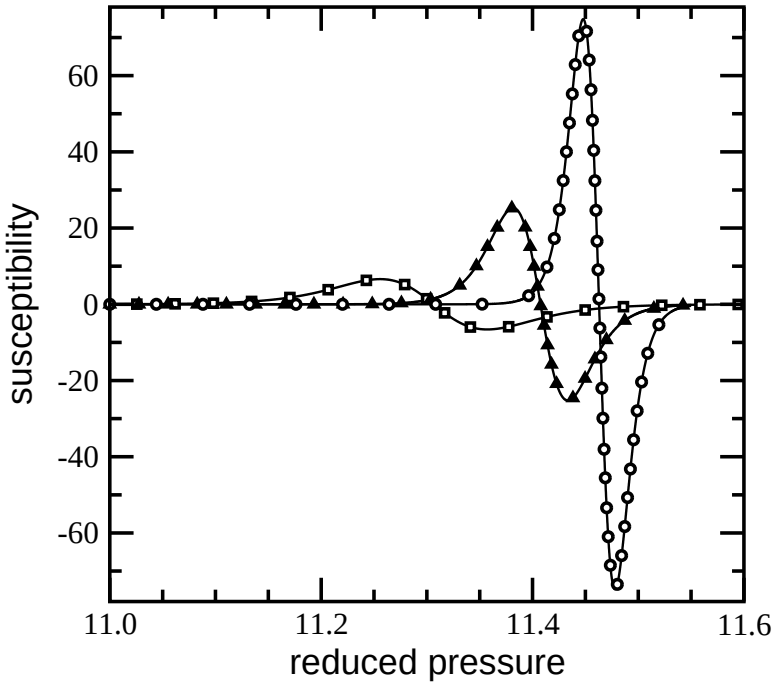


Figure 2.8: Variation of the 3rd-order susceptibility, χ_3, with reduced pressure, $\beta p\sigma^3$, at the fluid-solid coexistence for the hard sphere system. The curves correspond to number of particles (□): $N = 256$, (▲): $N = 500$, and (○): $N = 864$. The curve that corresponds to $N = 108$ is not shown for clarity.

In the limit of infinite size, χ_2 is a δ-function and hence χ_3 diverges to $+\infty$ ($-\infty$) as the coexistence pressure is approached from below (above). For finite systems, these singularities are rounded. The behavior of χ_3 with pressure is shown in Fig. 2.8. This function is a continuous function of the pressure with two extrema points, a maximum and a minimum respectively. The pressures associated with these two extrema can be used to obtain an estimate of the coexistence pressure in the $N \to \infty$ limit according to the scaling ansatz, Eq. (2.10). Higher-order susceptibilities can also be considered to provide additional estimators. However, these high-order susceptibilities are calculated by evaluating high-order moments of the volume distribution function [see, e.g., Eq. (2.11)], which are not as accurate as low-order moments.

The value of the coexisting pressure in the thermodynamic limit can be obtained by extrapolating the N-dependent estimates towards $N \to \infty$. The revelent scaling plot is shown in Fig. 2.9. Since the fluid-solid transition of hard spheres is a first-order transition, the revelent scaling variable is $X = N^{-1}$. The data shown in Fig. 2.9 correspond to N-dependent pressures obtained from the equal area construction (see Fig. 2.4) and those associated with the two extrema of the third-order susceptibility χ_3 (see Fig. 2.8). The results so obtained conform to a linear fit with satisfactory precision and yield $\beta p\sigma^3 = 11.502 \pm 0.019$ in the limit of infinite size, which compares well with previous estimates for the same system [56, 17, 127, 29]. The three estimates shown in Fig. 2.9 approach the infinite system value from below, i.e., from lower pressures. This situation is different than that observed in second-order transitions for which one can define estimators that approach the infinite

system value from both directions; see, e.g., Fig. 5 of Ref. [97].

The volumes of the coexisting fluid and solid phases can be determined by a similar scaling procedure such as the one used for the pressure in Fig. 2.9. The uncertainty in the N-dependent value of the pressure, however, causes considerable scatter in the N-dependent values of the volumes. A similar situation has also been observed in the determination of the critical density in second-order phase transitions, see, e.g., Fig. 6 of Ref. [97]. Nonetheless, estimates for the coexisting volumes can be obtained by implementing constant-pressure simulations for the fluid and the solid phase at the best estimate of the coexisting pressure ($\beta p\sigma^3 = 11.502$) for large systems ($N = 1372$ to 4000). Accounting for the calculated uncertainty in the pressure, these computations yielded the following estimates: $v^*_{\text{fluid}} = 1.0659 \pm 0.0010$ and $v^*_{\text{fcc}} = 0.9651 \pm 0.0009$. These values are in good accord with previous estimates for the same system [56, 17, 127, 29].

2.4 Supercritical Lennard-Jones

Since the behavior of supercritical Lennard-Jones is analogous to that of the hard-sphere model, the previous methodology has been applied to a system of Lennard-Jones particles at a reduced temperature $T^* = 2$. Since the critical temperature of the Lennard-Jones system is $T^* \cong 1.3$ [126], the isotherm at $T^* = 2$ is supercritical. First consider now a system of N particles of diameter σ that interact according to a pair-wise additive potential, ϕ,

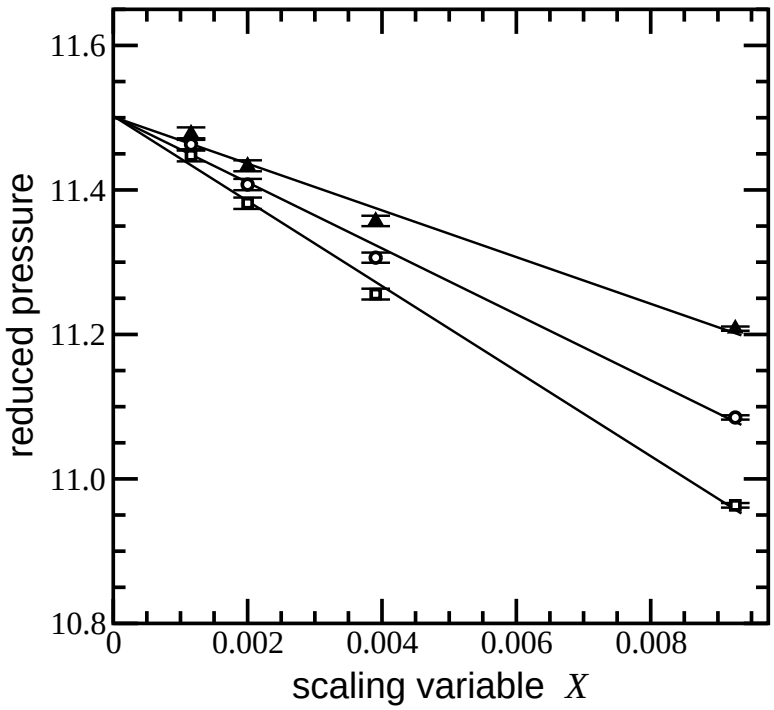


Figure 2.9: Determination of the fluid-solid coexistence pressure for the hard sphere system. The N-dependent coexistence pressures are plotted vs the scaling variable $X = N^{-1}$. The open circles, (○) correspond to the coexistence pressures obtained from the equal-area criterion (see Fig. 2.4). The filled triangles, (▲), and open squares, (□), correspond to the pressures at the extrema of the 3rd-order susceptibility (see Fig. 2.8). The solid lines are linear fits.

which is of the Lennard-Jones form, i.e.,

$$\phi(r) = 4\varepsilon \left[\left(\frac{\sigma}{r}\right)^{2n} - \left(\frac{\sigma}{r}\right)^{n} \right], \tag{2.12}$$

with $n = 6$. Also, in Eq. (2.12), r is the distance between two particles and $\varepsilon > 0$ is the depth of the potential in units of energy. As is common practice in simulations under periodic boundaries, the potential energy of a configuration is calculated by explicitly enumerating all pair terms for distances up to one-half of the length of the simulation box and adding a tail correction for all interactions associated with distances greater than one-half the box length [37]. The long-range correction term is approximate since it assumes that the radial distribution function is unity for distances greater than one-half of the length of the simulation box, which is not necessarily true for small systems or for the solid phase. However, since the long-range part of $\phi(r)$ decays rapidly to zero at large r (i.e., $\sim r^{-6}$), the effects associated with long-range interactions become progressively smaller as the size of the simulated system increases. For cases for which the long-range part of the pair potential decays slowly to zero at large r (such as ionic systems, for instance), the previous assertion is not true and special summation techniques (e.g., Ewald sums) are required.

As in the case of hard spheres (i.e., Section 2.3), there is a region of pressures, $5 \lesssim p^* \lesssim 6$, for which the volume distribution of the solid phase has a bimodal structure, see inset of Fig. 2.10, for instance. The shape of the volume distribution resembles that of a system undergoing some type of phase transition. The pressure–density isotherm of the solid phase exhibits an inflection point at pressure $p^* \cong 5.64$, which corresponds to a density $\rho \cong 0.70\rho_0$, where $\rho_0\sigma^3 = \sqrt{2}$ is the value of the reduced density at close packing. As in

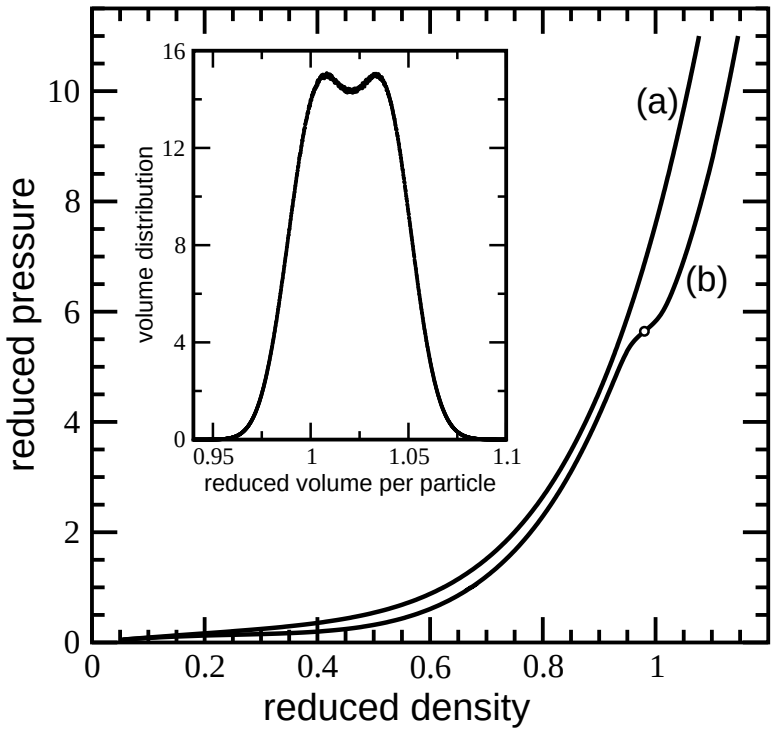


Figure 2.10: Reduced pressure, p^*, vs average reduced density, $\langle \rho^* \rangle$, for $N = 256$ Lennard-Jones particles at $T^* = 2$. (a): fluid; (b): constrained cell model. The inset shows the distribution of the reduced volume for the constrained cell model at $p^* = 5.644$. The state shown in the inset corresponds to the open circle, ($\circ$), on curve (b) in the main figure.

the case of hard spheres, a solid phase at densities $\rho < 0.7\rho_0$, cannot survive without the confinement imposed by the walls of the Wigner-Seitz cells and it thus rapidly melts to a disordered, fluid-like phase.

The effect of the external field b, at constant pressure for $N = 256$ Lennard-Jones particles is shown in Fig. 2.11. Specifically, the variation of the field variable, b, in terms of the average fraction of singly occupied Wigner-Seitz cells, $-\langle M \rangle / N$, is plotted in Fig.

2.11 for five isobars at pressures $p^* = 5$, 6, 7, 8, and 9. Once again we see similar behavior to the system of hard-spheres. That is, as b is reduced at fixed pressure from high values ($b \sim 4$) towards zero, the transformation of the solid to the fluid phase is continuous at low and intermediate pressures, see, e.g., isobar $p^* = 5$ in Fig. 2.11. Additionally, at high-field values, only single occupancy (i.e., $|M| = N$) configurations survive and the system is in the solid phase. At higher pressures, $p^* \gtrsim 6$, the shape of the isobars is similar to that of a system undergoing a first-order phase transition. The simulations indicate that there is a range of values of the field b, for which the distribution of volume V, and of the number of singly occupied cells $|M|$, both consist of two maxima (peaks) separated by a minimum. The inset of Fig. 2.11 shows the shape of the volume distribution for $b = 0.218$ and $p^* = 9$. For the state shown in the inset of Fig. 2.11, the average fraction of singly occupied cells is 90%. The specific value of b ($b = 0.218$) was found through histogram reweighting. The low-volume peak of the distribution shown in the inset of Fig. 2.11 is centered at $v^* \cong 0.905$ and it corresponds to a phase with one particle per Wigner-Seitz cell, i.e., solid. The high-volume peak, centered at $v^* \cong 0.965$ corresponds to a disordered phase with an average fraction of singly occupied cells $\sim 76\%$. Fig. 2.11 thus indicates that at high pressures, $p^* \gtrsim 6$, the transformation of the high-b, ordered, solid-like phase to the low-b, disordered, fluid-like phase occurs via a first-order phase transition. Furthermore, as the pressure increases, the order-disorder phase transition occurs at lower field values. This point is reinforced by Fig. 2.12 which shows the results for the coexisting densities of the ordered (solid-like) and the disordered (fluid-like) phase. This figure also includes the

relevant portion of the p–ρ isotherms of the fluid and the solid phase of the Lennard-Jones model at $T^* = 2$. Phase coexistence for a given b was defined by the equal-weight criterion, as said earlier.

The coexisting densities shown in Fig. 2.12, once again appear to be converging towards a terminal point in the neighborhood of the mechanical stability point of the p–ρ isotherm of the solid phase which is shown as a bullet, ($\bullet$), in Fig. 2.12.

Before analyzing the fluid-solid transition of the Lennard-Jones model in the high-p, $b \to 0$ limit of the field-induced, order-disorder transition of the modified cell model, it is instructive to comment on the usefulness of this model in thermodynamic integration techniques. In the context of constant-pressure simulations, one typically calculates free energy (e.g., chemical potential) differences in terms of pressure, using histogram reweighting, for both the fluid and the solid phase, for a series of states starting from a reference state of known free energy. The reference state is often a $p \to 0$, $\rho \to 0$ ideal-gas state. As already commented, the main drawback of these techniques is the large number of simulated states. Since there is no direct path that can link the two phases at a finite pressure (or density), thermodynamic integration requires the complete calculation of the fluid- and solid-phase isotherms from the low-p ideal-gas region, to the high-p nearly incompressible region, see, e.g., Fig. 2.10. Hoover and Ree [55] introduced the modified cell model as a means to provide a constant-p (or a constant-ρ) path that can connect the solid and the fluid phase, thus reducing the number of simulated states in thermodynamic integration techniques. Referring to Fig. 2.12 for instance, one can independently simulate the fluid and the solid phase

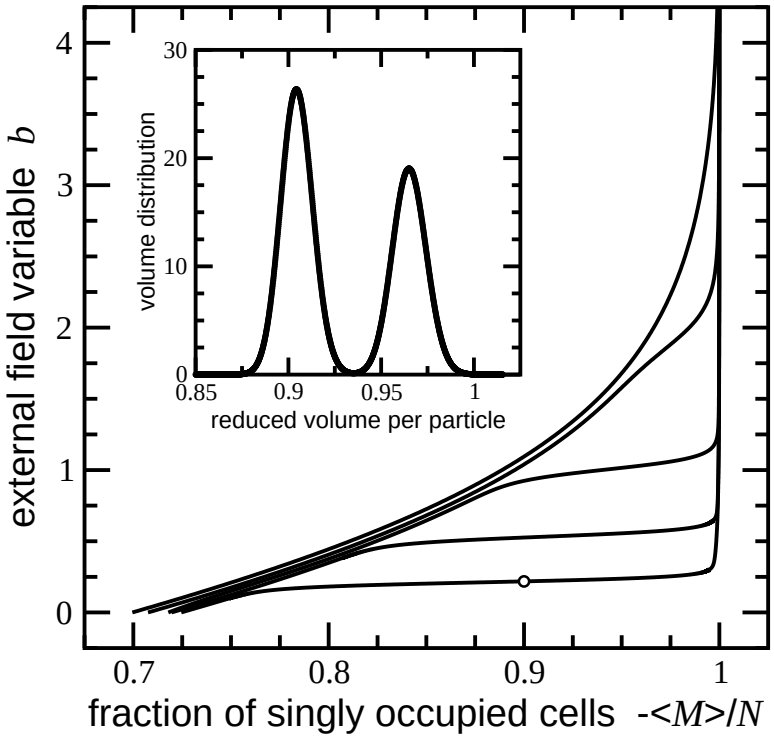


Figure 2.11: External field variable b, vs average fraction of singly occupied cells $-\langle M\rangle/N$, for $N = 256$ Lennard-Jones particles at $T^* = 2$. The curves from top to bottom correspond to reduced pressures $p^* = 5$, 6, 7, 8, and 9, respectively. The inset shows the volume distribution at $p^* = 9$ and $b = 0.218$. For this state, the average fraction of singly occupied cells is 90%. This state corresponds to the open circle, ($\circ$), on the $p^* = 9$ isobar of the main figure.

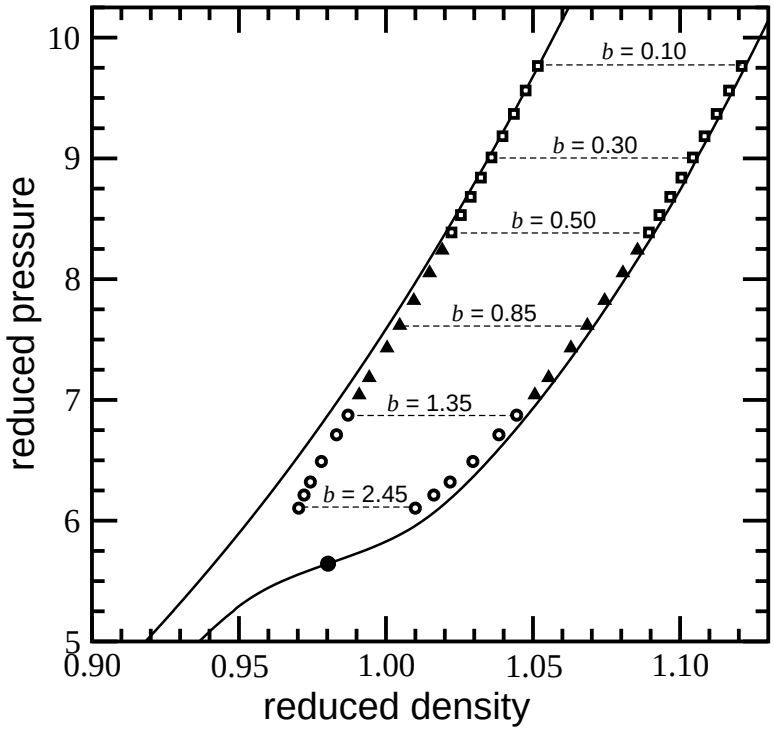


Figure 2.12: Phase diagram of the modified cell model for the Lennard-Jones system at $T^* = 2$ in the pressure-density plane. The solid lines correspond to the fluid and the constrained solid phases [i.e., curves (a) and (b) in Fig. 2.10]. The coexisting solid and fluid phases for finite values of b are shown as points and they correspond to system sizes, N, of ($\circ$): 2048, ($\blacktriangle$): 1372, and ($\square$): 864. The dashed horizontal lines are tie-lines and they are drawn for clarity. The point shown as bullet, ($\bullet$), corresponds to the inflection point of the $T^* = 2$ isotherm of the constrained cell model.

for pressures $p^* \geqslant 5$ and connect the chemical potentials of both phases at $p^* = 5$ through simulations of the modified cell model. In fact, the previous task can be performed at a pressure which is higher than $p^* = 5$, provided that the averaging is done over the appropriate bimodal distributions (such as that shown in the inset of Fig. 2.11, for instance).

2.4.1 Analysis of the freezing transition of supercritical Lennard-Jones

In the present subsection, the high-p, $b \to 0$ limit of the field-induced order-disorder transition of the modified cell model (see Fig. 2.12) is used to analyze the freezing transition of the Lennard-Jones model system on the $T^* = 2$ isotherm. The volume distributions at fluid-solid (fcc) coexistence, obtained through the equal-weight criterion [12], are shown in Fig. 2.13 for $N = 864$ and 2048 particles. To avoid clutter, the distributions for the other systems are not shown. In the context of the equal-weight criterion, size-dependent phase coexistence may be defined by locating (through histogram reweighting) the pressure for which the resulting volume distribution consists of two peaks of equal area. The N-dependent pressures and densities so obtained, can be extrapolated towards the thermodynamic limit ($N \to \infty$) using finite-size scaling techniques for first-order phase transitions [9, 8, 6, 7]. Similar to the previous case of hard-spheres, additional estimators for the pressure and the densities based on second- and higher-order derivatives of the partition function (or the Gibbs free energy), will be investigated to attain satisfactory precision.

The behavior of the first-order derivatives of the partition function Δ_m [cf. Eq. (2.3)],

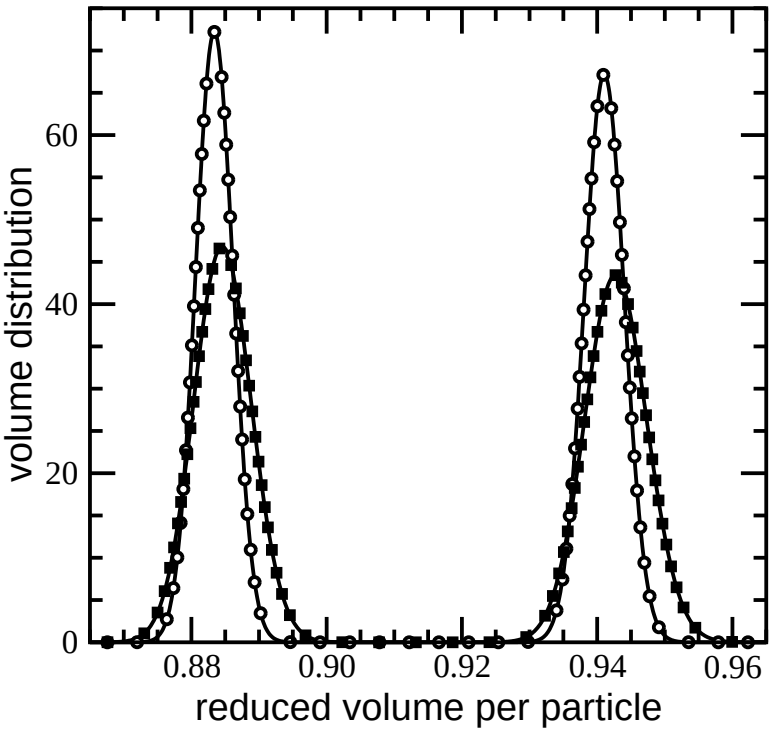


Figure 2.13: Distribution of reduced volume per particle, v^*, for the Lennard-Jones system at fluid-solid (fcc) coexistence at $T^* = 2$, determined through the equal-weight criterion. The distributions correspond to (○): $N = 2048$, and (■): $N = 864$.

the average volume per particle and the average fraction of singly occupied cells,

$$\langle v \rangle = -\frac{1}{N}\frac{\partial \ln \Delta}{\partial \beta p} \quad \text{and} \quad -\frac{\langle M \rangle}{N} = \frac{1}{N}\frac{\partial \ln \Delta}{\partial b}, \tag{2.13}$$

is shown in Figs. 2.14 and 2.15 in terms of the reduced pressure, p^*. In the thermodynamic limit ($N \to \infty$), both $\langle v \rangle$ and $\langle M \rangle / N$ are discontinuous at the transition. For finite systems, however, they vary continuously at the transition and this gradual variation gets progressively sharper as N increases. The fraction of singly occupied cells, shown in Fig. 2.15, increases continuously from $\sim 73\%$ in the low-p fluid phase to 100% in the high-p solid phase. The behavior of the functions shown in Figs. 2.14 and 2.15 is generally referred to as the rounding and shifting effect [9, 8, 6, 7] associated with phase transitions for finite systems.

In the context of hard-sphere freezing, the second- and third-order derivatives (susceptibilities) of the partition function are given by 2.8 and 2.11. The second-order susceptibility, χ_2, which is the negative slope of the v–p isotherms in Fig. 2.14, is shown in Fig. 2.16 in terms of pressure, p^*. In the limit of infinite size ($N \to \infty$), the v–p isotherm has a jump discontinuity at the transition and hence χ_2 is a δ-function. For finite N however, χ_2 is a rounded and shifted δ-function as is evident in Fig. 2.16. The third-order susceptibility, χ_3, shown in Fig. 2.17, is the slope of χ_2 [cf. Eq. (2.11)]. Since χ_2 is a δ-function in the thermodynamic limit, χ_3 diverges to $+\infty$ ($-\infty$) as the transition is approached from below (above). For finite systems, these $\pm\infty$ singularities are rounded and shifted, as Fig. 2.17 indicates. Specifically, χ_3 is a continuous function with two extrema points, a low-p maximum and a high-p minimum, respectively. The location of the extrema points can

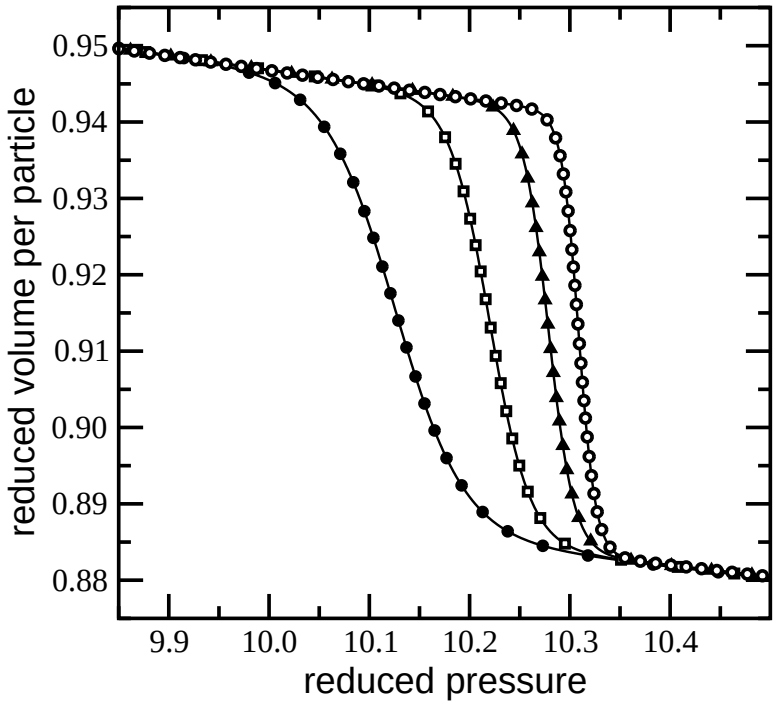


Figure 2.14: Variation of the average reduced volume per particle, $\langle v^* \rangle$, with reduced pressure, p^*, at fluid-solid (fcc) coexistence for the Lennard-Jones system at $T^* = 2$. From left to right, the curves correspond to number of particles (●): $N = 500$, (□): $N = 864$, (▲): $N = 1372$, and (○): $N = 2048$.

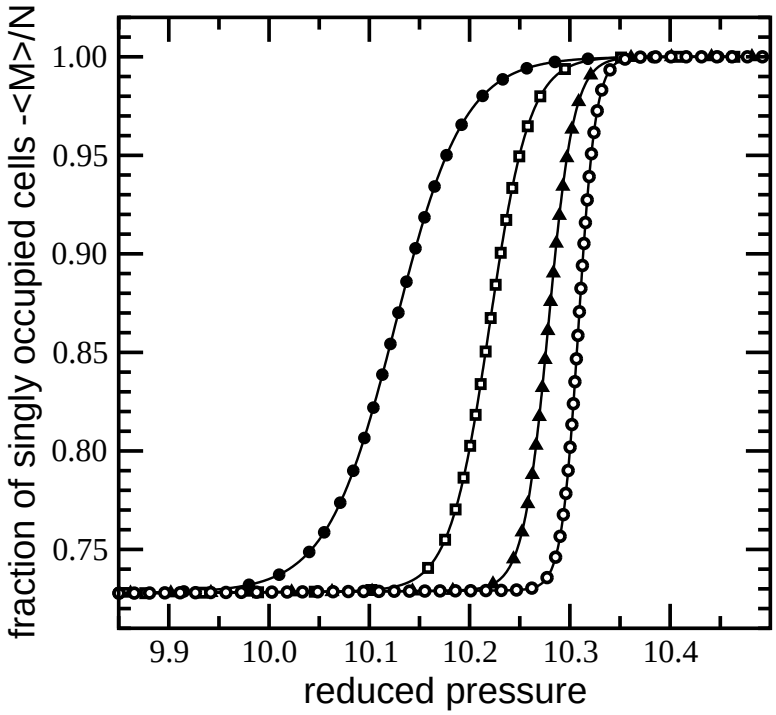


Figure 2.15: Variation of the average fraction of singly occupied cells, $-\langle M\rangle/N$, with reduced pressure, p^*, at fluid-solid (fcc) coexistence for the Lennard-Jones system at $T^* = 2$. The labeling of the curves is the same as that of Fig. 2.14.

be used to obtain estimates of the transition pressure through finite-size scaling techniques for first-order phase transitions [9, 8, 6, 7]. Additional susceptibilities may be defined by evaluating the corresponding derivatives of the partition function with respect to the field, b. The resulting extrema points however, are nearly indistinguishable from those shown in Figs. 2.16 and 2.17. Furthermore, fourth- and higher-order derivatives of the partition function may be defined to provide additional extrema points. However, these high-order derivatives are evaluated through high-order moments of the distribution of states [see, e.g., Eq. (2.11) for χ_3] which are not as accurate as low-order moments.

Finite-size scaling theory [9, 8, 6, 7] is a phenomenological approach that describes the rounding and shifting effects associated with phase transitions for finite systems (see, e.g., Figs. 2.14–2.17) in terms of universal scaling laws. It can be used to obtain estimates of the coexisting properties from data obtained on finite systems through extrapolation towards the limit of infinite size. Finite-size scaling theory for first-order phase transitions, assumes that the distribution of the order parameter (e.g., volume, density) is a superposition of two gaussian distributions centered at the two coexisting phases, see, e.g., Fig. 2.13. Based on the double-gaussian approximation for the distribution of the order parameter, it can be readily demonstrated that the extremum (maximum, say) of a susceptibility, $\chi^{(\max)}$, and the pressure at the extremum, $p^{(\max)}$, have the following scaling behavior: In Eq. (2.10), $p^{(\infty)}$ is the coexistence pressure in the $N \to \infty$ limit. The scaling laws, Eqs. (2.9) and (2.10), are strictly valid for large N, since terms associated with corrections to scaling are neglected. These corrections to scaling terms are important for small/moderate values of N and they

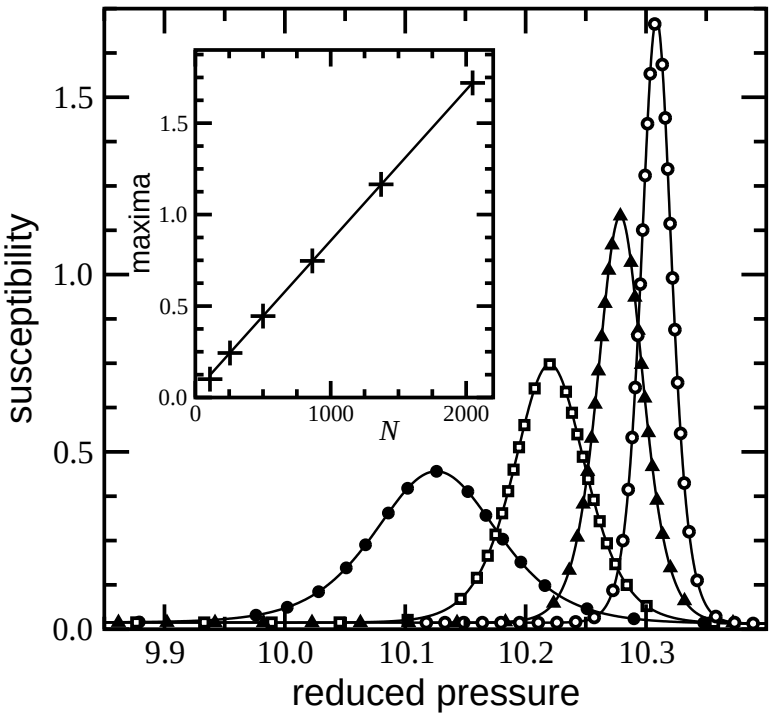


Figure 2.16: Variation of the susceptibility, χ_2, with reduced pressure, p^*, at fluid-solid coexistence for the Lennard-Jones system at $T^* = 2$. The labeling of the curves is the same as that of Fig. 2.14. The inset shows the behavior of the susceptibility maxima with system size, N.

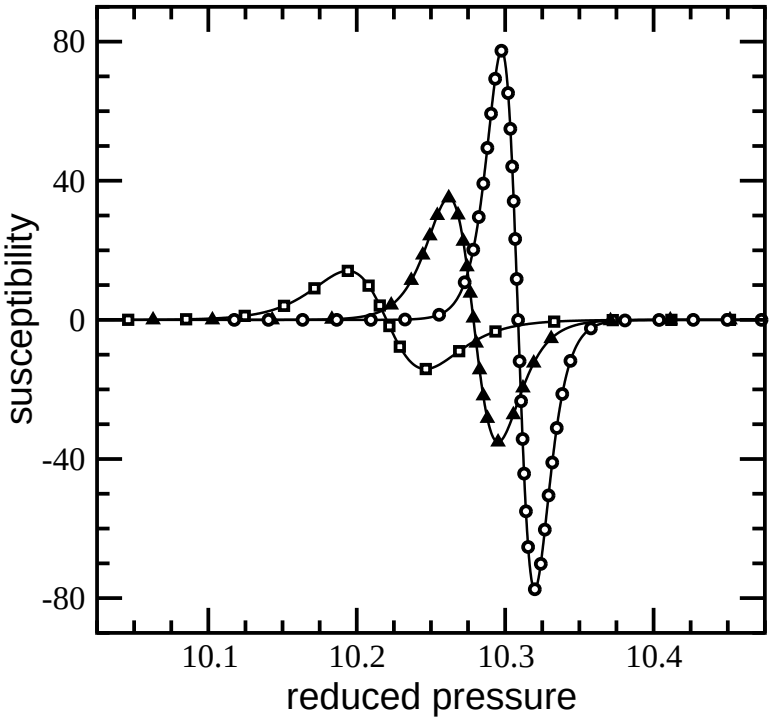


Figure 2.17: Variation of the 3rd-order susceptibility, χ_3, with reduced pressure, p^*, at fluid-solid coexistence for the Lennard-Jones system at $T^* = 2$. The curves correspond to number of particles (□): $N = 864$, (▲): $N = 1372$, and (○): $N = 2048$.

originate primarily from mixed-phase configurations that are intermediate (e.g., between the two peaks, see Fig. 2.13) to those of the two bulk coexisting phases.

The scaling behavior of the maximum of χ_2 is shown in the inset of Fig. 2.16 for $N = 108$ to 2048. The maximum of χ_2 obeys the scaling ansatz, Eq. (2.9), with satisfactory accuracy. As already emphasized, the coexistence pressure in the thermodynamic limit, $p^{(\infty)}$, can be estimated by extrapolating the N-dependent values, $p^{(\max)}$, to $N \to \infty$ according to Eq. (2.10). The relevant scaling plot is found in Fig. 2.18. In that figure, three pressure estimators are shown: the one obtained through the equal-weight criterion and the positions of the two extrema of χ_3 (see Fig. 2.17). The pressure associated with the maximum of χ_2 is nearly indistinguishable from that obtained through the equal-weight criterion. The anticipated linearity of $p^{(\max)}$ vs $1/N$ [cf. Eq. (2.10)] is observed for $N > 256$. This behavior is distinctly different from that seen for hard spheres for which linearity could be seen for N as low as 108.

The main contributing factor for the peculiar behavior of the pressure estimators in Fig. 2.18 for small N is the approximate nature of accounting for tail (e.g., long range) corrections in the evaluation of the energy. A similar behavior was also seen by Errington [29] in simulations of the freezing transition of the same system for $N = 108$, 256, and 500, via the phase switch method. Errington accounted for tail corrections via the standard way (as was done in this work) and via a lattice-based type of correction for solid-like configurations. The two results are different but the difference gets progressively smaller as N increases. For example, for $N = 500$ the two results differ by $\sim 0.2\%$. Hence, in

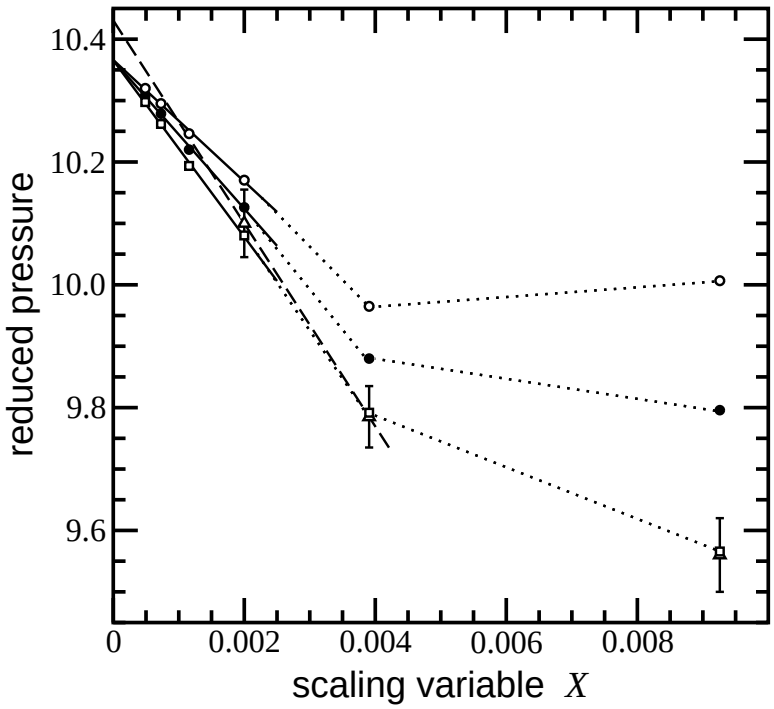


Figure 2.18: Determination of the fluid-solid coexistence pressure for the Lennard-Jones system at $T^* = 2$. The N-dependent coexistence pressures are plotted vs the scaling variable $X = N^{-1}$. The filled circles, (•), correspond to the coexistence pressures obtained from the equal-area criterion (see Fig. 2.13). The open circles, (○), and squares, (□), correspond to the pressures at the extrema of the 3rd-order susceptibility (see Fig. 2.17). The solid lines are linear fits. The points that correspond to $N = 108$ and 256 particles are excluded from the fit. Dotted lines are drawn for visual clarity. Statistical uncertainties do not exceed the size of the symbols. The points shown as open triangles, (△), are the results of the phase switch method [29] for the same system.

order to obtain the coexistence pressure, one must consider sizes $N \gtrsim 500$, a point that was also emphasized by Errington [29]. Consequently, only the results for $N = 500$, 864, 1372, and 2048 are fitted in terms of the scaling variable $X = N^{-1}$. As it can be seen in Fig. 2.18, the data for $N \geqslant 500$ conform to a linear fit with satisfactory precision and yield $p^* = 10.366 \pm 0.004$ in the limit of infinite size.

The determination of the densities of the two coexisting phases can be done by an analogous procedure as that of the pressure shown in Fig. 2.18. The simplest density estimator is associated with the N-dependent values of the first-moments of the two peaks of the volume distribution at coexistence, determined through the equal-weight criterion, see Fig. 2.13. Additional estimators may be formulated by considering dimensionless moment ratios of the order parameter. These ratios were introduced in phase transition studies of magnetic models [112, 11, 24, 10] (such as Ising, for instance) and were used by Fisher and coworkers [64, 63] to develop scaling algorithms for precise simulation of criticality in asymmetric fluids. The most common of these ratios is the 4-th order ratio, Q_4, which, in the context of this work, may be defined as follows:

$$Q_4 = \frac{\langle y^2 \rangle^2}{\langle y^4 \rangle} \quad \text{with} \quad y = v - \langle v \rangle. \tag{2.14}$$

For single-phase states, $Q_4 \to 1/3$ as $N \to \infty$ [112, 11, 24, 10, 64, 63]. For two-phase states, analysis based on the double-gaussian approximation indicates that this ratio has a singular behavior [112, 11, 24, 10, 64, 63]. Specifically, as the transition is approached from the single-phase region, Q_4 jumps discontinuously from $1/3$ to zero. Furthermore, for two-phase states, Q_4 varies between 0 and 1. The maximum value of unity corresponds to an

average density, $\langle\rho\rangle$, which is the arithmetic mean of the densities for the two coexisting phases. For finite systems, however, the behavior of Q_4 around the discontinuities and the maximum is rounded and shifted.

The behavior of Q_4 is shown in Fig. 2.19 in terms of the average reduced density $\langle\rho^*\rangle = 1/\langle v^*\rangle$ for several values of N. While at the thermodynamic limit the jump discontinuities from $Q_4 = 1/3$ to $Q_4 = 0$ occur at the respective coexisting densities, they are rounded and shifted for finite systems, as is evident in Fig. 2.19. The same is also true for the maximum of Q_4. The locations of the two minima in Fig. 2.19 serve as an estimator for the densities of the two coexisting phases. In addition to Q_4, other types of ratios can also be constructed. In the present work, the following 6-th order ratio has also been considered:

$$Q_6 = \frac{\langle y^2\rangle^3}{\langle y^6\rangle} \quad \text{with} \quad y = v - \langle v\rangle. \tag{2.15}$$

The behavior of Q_6 is analogous to that of Q_4 shown in Fig. 2.19. Specifically, it consists of two rounded and shifted minima separated by a maximum. The densities associated with the two minima yield another estimator for the densities of the two coexisting phases. All these estimators can be extrapolated towards $N \to \infty$ according to the scaling ansatz, Eq. (2.10).

The scaling plots of densities of the coexisting phases are shown in Fig. 2.20 (fluid) and Fig. 2.21 (solid). The three estimators shown in these two figures correspond to densities obtained through the equal-weight criterion (cf. Fig. 2.13) and to those associated with the minima of Q_4 (cf. Fig. 2.19) and Q_6. As in the case of the pressure in Fig. 2.18, the peculiar behavior of the data in Figs. 2.20 and 2.21 for small N is due to the approximate

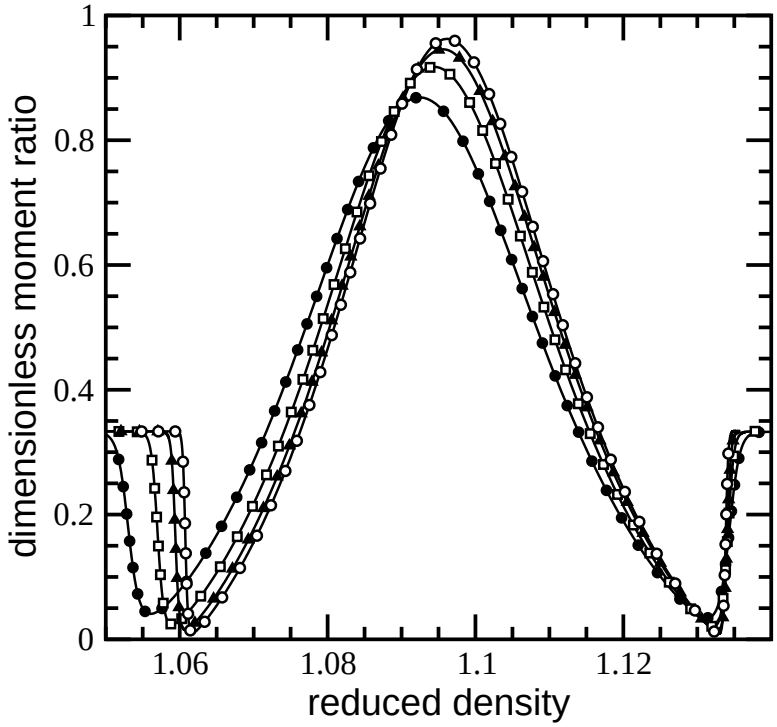


Figure 2.19: Dimensionless moment ratio, Q_4, vs average reduced density, $\langle \rho^* \rangle$, for the Lennard-Jones system at $T^* = 2$. The curves correspond to number of particles (•): $N = 500$, (□): $N = 864$, (▲): $N = 1372$, and (○): $N = 2048$.

nature of accounting for tail corrections in the energy evaluations. The points associated with $N = 108$ and 256 are thus excluded from further analysis. The $N \geq 500$ data for the densities appear to obey the anticipated linearity with respect to the scaling variable $X = N^{-1}$. A close look at the ordinates in Figs. 2.20 and 2.21, however, indicates that the densities associated with the minima of Q_4 and Q_6 extrapolate to a value that is slightly different than that of the density obtained through the equal-weight criterion. The reason for this behavior is related to the way with which these estimators are calculated. Recall that the calculation of Q_4 and Q_6 entails high-order moments [cf. Eqs. (2.14) and (2.15)] which are not as accurate as low-order moments. The densities associated with the squares in Figs. 2.20 and 2.21 are the first moments of the peaks of the distributions shown in Fig. 2.13. Thus, the best estimate of the coexisting densities can be obtained by considering data associated with low-order moments, e.g., equal-weight data. The difference between the estimate originating from the equal-weight data and that associated with Q_4 and Q_6, is indicative of the precision (or the resolution) achieved with the present level of analysis and it also serves as an error estimate of the respective values. The equal-weight data thus yield: $\rho^*_{\text{fluid}} = 1.0638 \pm 0.0004$ and $\rho^*_{\text{fcc}} = 1.1330 \pm 0.0005$. Better precision would require much longer simulations for system sizes much larger than $N = 2048$.

In Figs. 2.18, 2.20 and 2.21, the results of this work are compared to those of the phase switch method [29]. As already commented, previous work via the phase switch method considered systems of $N \leqslant 500$ particles and only used the equal-weight construction to obtain N-dependent estimates of the coexisting pressures and densities. The system sizes

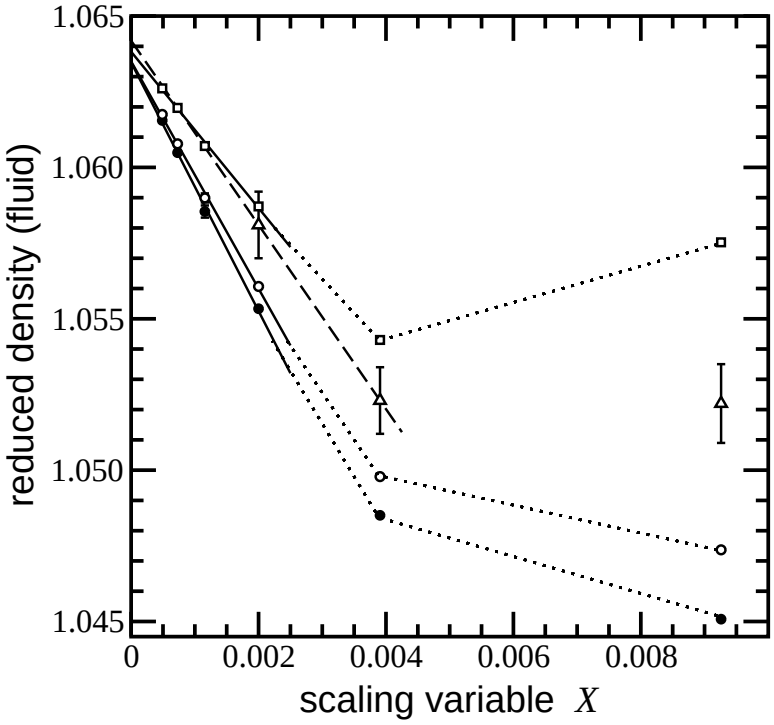


Figure 2.20: Determination of the coexistence density of the fluid phase for the Lennard-Jones system at $T^* = 2$. The N-dependent densities are plotted vs the scaling variable $X = N^{-1}$. The squares, (□), correspond to densities obtained from the equal-weight criterion. The filled, (•), and open, (○), circles correspond to densities obtained from the dimensionless ratios Q_4 and Q_6. The solid lines are linear fits. The points associated with $N = 108$ and 256 particles are excluded from the fit. Dotted lines are drawn for visual clarity. Statistical uncertainties do not exceed the size of the symbols. The points shown as open triangles, (△), are the results of the phase switch method [29] for the same system.

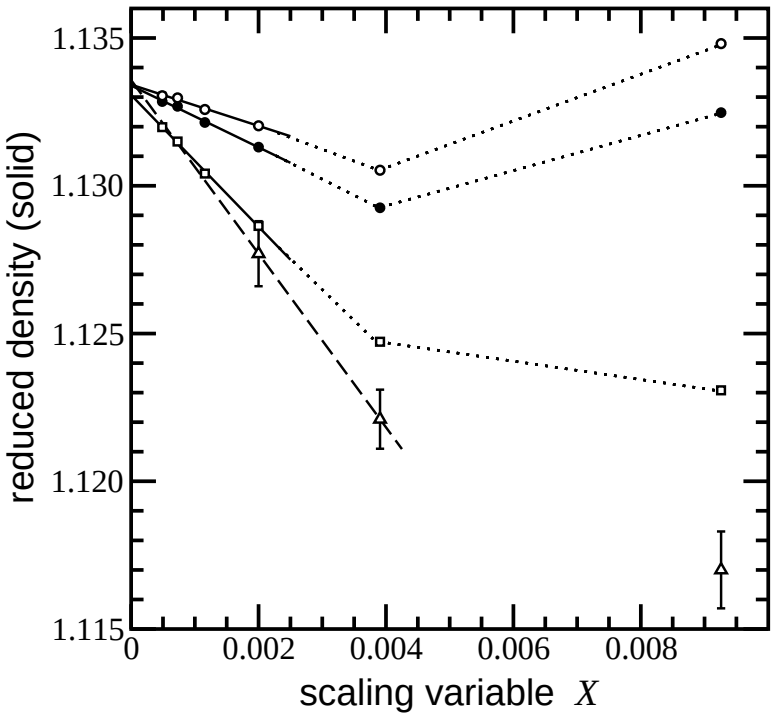


Figure 2.21: Determination of the coexistence density of the solid (fcc) phase for the Lennard-Jones system at $T^* = 2$. The labeling is the same as that of Fig. 2.20.

considered in previous work [29, 78] do not scale with $X = N^{-1}$. Hence, a satisfactory estimate of the pressure and of the densities in the limit of infinite size with which the present results can be compared against, could not be obtained. The size-dependent results for $N = 500$ are in good accord to each other as is evident in Figs. 2.18, 2.20 and 2.21. Note that since the reduced pressures are defined differently, the pressures obtained from the phase switch method [29] must be divided by 2 to ensure consistency. Regarding the data associated with small systems (i.e., $N \leqslant 256$), larger discrepancies are expected due to effects associated with long-range corrections and, more importantly, due to the fact that the volume distribution does not comprise two well-separated peaks. In such cases, the effects of intermediate, mixed-phase configurations render determination of the coexisting pressures and densities via the equal-weight construction ambiguous. Nonetheless, a rough estimate of phase coexistence in the limit of infinite size for the phase switch method may be obtained by connecting the points that correspond to $N = 256$ and $N = 500$ particles with a straight line. Such a construction is shown with a dashed line in Figs. 2.18, 2.20 and 2.21. The intersection of the dashed line with the ordinate provides a rough estimate of phase coexistence in the thermodynamic limit. As is evident in Figs. 2.18, 2.20 and 2.21, the values so obtained are in reasonable accord with the results of this work.

2.5 Subcritical Lennard-Jones

As is well-known [91, 118, 100, 126, 18], the Lennard-Jones fluid undergoes a first-order, gas-liquid phase transition that terminates at a second-order Ising critical point at $T^* \cong$

1.3 and $\rho^* \cong 0.3$. The values of the critical parameters are somewhat sensitive [118] to the details of the truncation of the Lennard-Jones potential. In the present section, it is shown that the corresponding constrained cell model of a system of Lennard-Jones particles exhibits a similar type of phase transition between a dilute, gas-like and a dense, liquid-like phase. The resulting phase diagram of the constrained model in the temperature-density plane is shown in Fig. 2.22 and it is compared against that of the unconstrained system. For both cases, the phase envelopes have been obtained from constant-pressure simulations for $N = 256$ particles using tempering techniques [77, 76] and histogram reweighting [31, 32]. For a given temperature, the densities of the coexisting gas and liquid phases, ρ_{gas} and ρ_{liq}, were obtained through the equal-area criterion [12]. Since the simulated system size is not large, the data does not extend sufficiently close to the critical point. Nonetheless, a rough estimate for the critical temperature, T_c, and density, ρ_c, may be obtained by fitting [37] the coexistence data to the leading power law for the order parameter (density difference) and the law of the rectilinear diameters, i.e.,

$$\rho_{\text{liq}} - \rho_{\text{gas}} \sim (T_c - T)^{\beta}, \tag{2.16}$$

$$\frac{1}{2}\left(\rho_{\text{liq}} + \rho_{\text{gas}} \right) - \rho_c \sim T_c - T, \tag{2.17}$$

where it has been assumed that $\beta = 0.326$, appropriate for the Ising universality class [73]. Application of this approach gives: $T_c^* \cong 1.70$, $\rho_c^* \cong 0.287$ (constrained model) and $T_c^* \cong 1.295$, $\rho_c^* \cong 0.316$ (unconstrained system). As shown in Fig. 2.22, the results for the unconstrained system are in good accord with those obtained by Chen *et al.* [18] using grand canonical simulations and histogram reweighting. Better resolution for the

critical parameters requires data collection for a series of progressively increasing values of the system size and extrapolation towards the limit of infinite size according to finite-size scaling techniques for second-order Ising criticality [97, 64]. As evident in Fig. 2.22, the phase envelope of the unconstrained system lies "inside" that of the constrained system. In other words, the single occupancy constraint causes the gas-liquid transition to persist at higher temperatures and pressures than the unconstrained system. The primary reason for this behavior is the reduction of the entropy of the system caused by the single occupancy constraint.

The pressure-density isotherms of the fluid (unconstrained) and the solid (constrained) phase at $T^* = 1$ are shown in Fig. 2.23. These isotherms have been obtained from constant-pressure simulations of a system of $N = 256$ particles using histogram reweighting [31, 32]. At very low densities and pressures both systems behave as an ideal gas and, hence, the corresponding p–ρ isotherms are nearly-straight and lie very close to each other. As the pressure increases, both systems undergo a first-order, gas-liquid transition and both isotherms are thus nearly horizontal. The region that corresponds to the phase transition has been constructed from the tempering simulations used to obtain the phase diagrams shown in Fig. 2.22. At higher pressures, both systems are nearly-incompressible and, hence, the p–ρ isotherms are nearly vertical. As was also seen previously for hard- and soft-sphere models, the p–ρ isotherm of the solid (constrained) phase exhibits an inflection point at pressure $p^* \cong 1.096$, which corresponds to a density $\rho \cong 0.61\rho_o$, where $\rho_o\sigma^3 = \sqrt{2}$ is the value of the reduced density at close packing. Specifically, there is a range of pressures,

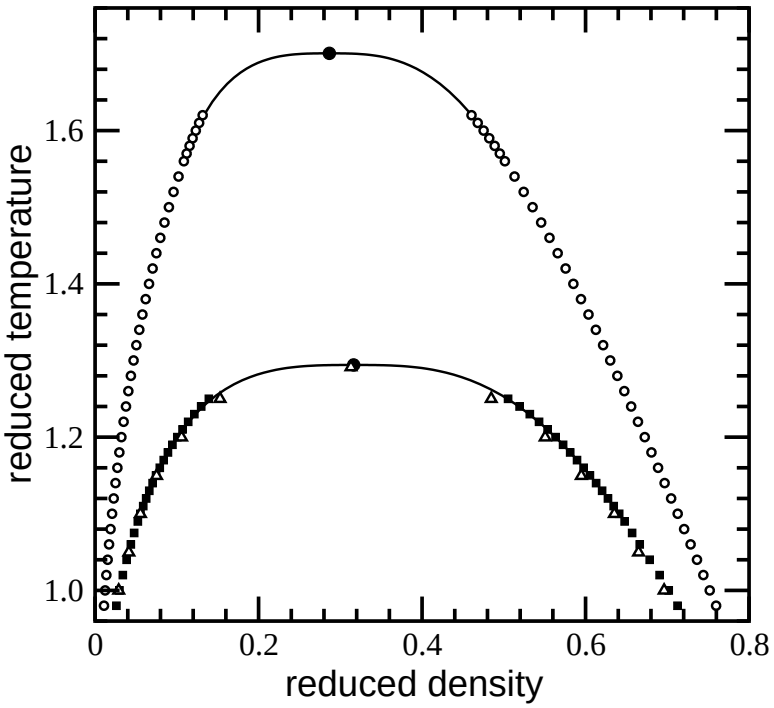


Figure 2.22: Gas-liquid phase diagram for the Lennard-Jones model in the temperature-density plane. The open circles, (○), correspond to the coexisting densities of the constrained cell model. The coexisting densities of the unconstrained system are shown as filled squares, (■). For both cases, the critical points are shown as filled circles, (●). The solid lines are fits to the scaling laws for the order parameter and the coexistence curve diameter, see Eqs. (2.16) and (2.17). The open triangles, (△), are the results of Chen *et al.* [18] for the unconstrained Lennard-Jones system.

$1.0 \lesssim p^* \lesssim 1.25$, for which the volume distribution of the constrained system has a double-peak structure. Such a distribution is shown in the inset of Fig. 2.23. The maxima of this distribution have equal height and the corresponding value of the pressure, obtained through histogram reweighting, is marked with the open circle in the p–ρ isotherm of the solid (constrained) phase in the main figure. As explained in previously, the inflection point is the limit of mechanical stability of the solid. At densities $\rho < 0.61\rho_o$, the solid cannot survive without the confinement imposed by the cell walls and it will quickly disintegrate to a disordered fluid phase. A behavior similar to that shown in Fig. 2.23 was also seen by Hansen and Verlet [52] for the Lennard-Jones model at $T^* = 1.15$. Since they, however, used constant-volume simulations, the p-ρ isotherms of both phases exhibited van der Waals type of loops around the gas-liquid transition region. Furthermore, the inflection point associated with the mechanical stability of the solid phase appeared as an angular point (or kink) in the work of Hansen and Verlet.

In the present work, the modified cell model with partition function given by Eq. (2.3) has been simulated at $T^* = 1$ for pressures $p^* \geqslant 1$. The effect of the external field, b, for a system of $N = 256$ Lennard-Jones particles at $T^* = 1$ is shown in Fig. 2.24. In that figure, the variation of the field, b, is plotted against the fraction of singly occupied cells, $-\langle M \rangle / N$, for five isobars that correspond to pressures $p^* = 1.0$, 1.5, 2.0, 2.5, and 3.0, respectively. These five isobars shown in Fig. 2.24 were obtained by analyzing the histograms obtained from standard constant-pressure simulations through histogram reweighting [31, 32]. As seen previously, the behavior of these five isobars is reminiscent to that of a system un-

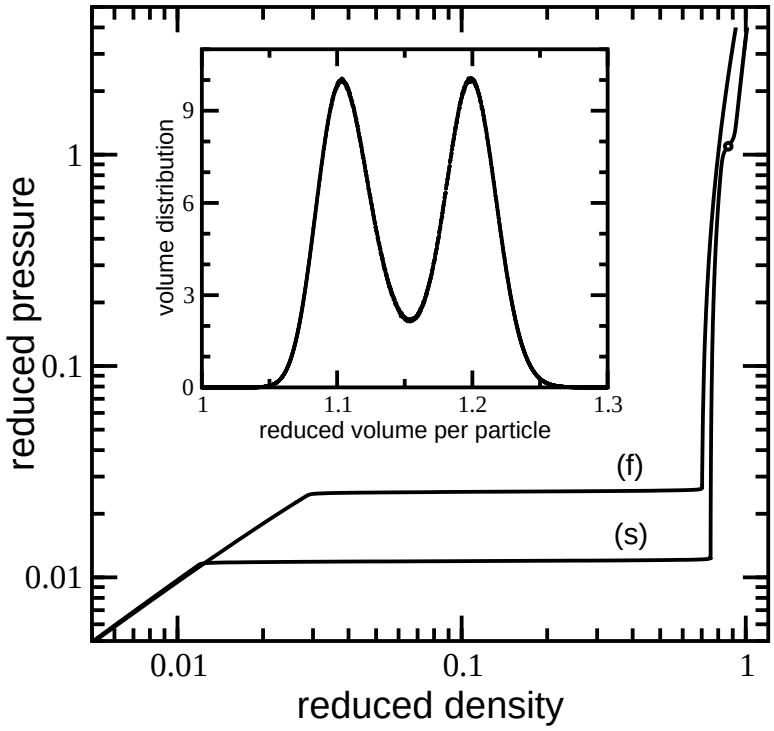


Figure 2.23: Reduced pressure, p^*, vs average reduced density, $\langle \rho^* \rangle$, for $N = 256$ Lennard-Jones particles at $T^* = 1$. (f): fluid (unconstrained model); (s): solid (constrained cell model). The inset shows the distribution of the reduced volume for the constrained cell model at $p^* = 1.096$. The state shown in the inset corresponds to the open circle, ($\circ$), on the p–ρ isotherm of the solid phase in the main figure.

dergoing a first-order phase transition that terminates at a second-order critical point. As already commented, at low and intermediate values of the pressure, the transformation of the solid to the fluid phase is continuous, see, e.g., isobar $p^* = 1.0$ in Fig. 2.24. At higher pressures, $p^* \gtrsim 2.0$, the simulations indicate that there is a range of values of the field, b, for which the distribution of the volume, V, and of the number of singly occupied cells, $|M|$, has a double-peak structure. The inset of Fig. 2.24 shows the volume distribution at $p^* = 2.5$ and $b = 0.508$. The specific value of b was found through the equal-area construction [12]. Hence, the volume distribution shown in the inset of Fig. 2.24 corresponds to finite-system coexistence between a solid phase with one particle per Wigner-Seitz cell and a fluid phase for which the fraction of cells that contain one particle is $\sim 82\%$.

The phase transition implied in Fig. 2.24 has been studied by standard flat-histogram techniques, namely tempering [77, 76], for systems of $N = 864$, 1372, and 2048 particles. The results for the coexisting densities are shown in Fig. 2.25 together with the relevant part of the p–ρ isotherm of the solid (constrained) and the fluid (unconstrained) phase of the Lennard-Jones model at $T^* = 1$. The coexisting densities of the fluid-like and the solid-like phase of the modified cell model, obtained using the equal-area construction [12], lie in between the isotherms of the unconstrained and the constrained system. As p increases, the value of b at coexistence decreases and as $b \to 0$, the coexisting densities approach those of the fluid (unconstrained) and solid (constrained) phase, respectively.

As already commented, fluid-solid coexistence of a system of interest can be obtained by analyzing the field-induced phase transition of the corresponding modified cell model

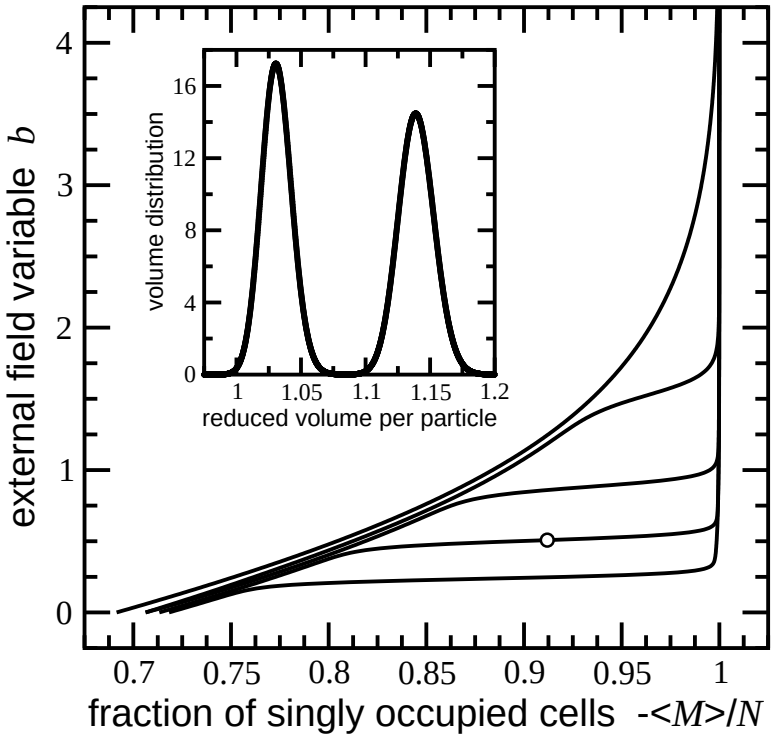


Figure 2.24: External field variable b, vs average fraction of singly occupied cells $-\langle M\rangle/N$, for $N = 256$ Lennard-Jones particles at $T^* = 1$. The curves from top to bottom correspond to reduced pressures $p^* = 1.0$, 1.5, 2.0, 2.5, and 3.0, respectively. The inset shows the volume distribution at $p^* = 2.5$ and $b = 0.508$. The specific value of b was found through the equal-area construction [12]. This state corresponds to the open circle, ($\circ$), on the $p^* = 2.5$ isobar of the main figure.

(cf. Fig. 2.25) in the high-p, $b \to 0$ limit. Alternatively, one may implement a procedure that is similar to thermodynamic integration. Specifically, the fluid (unconstrained) and the solid (constrained) phase are both simulated for a range of pressures, starting from a reference pressure, p_o. Analysis of the simulation data via histogram reweighting, leads to the following expressions for the chemical potentials of the two phases

$$\widetilde{\mu}_f(p) - \widetilde{\mu}_f(p_o) = -\frac{1}{N} \ln \left[\frac{\Delta_f(p)}{\Delta_f(p_o)} \right], \tag{2.18}$$

$$\widetilde{\mu}_s(p) - \widetilde{\mu}_s(p_o) = -\frac{1}{N} \ln \left[\frac{\Delta_s(p)}{\Delta_s(p_o)} \right]. \tag{2.19}$$

The partition functions in Eqs. (2.18) and (2.19) are evaluated thought the density of states, obtained from histogram reweighting. In the absence of a direct thermodynamic path that connects the fluid with the solid phase, the simulations must be extended in the low-density, ideal-gas region for which the chemical potentials of the reference states, $\widetilde{\mu}_f(p_o)$ and $\widetilde{\mu}_s(p_o)$, can be evaluated analytically. The modified cell model can be used to provide the desired thermodynamic path that connects the two phases, thus eliminating the need to extend the simulations all the way to the ideal-gas region. Specifically, the modified cell model is simulated on the constant-pressure path $p = p_o$ for a series of progressively increasing values of the field, $0 \leq b \leq b_\infty$ ($= 10$, say). Analysis of the simulation data via histogram reweighting, leads to the following expression for the chemical potential differences

$$\widetilde{\mu}_m(p_o, b_\infty) - \widetilde{\mu}_m(p_o, 0) = -\frac{1}{N} \ln \left[\frac{\Delta_m(p_o, b_\infty)}{\Delta_m(p_o, 0)} \right] . \tag{2.20}$$

The value $b = 0$ corresponds to an unconstrained system and thus $\widetilde{\mu}_m(p_o, 0) = \widetilde{\mu}_f(p_o)$. Since at large fields, b_∞, only single occupancy configurations are possible, inspection

of Eq. (2.3) indicates that the partition function of the modified cell model, Δ_m, may be written as $\Delta_m(p_o, b_\infty) = \Delta_s(p_o)\exp(b_\infty N)$, which allows the chemical potential of the solid at pressure p, $\widetilde{\mu}_s(p)$, to be written in a form that contains the same reference point as that of the fluid, $\widetilde{\mu}_f(p)$:

$$\widetilde{\mu}_s(p) - \widetilde{\mu}_f(p_o) = b_\infty - \frac{1}{N}\ln\left[\frac{\Delta_m(p_o, b_\infty)}{\Delta_m(p_o, 0)} \cdot \frac{\Delta_s(p)}{\Delta_s(p_o)}\right] . \tag{2.21}$$

Again, the partition functions that appear in Eq. (2.21) are evaluated through histogram reweighting. The chemical potentials of the fluid and the solid, given by Eqs. (2.18) and (2.21), can now be used to define fluid-solid coexistence at temperature T by locating the common intersection of the two μ–p curves. In the present work, the previous thermodynamic integration technique has been used to determine fluid-solid coexistence for $N = 108$, 256, and 500 Lennard-Jones particles at $T^* = 1$. The reference pressure was taken to be $p_o^* = 1$. For $N = 864$, 1372, and 2048 particles, fluid-solid coexistence was determined by simulating the field-induced phase transition of the modified cell model using tempering techniques.

2.5.1 Finite-size scaling analysis of fluid-solid transitions

In the present section, the fluid-solid coexistence data for finite systems of Lennard-Jones particles at $T^* = 1$, obtained according to the procedures described in the previous section, are analyzed according to finite-size scaling techniques for first-order phase transitions. The methodology used here closely follows that of the previous section which is based on analyzing the size-dependent behavior of second- and higher-order derivatives (suscepti-

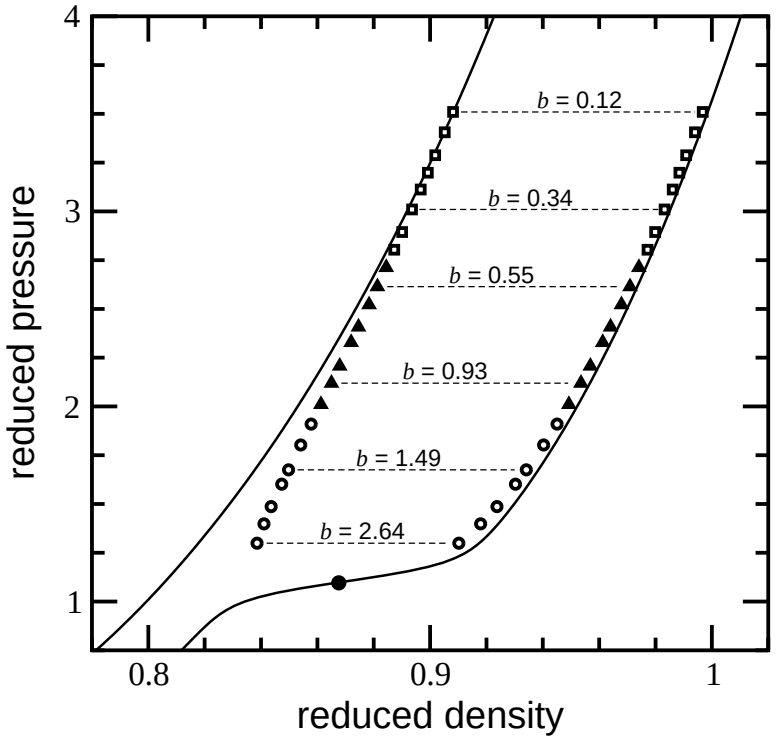


Figure 2.25: Phase diagram of the modified cell model for the Lennard-Jones system at $T^* = 1$ in the pressure-density plane. The solid lines correspond to the fluid and the constrained solid phases [i.e., curves (f) and (s) in Fig. 2.23]. The coexisting solid and fluid phases for finite values of b are shown as points and they correspond to system sizes, N, of ($\circ$): 2048, ($\blacktriangle$): 1372, and ($\square$): 864. The dashed horizontal lines are tie-lines and they are drawn for clarity. The point shown as bullet, ($\bullet$), corresponds to the inflection point of the $T^* = 1$ isotherm of the constrained cell model.

bilities) of the Gibbs free energy as well as of dimensionless moment ratios. As previously defined, the second- and third-order susceptibilities are defined by Eqs. 2.8 and 2.11. As the transition pressure is approached either from high- or low-p, χ_2 diverges to $+\infty$. For finite systems, however, χ_2 only attains an N-dependent maximum. The size-dependent value of the pressure associated with that maximum, provides a coexistence pressure estimator. The third-order susceptibility, χ_3, which is the slope of χ_2, diverges to $+\infty$ ($-\infty$) as the transition pressure is approached from low (high) values. The size-dependent shape of χ_3 thus contains two extrema points, a low-p maximum and a high-p minimum, respectively. The size-dependent values of the pressures associated with the extrema of χ_3 provide two additional coexistence pressure estimators. Higher-order susceptibilities have not been considered since, as the order increases, the precision with which the corresponding pressures can be located, decreases.

According to finite-size scaling theory for first-order phase transitions [9, 8, 6, 7], the pressure, $p^{(N)}$, associated with an extremum point, has the following scaling behavior in terms of N

$$p^{(N)} - p^{(\infty)} \sim \frac{1}{N}, \tag{2.22}$$

where $p^{(\infty)}$ is the coexistence pressure in the limit of infinite size. The previous scaling form, obtained by assuming the probability distribution of the order parameter is a superposition of two gaussian distributions, ignores corrections-to-scaling terms and the effects of intermediate, mixed-phase configurations. The relevant scaling plot for the coexistence pressure is shown in Fig. 2.26. The three pressure estimators shown in Fig. 2.26 corre-

spond to N-dependent pressures obtained through the equal-area criterion (which is almost indistinguishable from the pressures associated with the maximum of χ_2) as well as the pressures associated with the two extrema of χ_3. As is evident in Fig. 2.26, linearity with respect to the scaling variable $X = 1/N$ is only observed for $N \geq 500$. The irregular behavior for small N (= 108 and 256) is due to the effect of tail (i.e., long-range) corrections in the energy calculation. Since the effect of tail corrections becomes progressively less important as the system size increases, only data associated with moderate and large systems (e.g., $N \geq 500$) should be considered in the analysis. As Fig. 2.26 indicates, the N-dependent pressures for $N \geq 500$ conform to the linear fit with satisfactory precision and yield $p^* = 3.937 \pm 0.005$ in the limit of infinite size.

In the context of the determination of the coexisting densities, we will once again consider the use of dimensionless moment ratios [112, 11, 24, 10] of the order parameter. These ratios have been remarkably successful in simulation studies of second-order criticality in asymmetric fluids [64]. The most widely used ratio, is the so-called 4th-order ratio, Q_4, which, in the context of the present work is defined by Eq. 2.14. As a first-order transition is approached from both sides at fixed T, Q_4 jumps discontinuously from $1/3$ to zero [64]. In the two-phase region, Q_4 varies between zero and one and attains the maximum value of unity at a density which is the arithmetic mean of the densities of the two coexisting phases [64]. The relevant plot for Q_4 is found in Fig. 2.27. As Fig. 2.27 indicates, for finite systems, Q_4 develops two size-dependent minima separated by a size-dependent maximum. As before the following 6th-order ratio is considered by Eq. 2.15 which exhibits a similar

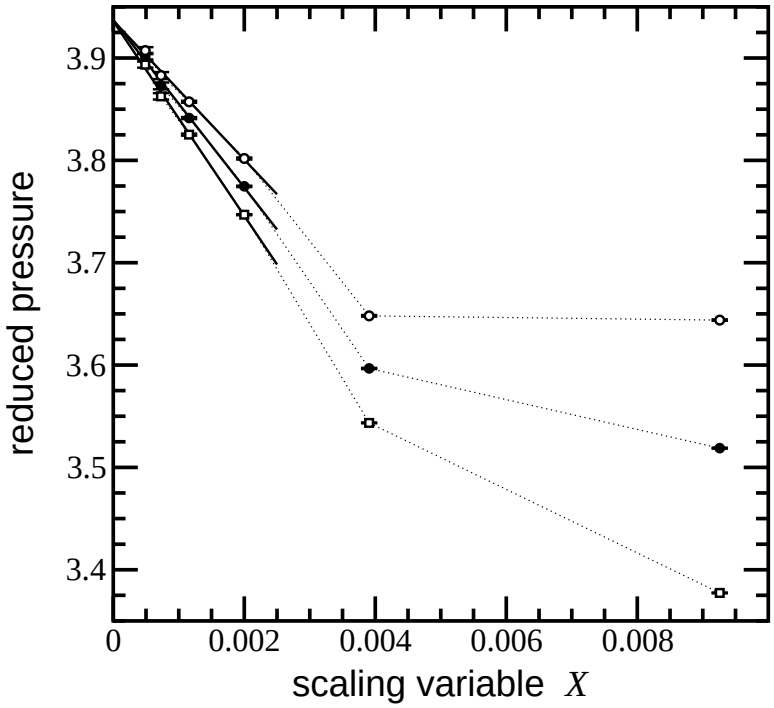


Figure 2.26: Determination of the fluid-solid coexistence pressure for the Lennard-Jones system at $T^* = 1$. The N-dependent coexistence pressures are plotted vs the scaling variable $X = N^{-1}$. The filled circles, (•), correspond to the coexistence pressures obtained from the equal-area criterion. The open circles, (○), and squares, (□), correspond to the pressures at the extrema of the 3rd-order susceptibility. The solid lines are linear fits. The points that correspond to $N = 108$ and 256 particles are excluded from the fit. Dotted lines are drawn for visual clarity.

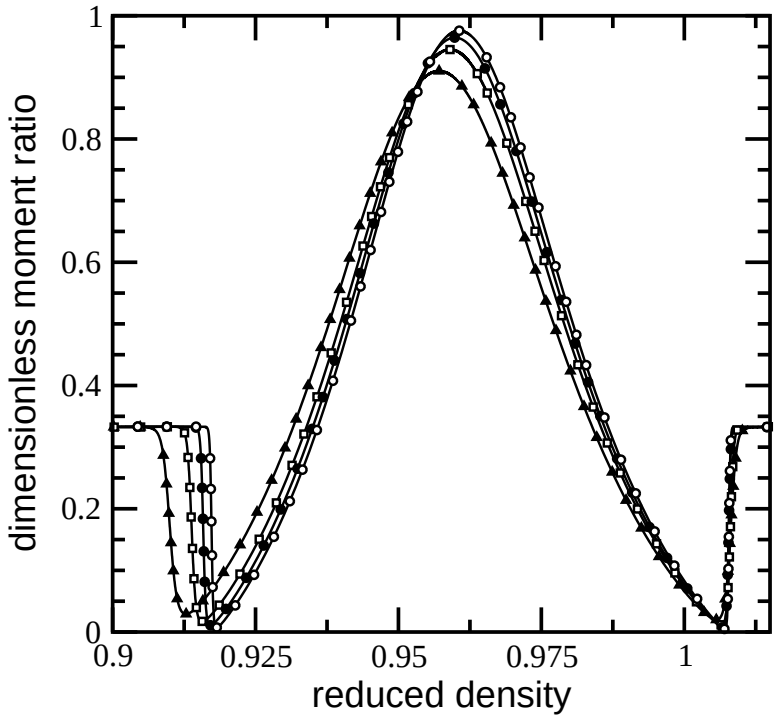


Figure 2.27: Dimensionless moment ratio, Q_4, vs average reduced density, $\langle \rho^* \rangle$, for the Lennard-Jones system at $T^* = 1$. The curves correspond to number of particles (•): $N = 500$, (□): $N = 864$, (▲): $N = 1372$, and (○): $N = 2048$.

size-dependent behavior as that of Q_4. The coexisting densities of the bulk system can be obtained by extrapolating the N-dependent densities associated with the minima of Q_4 and Q_6 towards $N \to \infty$ according to the scaling ansatz, Eq. (2.22).

The scaling plots for the coexisting densities are shown in Figs. 2.28 (fluid) and 2.29 (solid). The open symbols correspond to the densities associated with the minima of Q_4 and Q_6. The filled circles correspond to densities obtained from the first-moments of the volume distributions of the fluid and the solid phase at a pressure obtained through the

equal-area construction (or the maximum of χ_2). The behavior of the densities in Figs. 2.28 and 2.29 is analogous to that of the pressures in Fig. 2.26. As in the case of the pressure estimation, only the points associated with $N \geq 500$ are considered in the scaling analysis. However, the densities associated with the minima of Q_4 and Q_6 extrapolate to a density that is slightly different than that obtained from the first-moments of the distribution function. As also explained in previously, this happens because Q_4 and Q_6 require the evaluation of higher-order moments, see Eqs. (2.14) and (2.15), which are not as accurate as low-order moments. The best estimators for the coexisting densities can thus be obtained from the first-moments of the distribution function, i.e., the filled circles in Figs. 2.28 and 2.29. The difference between this estimate and that obtained from the minima of Q_4 and Q_6 is indicative of the precision that can be attained by this analysis and can also be considered as providing a measure of the uncertainty of the estimate. The data associated with the first moments thus give $\rho^*_{\text{fluid}} = 0.9201 \pm 0.0007$ and $\rho^*_{\text{fcc}} = 1.0072 \pm 0.0008$. As emphasized by previously, better precision requires simulations of systems much larger than $N = 2048$.

The freezing transition of the Lennard-Jones model has been investigated via thermodynamic [52, 120] and Gibbs-Duhem [1] type of integration techniques as well as with the phase switch method [29, 78] . Hansen and Verlet [64] obtained fluid-solid coexistence of a system of $N = 864$ Lennard-Jones particles at $T^* = 2.74$, 1.15, and 0.75, using thermodynamic integration. Since both the fluid and the solid are nearly incompressible, linear interpolation between their reported densities at $T^* = 0.75$ and 1.15 gives $\rho^*_{\text{fluid}} \cong 0.913$ and $\rho^*_{\text{fcc}} \cong 1.005$ at $T^* = 1$. Agrawal and Kofke [1] used Gibbs-Duhem integration to obtain

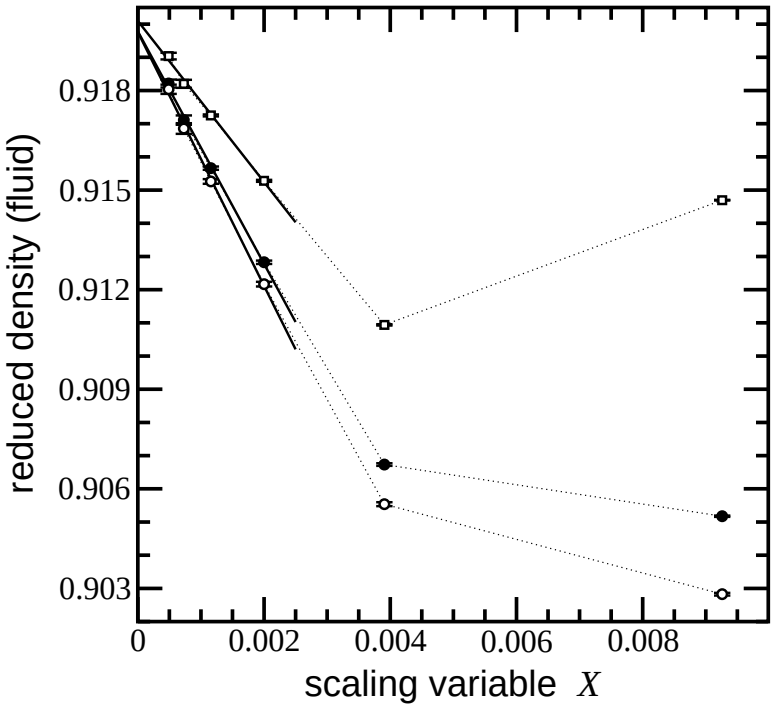


Figure 2.28: Determination of the coexistence density of the fluid phase for the Lennard-Jones system at $T^* = 1$. The N-dependent densities are plotted vs the scaling variable $X = N^{-1}$. The squares, (□), correspond to densities obtained from the equal-weight criterion. The filled, (•), and open, (○), circles correspond to densities obtained from the dimensionless ratios Q_4 and Q_6. The solid lines are linear fits. The points associated with $N = 108$ and 256 particles are excluded from the fit. Dotted lines are drawn for visual clarity.

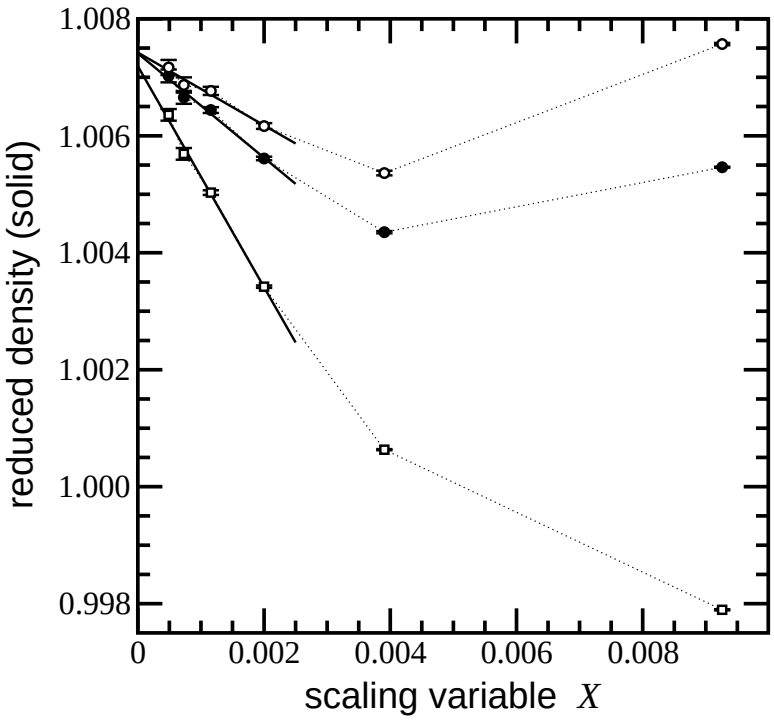


Figure 2.29: Determination of the coexistence density of the solid phase for the Lennard-Jones system at $T^* = 1$. The labeling is the same as that of Fig. 2.28.

the fluid-solid transition of $N = 500$ Lennard-Jones particles for a range of temperatures. Interpolation between the entries at $1/T^* = 0.943$ and $1/T^* = 1.057$ gives $\rho^*_{\text{fluid}} \cong 0.925$ and $\rho^*_{\text{fcc}} \cong 1.011$ for the coexisting densities at $T^* = 1$. The coexistence pressure cannot be determined from such simple interpolation schemes because it varies rapidly with the temperature. The phase switch method is implemented under constant pressure and uses multicanonical sampling and a series of "smart" moves to facilitate transitions between the two coexisting phases. Implementation of the phase switch method in the freezing transition of the Lennard-Jones model gave [78] $p^* = 3.546$, $\rho^*_{\text{fluid}} \cong 0.9098$ and $\rho^*_{\text{fcc}} \cong 0.999$ for $N = 256$ particles at $T^* = 1$. All these values, obtained on finite systems are in reasonable accord with the results of this work for $N \leqslant 864$. Most previous simulations have not implemented a thorough analysis of system-size effects. Figs. 2.26, 2.28, and 2.29, however, demonstrate the importance of accounting for size effects in simulation studies of fluid-solid transitions.

2.6 Conclusions

The main advantage of the modified cell model in simulation studies of fluid-solid transitions is its simplicity. The extra effort (computational as well as programming) associated with the introduction of the external field variable into the ordinary isothermal-isobaric ensemble is minor. Unlike other methods that are based on complicated order parameters, the conjugate variable associated with the external field can be simply and transparently identified as the number of Wigner-Seitz cells containing a single particle. In the context

of simulations of fluid-solid equilibria via thermodynamic integration, the modified cell model provides a path that connects the fluid with the solid phase, thus reducing the number of simulated states. Alternatively, the fluid-solid transition of the system of interest can be obtained by analyzing the field-induced phase transition of the corresponding modified cell model by standard flat-histogram techniques. Furthermore, a crucial element stemming from this section is the importance of accounting for and analyzing size effects in simulation studies of fluid-solid transitions. No comparable scaling analysis comprising more than a single pressure and density estimator currently exists in the literature.

Chapter 3

Phase transitions, criticality, and three-phase coexistence using constrained cell models

3.1 Introduction

In Chapter 2, the constrained and generalized cell models were introduced, and then used to find the coexistence properties for systems of hard spheres and Lennard-Jones particles. In this chapter, the phase diagram of the constrained cell model for a system of Lennard-Jones particles is determined from Monte Carlo simulations. Understanding of the phase diagram of such a system is essential in devising simulation algorithms for direct determination of sublimation and triple-point coexistence. The simulations are implemented under

constant pressure using tempering [77, 76] and histogram reweighting [31, 32] techniques. The coexisting phases are determined by analyzing the structure of the volume distribution which was obtained as a histogram of observations from the simulations. The topology of the phase diagram of the constrained system is similar to that of the unconstrained system. Specifically, the constrained system is found to contain a critical and a triple point just as the unconstrained system. The values of the temperature and pressure at the critical point are higher than those of the unconstrained system.

3.2 Phase diagram of the constrained cell model

Once again we consider a system of Lennard-Jones particles acting according to Eq. 2.12 with $n = 6$. In Fig. 3.1, four pressure–density isotherms for the solid phase (as modeled through the constrained cell model) for a system of $N = 256$ Lennard-Jones particles are shown. These isotherms have been obtained from constant-pressure simulations using tempering techniques [77, 76]. The main output of these simulations comprises the joint distribution (histogram) of the volume and energy. These distributions were analyzed according to histogram reweighting [31, 32] techniques. The details of these simulations have been discussed in the previous chapter. As is evident in Fig. 3.1, at very low densities and pressures, the system behaves as an ideal gas, i.e., $p^* \cong \rho$. At high pressures the system is nearly incompressible and, hence, the p–ρ isotherms are almost vertical. As the pressure decreases from high values, the four isotherms shown in Fig. 3.1 develop inflection points, shown as open circles in the figure, at densities that are 60–70% of the density at close

packing. The inset of Fig. 3.1 shows the resulting volume distributions at the inflection points. The two-peak shape of these distributions resembles that of a system undergoing some type of phase transition. The specific value of the pressure was determined through histogram reweighting to yield a volume distribution for which the two peaks have the same height. In constant-volume simulations of hard- and soft-sphere and Lennard-Jones systems [55, 56, 52, 57], the inflection points shown in Fig. 3.1 appeared in the form of a kink or cusp or an angular point. As was first emphasized by Hoover and Ree [56], the kink (or the inflection) can be thought of as the limit of mechanical stability of the solid phase. At densities that are higher than the density associated with the kink/inflection, the solid phase can survive for substantial time intervals without the confinement imposed by the walls of the Wigner-Seitz cells. In contrast, at lower densities the presence of the cell walls is necessary to prevent the solid phase from rapidly disintegrating to a disordered fluid phase.

As the pressure is further reduced, the shape of the p–ρ isotherms shown in Fig. 3.1 indicates that at temperatures $T^* < 2$ the solid phase (i.e., the constrained cell model) undergoes a phase transition between an expanded and a compressed phase. This phase transition is the analogue of the gas-liquid transition of the unconstrained system. In the neighborhood of the transition, the shape of the volume distribution is bimodal (i.e., two maxima separated by a minimum, see also Fig. 3.2 below) and averaging over this distribution gives the nearly horizontal segments shown in Fig. 3.1. The coexisting phases are of the same symmetry and, hence, this transition terminates at a second-order critical point as is evident

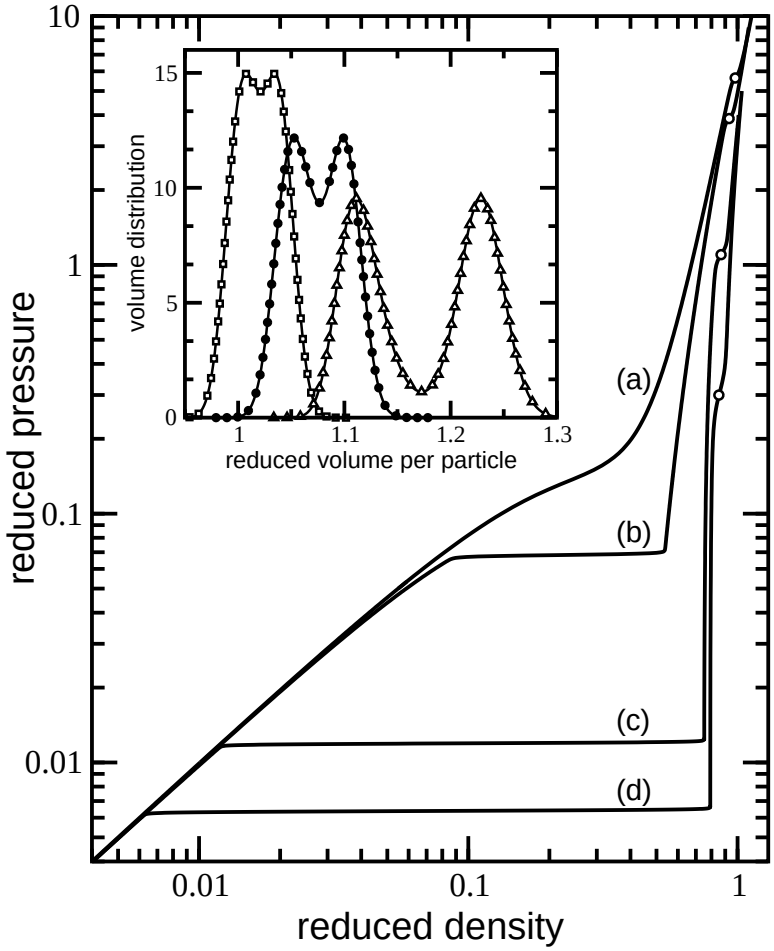


Figure 3.1: Reduced pressure, p^*, vs average reduced density, $\langle \rho^* \rangle$, for the solid phase (i.e., the constrained cell model) of a system of $N = 256$ Lennard-Jones particles. The four p–ρ isotherms shown in the figure correspond to reduced temperatures (a): $T^* = 2$, (b): $T^* = 1.5$, (c): $T^* = 1$, and (d): $T^* = 0.9$. The open circles ($\circ$) are estimates of the limits of mechanical stability. The inset shows the distribution of the reduced volume at the mechanical stability points, obtained from constant–p simulations. The three volume distributions shown in the inset correspond to temperatures ($\square$): $T^* = 2$, ($\bullet$): $T^* = 1.5$, and ($\triangle$): $T^* = 0.9$. The volume distribution associated with the $T^* = 1$ p–ρ isotherm [i.e., curve (c) in the main figure] has been omitted for clarity.

in Fig. 3.1. On physical grounds, such a behavior should be expected since the potential energy function contains repulsive as well as attractive terms. A behavior similar to that shown in Fig. 3.1 was also detected by Hansen and Verlet [52] for the same system. In that case, however, the p–ρ isotherms exhibited van der Waals type of loops because the simulations were implemented under constant volume conditions.

The resulting phase diagram for the constrained cell model for a system of $N = 256$ Lennard-Jones particles in the T–ρ plane is shown in Fig. 3.2. The phase coexistence between the expanded and the compressed solid phases (i.e., the nearly-horizontal segments of the p–ρ isotherms in Fig. 3.1) is shown as open circles in Fig. 3.2. The inset of Fig. 3.2 shows the resulting volume distribution at such coexistence at $T^* = 1.62$. For a given temperature, the coexisting pressure and densities were determined by the equal-area construction [12]. As already emphasized, this first-order phase transition is the analogue of the gas-liquid transition of the unconstrained system and, since the coexisting densities are of the same symmetry, it terminates at a second-order critical point which is shown as a filled square in Fig. 3.2. The critical values of the temperature and density ($T_c^* \cong 1.70$, $\rho_c^* \cong 0.287$) have been estimated by fitting the coexisting densities to the scaling law for the order parameter and the law of the rectilinear diameters assuming Ising criticality.

As already commented, at the limit of mechanical stability, the shape of the resulting volume distribution resembles that of a system undergoing a phase transition, see inset of Fig. 3.1 for instance. Molecular dynamics simulations on hard spheres suggest that this transition is a weak first-order transition [128]. The relative height of the minima of

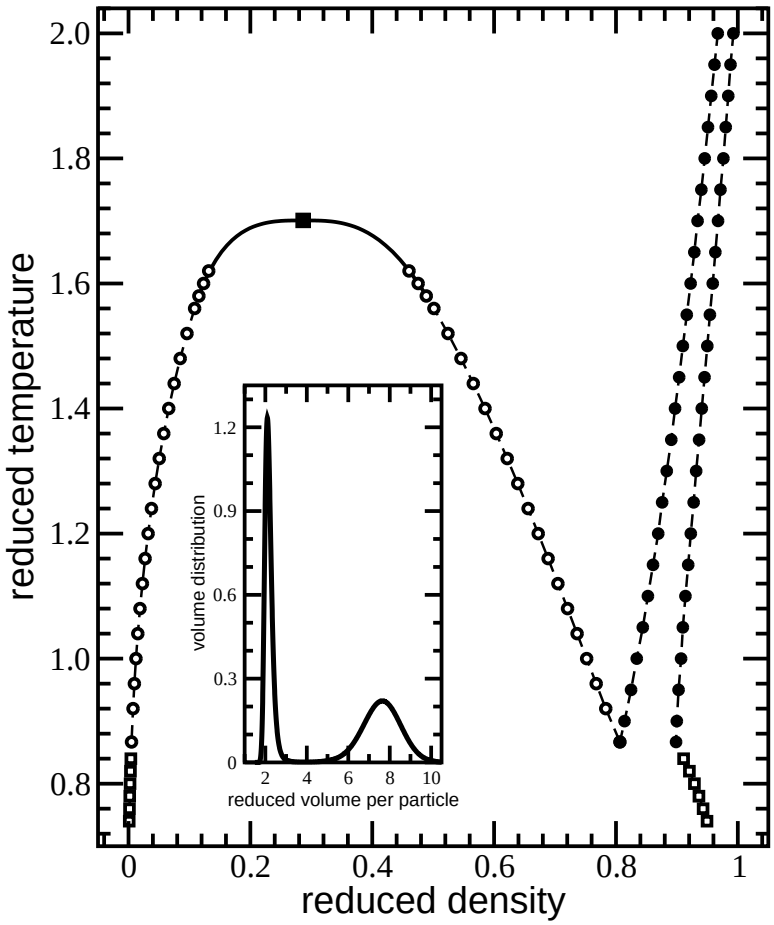


Figure 3.2: Phase diagram of the constrained cell model for a system of $N = 256$ Lennard-Jones particles in the temperature-density plane. The filled circles, (•), have been obtained from the positions of the maxima of the volume distributions at the limit of mechanical stability (see also inset of Fig. 3.1). The open circles (○), correspond to coexistence (obtained through the equal area construction) between an expanded and a compressed solid phase. The inset shows the volume distributions at such coexistence at $T^* = 1.62$. This coexistence terminates at a critical point shown as a filled square, (■). The solid line is a fit to scaling laws. The filled and open circles intersect at $T^* \cong 0.867$. Phase coexistence for $T^* < 0.867$, shown as open squares (□), resembles sublimation equilibria in unconstrained systems.

the volume distributions shown in the inset of Fig. 3.1, compared to the maxima, seem to support this assertion. Determination of the order of this transition requires scaling analysis and computational effort that involves system sizes much greater than $N = 256$. Assuming that a phase transition does indeed occur at the limit of mechanical stability, a rough estimate for the coexisting densities at a given T may be obtained by locating the positions of the maxima of the volume distributions shown in the inset of Fig. 3.1. The results of this procedure, commonly referred to as the equal-height construction [12], are shown as filled circles in Fig. 3.2. The form of this coexistence is reminiscent of liquid-solid coexistence in an unconstrained system. However, there is a crucial difference. Liquid-solid coexistence is associated with phases of different symmetry (liquid vs solid) and, hence, there is no critical point. In contrast, the coexistence implied by the filled circles in Fig. 3.2 involves phases of the same symmetry and, hence, the appearance of a critical point as the temperature increases cannot be precluded. The existence of such criticality is related to the behavior of the solid around the mechanical stability point as seen through the shape of the volume distribution. This question merits its own separate investigation via scaling analysis for finite systems.

The p–T phase boundary associated with gas-like and liquid-like coexistence (i.e., the open circles in Fig. 3.2) intersects the p–T stability line (i.e., the filled circles in Fig. 3.2) at $T^* \cong 0.867$. The volume distribution in the vicinity of the intersection of the two p–T curves contains three peaks separated by two minima. Such three-peaked distributions are shown in Figs. 3.3 and 3.4. This case is associated with three-phase coexistence and the

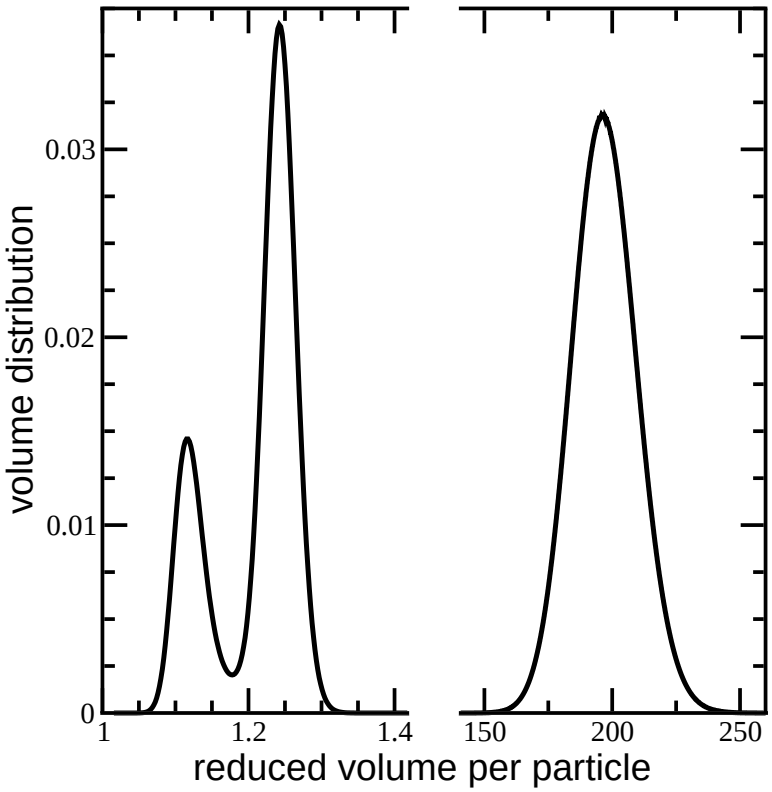


Figure 3.3: Volume distribution of the constrained cell model for a system of $N = 256$ Lennard-Jones particles at $T^* = 0.87$ and $p^* = 0.005$.

intersection at $T^* \cong 0.867$ plays the role of a triple point. For $T^* < 0.867$ the middle phase becomes metastable (see Fig. 3.4, for instance) and the resulting two-phase coexistence is reminiscent of sublimation in an unconstrained system.

Despite the fact that the phase diagram of the constrained cell model, shown in Fig. 3.2, resembles that of the unconstrained system, there are a few crucial differences. First, the critical temperature ($T_c^* \cong 1.7$) and the triple-point temperature ($T_{tr}^* \cong 0.867$) are about 25-30% higher than those of the unconstrained system [100, 1, 18] ($T_c^* \cong 1.3$, $T_{tr}^* \cong 0.69$)

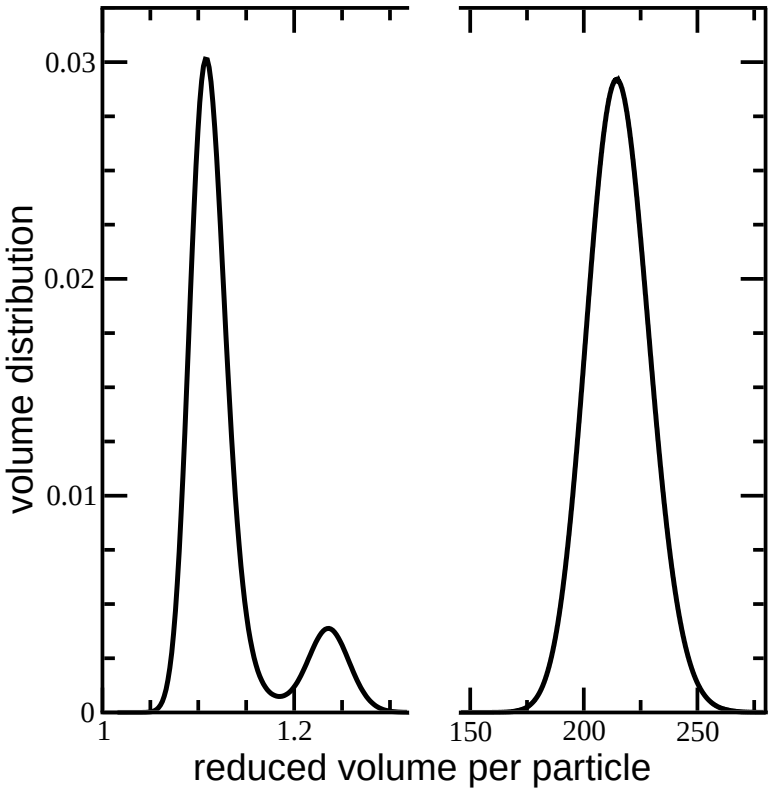


Figure 3.4: Volume distribution of the constrained cell model for a system of $N = 256$ Lennard-Jones particles at $T^* = 0.86$ and $p^* = 0.0046$.

for the Lennard-Jones model. The main reason for this behavior is the decrease in the entropy of the system caused by the single occupancy constraint. Further, since all coexisting phases are of the same symmetry it is possible to observe two- and three-phase coexistence from a single simulation. Such a task is considerably more difficult in the unconstrained system.

3.3 Conclusion

The constrained cell model is a limiting case of a more general cell model. Such a generalized or modified cell model, first proposed by Hoover and Ree [55], was used in the previous chapter to devise simulation techniques appropriate for liquid-solid transitions under constant pressure. The modification is associated with the introduction of a homogenous external field variable that interacts with the Wigner-Seitz cells. High values of the external field force single occupancy configurations and thus stabilize the solid phase. We showed that this modified cell model can be used to link the fluid with the solid phase by progressively increasing the strength of the field. This technique has hitherto been used for liquid-solid coexistence studies. Of particular interest is the question of whether the modified cell model can be used in the direct determination of gas-solid and gas-liquid-solid coexistence which will be presented in the following two Chapters. Direct determination of gas-solid and gas-liquid-solid coexistence is considerably more challenging than liquid-solid coexistence because the volume range that must be covered in a single simulation is large. Furthermore, an added complication has to do with the range of temperatures

and pressures for which the phase transitions of the constrained cell model, shown in Figs. 3.2-3.4, occur. The first step in this direction is associated with understanding the phase diagram of the constrained cell model, which was provided in this work. Further work comprises the construction of generalized cell models and appropriate sampling techniques capable of linking the coexisting phases of the constrained cell model shown in Figs. 3.2-3.4 with those of the unconstrained system.

Chapter 4

Direct determination of triple-point coexistence through cell model simulation

4.1 Introduction

In the previous chapters, a series of simple and efficient direct and indirect simulation techniques for accurate simulation of fluid-solid coexistence is proposed. These approaches are based on extensions or generalizations of the constrained cell model [55, 56] description of the solid phase. Specifically, the generalization entails the introduction of a homogeneous, uniform external field variable that interacts with the Wigner-Seitz cells and thus controls the relative stability of the solid against the fluid phase. High values of the field variable

force configurations with one particle per cell and, hence, favor the solid phase. Thus, by varying the strength of the field at fixed density or pressure, it is possible to devise a path that connects the solid with the fluid phase.

Specifically, in Chapter 2, these simulation techniques were used to calculate fluid-solid coexistence points at a fixed temperature, and in Chapter 3, the phase boundary was determined for the constrained cell model through a thermodynamic integration technique based on histogram reweighting [31, 32]. For the cases for which the phase diagram of the system of interest contained a triple point, the corresponding triple-point temperature and pressure was estimated from the common intersection of the various saturation curves (i.e., evaporation and melting). In this chapter, the information from the previous two chapters are used along the generalized cell model in order to determine coexistence at the triple-point.

4.2 The generalized cell model

In the present work, the phase diagram of the generalized cell model is investigated through multicanonical simulations [4] at constant pressure and histogram reweighting techniques [31, 32] for a system of 256 Lennard-Jones particles. At a given temperature and pressure, the resulting type of coexistence is deduced by analyzing the structure of the volume distribution (histogram). The phase diagram resembles that of the unconstrained and the constrained system. The simulation data is used to obtain an estimate of the triple-point

of the Lennard-Jones system. The triple-point parameters are obtained via a procedure analogous to the equal-area construction [12] for two-phase coexistence.

In the isothermal-isobaric ensemble, the number of particles N, the pressure p, and the temperature T [or the reciprocal temperature $\beta = 1/(k_B T)$], are fixed. The energy and the volume V, both vary and thus the reduced density $\rho = N/V$, and the volume per particle $v = 1/\rho = V/N$, fluctuate about mean values, $\langle \rho \rangle$ and $\langle v \rangle$ respectively. If $\varepsilon > 0$ is the depth of the attractive part of the pair potential, appropriate reduced variables may be defined as: $T^* = k_B T/\varepsilon$ (reduced temperature), $p^* = \beta p \sigma^3 = p\sigma^3/k_B T$ (reduced pressure), $\langle \rho^* \rangle = \langle \rho \rangle \sigma^3$ (average reduced density), and $\langle v^* \rangle = \langle v \rangle / \sigma^3$ (average reduced volume per particle).

In simulations involving solid phases, the constrained cell model is formed [55, 56] by dividing the volume V, into N Wigner-Seitz cells, each of volume V/N. Simulation techniques based on cell models cannot find the equilibrium solid phase of the system under consideration. Hence, the equilibrium solid phase must be known *a priori* or found by some other means. In the present work, it is assumed that the stable solid phase is of the face-centered cubic (fcc) type and, hence, the cells are rhombic dodecahedra. For a cubic simulation box, the number of particles, N, satisfies $N = 4k^3$ ($k \geqslant 2$). Each cell contains a single particle and particle displacements that violate the single-occupancy constraint are not allowed. The fluid phase (vapor and liquid) is modeled as an unconstrained system for which there is no restriction on the number of particles per cell. The generalized cell model is defined by introducing a homogenous, uniform external variable, B say, that interacts

with the N cells. The field variable B should be roughly thought of as the magnetic field in Ising spin-type of lattice models. Suppose again that the volume is partitioned into N Wigner-Seitz cells but consider arbitrary assignments of the N particles into the N cells. The occupation status (i.e., number of particles) of cell j ($j = 1, 2, \ldots, N$) can be described in terms of the spin-type of variable s_j that only takes two values: -1 if cell j contains one particle and zero otherwise. By analogy with a magnetic spin-type of system, suppose that the interaction of the field B with cell j can be described with the term Bs_j. An order parameter M, conjugate to B, can thus be defined as $M = \sum_j s_j$. Note that $|M|$ (or $-M$) is simply the number of singly occupied cells of a given N-particle configuration. If $b = \beta B$, the partition function for a system of N particles at T, p, and b is given by

$$\Delta_m(N, p, T, b) = e^{-\widetilde{\mu}_m N} = \sum_i e^{-\beta E_i - \widetilde{p} V_i - b M_i}, \tag{4.1}$$

where index i enumerates the states and E_i and V_i are the energy and volume in state i. Please note that this is also listed as Eq. 2.3, but has been given again for clarity. In Eq. (4.1), $|M_i|$ is the number of singly occupied cells in state i, $\widetilde{p} = \beta p$ and $\widetilde{\mu}_m = \beta \mu_m$, where μ_m is the chemical potential at p, T, and b. The elementary steps comprise particle displacements and volume changes that are accepted with standard Metropolis type of criteria [37]. As is evident from Eq. (4.1), the unconstrained system is recovered in the limit in which $b \to 0$. Furthermore, only single occupancy configurations are possible for large values of b. Hence, the constrained cell model is recovered in the limit in which $b \to \infty$. From the previous arguments it thus follows that the generalized cell model can be used to connect the fluid with the solid phase, since both of these phases, fluid (unconstrained) and

solid (constrained), are limiting cases of this generalized cell model.

4.3 Phase diagram of the generalized cell model

The generalized cell model has been simulated for a system of $N = 256$ Lennard-Jones particles. The potential energy of a configuration for a system of N particles in a periodic simulation box of volume $V = L^3$, is evaluated by calculating all pair terms for distances $r \leqslant L/2$ and adding a tail correction [37] for $r > L/2$. The simulations were implemented according to the multicanonical preweighting scheme [4] and the simulation output comprised three-dimensional histograms of energy, volume V, and order parameter M. The behavior of the Lennard-Jones system in terms of the field b was inferred by analyzing the structure of the volume distribution (obtained through histogram reweighting) and evaluating its moments. The length of each simulation was $10\text{–}100 \times 10^6$ sweeps (N displacements and N volume changes) depending on the range covered and the possibility of phase transition behavior within that range. Statistical uncertainties on the coexistence properties were calculated as two standard deviations of ten block averages.

The phase behavior of the Lennard-Jones model for $b = 0$ (unconstrained system) and $b \gg 0$ (constrained cell model) has been discussed in Chapters 2 and 3. In particular, the phase diagram of the constrained cell model for a system of Lennard-Jones particles is similar to that of the unconstrained system. Specifically, for $b \gg 0$ a critical and a triple-point is found at temperatures and pressures that are higher than those of the unconstrained

system ($b = 0$). The behavior of the system for intermediate values of b is shown in Fig. 4.1. In that figure, the pressure-density isotherm at $T^* = 1.2$ is shown for four values of b, namely, $b = 0$, 0.6, 1.5, and 20. Although it is possible to construct p–ρ curves for any value of b within the simulated range ($0 \leqslant b \leqslant 20$) using histogram reweighting, only four p–ρ isotherms are shown in Fig. 4.1 for reasons of clarity. At large values of b (i.e., $b \geqslant 10$ to 20) only single occupancy configurations are possible and, hence, the large-b behavior of the generalized cell model is indistinguishable from that of the constrained cell model. The nearly horizontal segments of the p–ρ curves indicate that there is an evaporation / condensation phase transition at moderate densities followed by a melting / freezing transition at higher densities. In the vicinity of these transitions, the corresponding volume distribution (histogram) consists of two peaks separated by a minimum, see Fig. 4.2 for instance.

Regarding the phase transitions shown in Fig. 4.1, coexistence at temperature T is defined by finding (through histogram reweighting) the value of the pressure for which the resulting volume distribution consists of two peaks of equal areas. The previous task is commonly referred to as the equal-area construction [12]. The coexisting densities are defined as the first moments of the two peaks. The volume distribution at gas-liquid and liquid-solid coexistence for the unconstrained system ($b = 0$) are shown in Fig. 4.2(a) and (b). Phase diagrams may be obtained by applying the equal-area rule for a range of temperatures. The resulting evaporation / condensation phase diagram in the temperature-density plane is shown in the inset of Fig. 4.1. This phase transition terminates at a critical

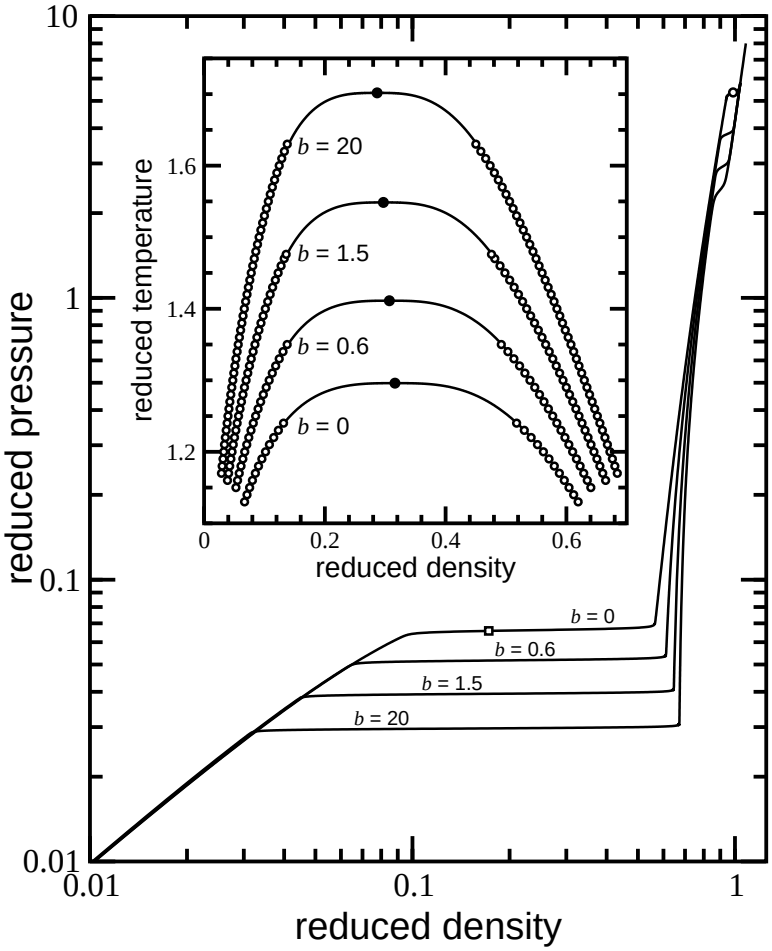


Figure 4.1: Reduced pressure, p^*, vs average reduced density, $\langle \rho^* \rangle$, for a system of $N = 256$ Lennard-Jones particles at $T^* = 1.2$. In the main figure, the p–ρ curves correspond to values of b (from top to bottom) of: 0, 0.6, 1.5, and 20. The nearly-horizontal segments of the p–ρ curves indicate phase transition behavior. For $b = 0$ (unconstrained system), phase coexistence is marked with an open circle (○) and square (□), respectively, and the resulting volume distributions are shown in Fig. 4.2. The phase transition at intermediate densities is an evaporation / condensation type of transition that terminates at a critical point. The inset shows the resulting gas-liquid type of coexistence in the T–ρ plane for $b = 0$, 0.6, 1.5, and 20.

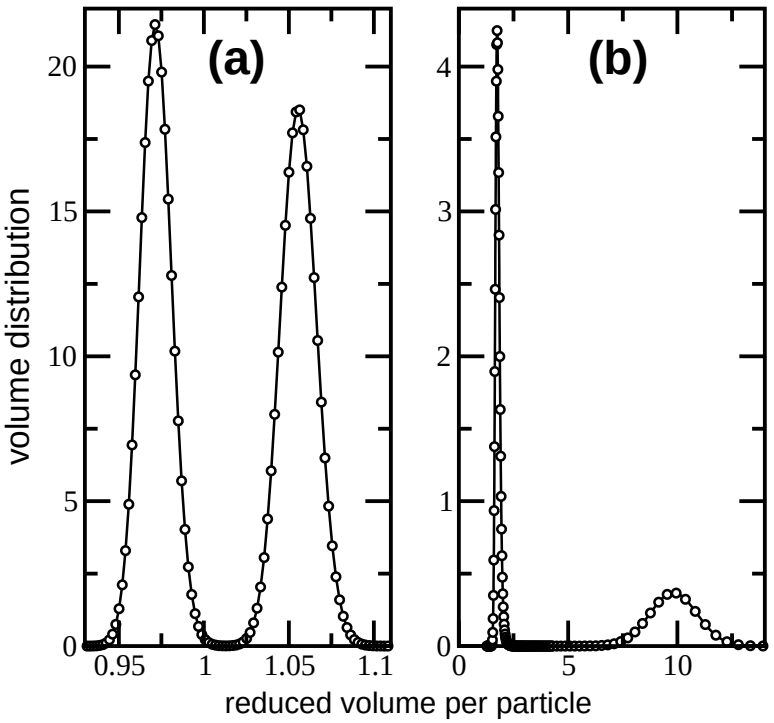


Figure 4.2: Volume distribution at the (a) melting ($p^* = 5.360$) and (b) condensation ($p^* = 0.0658$) transition for $N = 256$ Lennard-Jones particles at $T^* = 1.2$, obtained according to the procedure described in the text. The value of the pressure was calculated through the equal-area rule using histogram reweighting. These distributions correspond to points shown as open circle (○) and open square (□) in the $b = 0$, p–ρ isotherm in Fig. 4.1.

point which is marked with a filled circle in the inset of Fig. 4.1. The critical point is found by fitting the coexistence data (shown as open circles in the inset of Fig. 4.1) to the scaling law for the order parameter (assuming Ising criticality) and the law of rectilinear diameters. As the value of the field increases, the corresponding value of the critical temperature increases as well. The primary reason for the increase in the critical temperature is the reduction of the entropy caused by the increased strength of the field b that forces particle localization within the cells.

The high-density melting / freezing transition shown in Fig. 4.1 occurs at higher pressures as the strength of the field is reduced from high values towards zero. In the context of the constrained cell model (i.e., $b \gg 0$), this phase transition may be interpreted as the mechanical stability limit [56] of the solid phase (as modeled through the constrained cell model). Specifically, at pressures that are higher than the transition pressure, an ordered, solid-like configuration can exist for substantial time periods without the confinement imposed by the walls of the cells or by the external field. In contrast, at low pressures, a solid configuration cannot survive without the confining effect of the cell surfaces or of the field b and it rapidly disintegrates to a disordered fluid phase. As the value of the confining field b decreases, the magnitude of the confining effect decreases too, and, hence, the mechanical stability point (or, alternatively, the melting transition point) must occur at higher pressures and densities.

4.4 Three phase coexistence

The simulation data can be used to provide an estimate of the triple point of the Lennard-Jones system. The location of the triple point is found through a construction that is analogous to the equal-area rule for two-phase coexistence. Specifically, the simulation data indicates that for temperatures $0.68 \lesssim T^* \lesssim 0.72$, there is a range of pressures for which the corresponding volume distribution consists of three peaks separated by two minima. These peaks correspond to the three phases, namely, gas, liquid, and solid. The areas of the three peaks are calculated for a given T and p and the process is repeated for different T and p values in order to locate the unique point for which all three areas are equal. This generalized equal-area procedure gave $T^*_{\rm tr} = 0.7085 \pm 0.0005$, $p^*_{\rm tr} = (2.264 \pm 0.017) \times 10^{-3}$ for $N = 256$ particles. The resulting volume distribution at $T^*_{\rm tr}$ and $p^*_{\rm tr}$ is shown in Fig. 4.3. The first moment under each peak defines the density of the corresponding phase at the triple point. The coexisting densities are found to be $\rho^*_{\rm gas} = (2.298 \pm 0.018) \times 10^{-3}$, $\rho^*_{\rm liq} = 0.8405 \pm 0.0003$, and $\rho^*_{\rm fcc} = 0.9587 \pm 0.0002$.

As anticipated, the gas phase at the triple point is almost ideal since $p^*_{\rm tr} \cong \rho^*_{\rm gas}$. Furthermore, the estimated uncertainty of the pressure and of the gas density ($\sim 0.8\%$) is an order of magnitude higher than that of the temperature ($\sim 0.07\%$) and of the densities of the condensed phases ($\sim 0.03\%$). The origin of the discrepancy in these uncertainties may be understood by considering the μ–p curves of the three phases. The results of the equal-area construction at $T^* = 0.7085$ (i.e., the best estimate for $T_{\rm tr}$) are shown in Fig. 4.4. The area under each peak of the volume distribution (see, e.g., Fig. 4.3), which is proportional

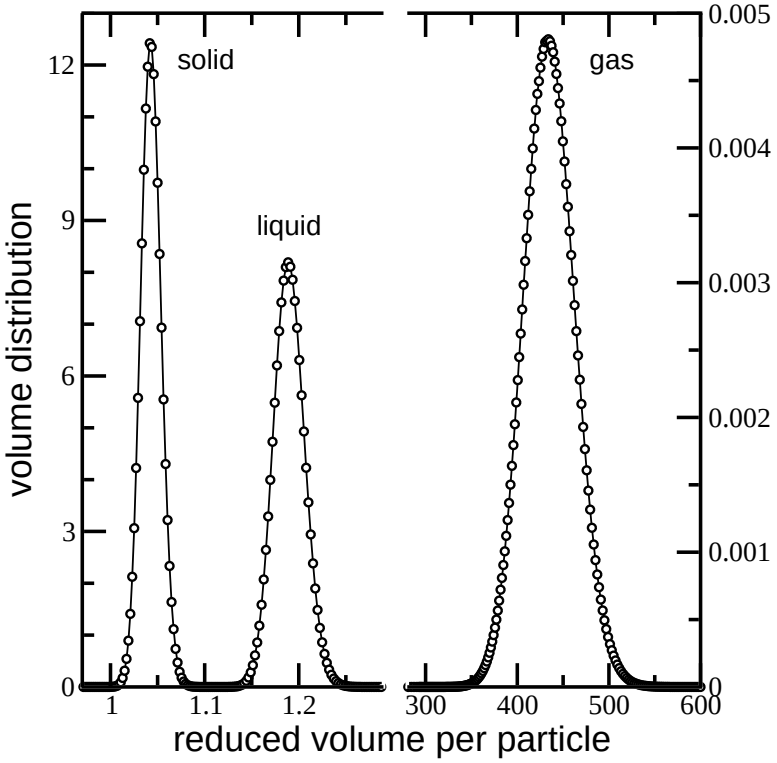


Figure 4.3: Volume distribution for a system of $N = 256$ Lennard-Jones particles at the triple point ($T^*_{\rm tr} = 0.7085$, $p^*_{\rm tr} = 2.264 \times 10^{-3}$). The values of T and p were found according to the equal-area rule as explained in the text.

to the partition function Δ, has been converted to chemical potential using $\widetilde{\mu} = -N^{-1} \cdot \ln \Delta$, cf. Eq. (4.1). In order to eliminate the proportionality constant, the chemical potential of the liquid at $p^* = 2 \times 10^{-3}$ was set to zero. Since both the liquid and the solid phase are nearly incompressible the corresponding μ–p curves are almost straight lines. Hence, a small uncertainty in the pressure and the temperature causes a much smaller uncertainty in the densities of the condensed phases. The μ–p curve of the gas is almost vertical, reflecting the very compressible nature of that phase. Consequently, a small uncertainty in the temperature, causes a larger uncertainty in the common intersection of the three μ–p curves (which defines the triple-point pressure) and of the gas density.

The estimate for the triple-point temperature is higher than that of previous work ($T_{\mathrm{tr}}^* =$ 0.67 to 0.69), obtained using thermodynamic and Gibbs-Duhem integration and Gibbs ensemble techniques [52, 71, 1, 18]. The likely sources for that discrepancy might be related to different system sizes used and to differences in the energy evaluation. This discrepancy can be resolved though a finite-size scaling analysis (see, e.g., refs. 1-5) which is, however, a computationally demanding task. The current estimates of the triple-point pressure and of the density of the gas phase are both higher by a factor of $\sim$ 15% to 30% compared to previous work. The main reasons for the higher values of p_{tr} and ρ_{gas}, compared to previous work, are the higher value of T_{tr} and the very compressible nature of the dilute phase. On the other hand, the present estimates for densities of the condensed phases are in reasonable accord with those of previous work ($\rho_{\mathrm{liq}}^* = 0.818$–$0.854$ and $\rho_{\mathrm{fcc}}^* = 0.955$–$0.963$), a fact that may be attributed to the nearly incompressible nature of the liquid and the solid

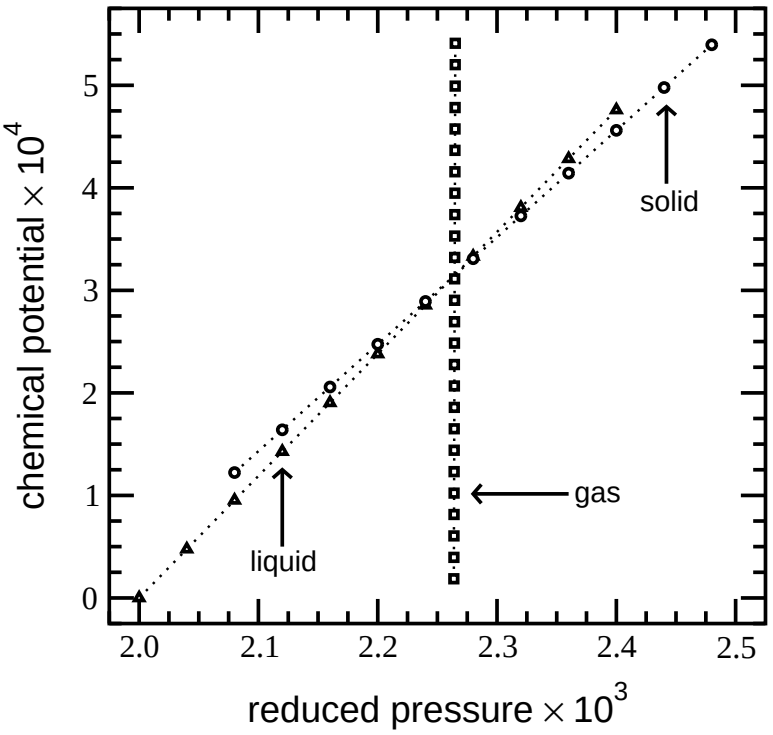


Figure 4.4: Chemical potential vs pressure for $N = 256$ Lennard-Jones particles at $T^* = 0.7085$. (□): gas; (△): liquid; (○): solid. The chemical potential of the liquid at $p^* = 2 \times 10^{-3}$ is assigned a value of zero. The common intersection of the three μ–p curves defines the triple point.

phase.

4.5 Conclusion

The partition function of an unconstrained system (i.e., $b = 0$) contains all possible configurations, both fluid and solid, at a given T and p. In simulations, however, isothermal compression of a low-ρ fluid phase often leads to formation of a non-equilibrium glassy type of structure rather than the equilibrium solid phase. A model for the solid phase (i.e., the constrained cell model) may be formed by considering only configurations with one particle per Wigner-Seitz cell (i.e, $b \gg 0$) and excluding all others. Since the relevant portions of the phase space associated with the disordered, fluid phase and the ordered, solid phase are distinctly different from each other, it is not obvious as to how one can construct a path that connects the fluid (unconstrained system, $b = 0$) with the solid (constrained cell model, $b \gg 0$). The present chapter (as well as previous chapters) in this dissertation demonstrates that such a connection can be achieved through the generalized cell model which provides a simple way to simulate liquid-solid (cf. Fig. 4.2), gas-solid, and gas-liquid-solid (cf. Fig. 4.3) coexistence. In the present Chapter, it was shown that this model can be used to provide direct estimates of triple-point equilibria. In previous work, the triple point was estimated by indirect methods, i.e., by locating the common intersection of the various p–T curves (such as condensation and melting) through extrapolations. Despite these advances, there are still a few crucial questions that must be addressed. One such crucial issue is the size dependence of the triple-point parameters. As already commented, size-dependence issues

can formally be resolved through finite-size scaling analysis for first-order phase transitions [98, 85, 86]. This task is, however, computationally demanding due to the substantial density range (from the low-ρ ideal-gas to the high-ρ incompressible regime) that must be sampled in a single simulation. Regarding system-size effects, previous work [86, 84, 87] associated with finite-size scaling analysis of liquid-solid coexistence at fixed T^* using cell models indicates that the estimate of the pressure for $N = 256$ differs by 4 to 6% from the $N = \infty$ value. These differences are expected to be somewhat larger for the triple-point parameters due to the highly compressible nature of the gas phase.

Chapter 5

Simulation of phase boundaries using constrained cell models

5.1 Introduction

In the previous chapter, the triple-point was determined for a system of Lennard-Jones particles interacting according to Eq. (2.12) with $n = 6$. The present chapter focuses on the determination of the phase diagram of a system of particles also interacting according to Eq. (2.12), however with a value of $n = 12$. Previous work [53, 125] indicates that the vapor-liquid transition becomes metastable against solidification for $n \gtrsim 10$ to 11. The vapor-liquid phase diagram is readily established from constant-pressure simulations and flat-histogram techniques [77, 76]. The fluid-solid phase diagram is also constructed from constant-pressure simulations utilizing constrained and generalized cell models. Specif-

ically, fluid-solid coexistence on the isotherm that corresponds to a reduced temperature of 2 is determined by analyzing the field-induced phase transition of the generalized cell model. The previous fluid-solid coexistence point is used as the reference point in the determination of the entire fluid-solid phase boundary through a thermodynamic integration type of technique based on histogram reweighting [31, 32]. As anticipated [53, 125], the vapor-liquid transition is metastable against crystallization. The phase diagram of the corresponding constrained cell model is also determined. The simulation results indicate that the phase diagram of the constrained cell model contains a stable critical point as well as a triple point.

5.2 Phase diagram of the constrained cell model

The present section is associated with the behavior of the constrained cell model and, in particular, with the possibility of phase transitions. In Fig. 5.1, three representative pressure-density isotherms of this model, obtained through constant-pressure simulations of a system of $N = 256$ particles, are shown. These isotherms correspond to reduced temperatures $T^* = 2$, 0.75, and 0.70. As is evident in Fig. 5.1, at low densities and pressures the system is nearly ideal and, hence, $p^* \cong \rho^*$. At high pressures and densities, the system is nearly incompressible and these p–ρ isotherms are, thus, nearly vertical. Furthermore, these p–ρ isotherms contain horizontal segments that are indicative of phase transition behavior. These different types of phase transitions are explained below.

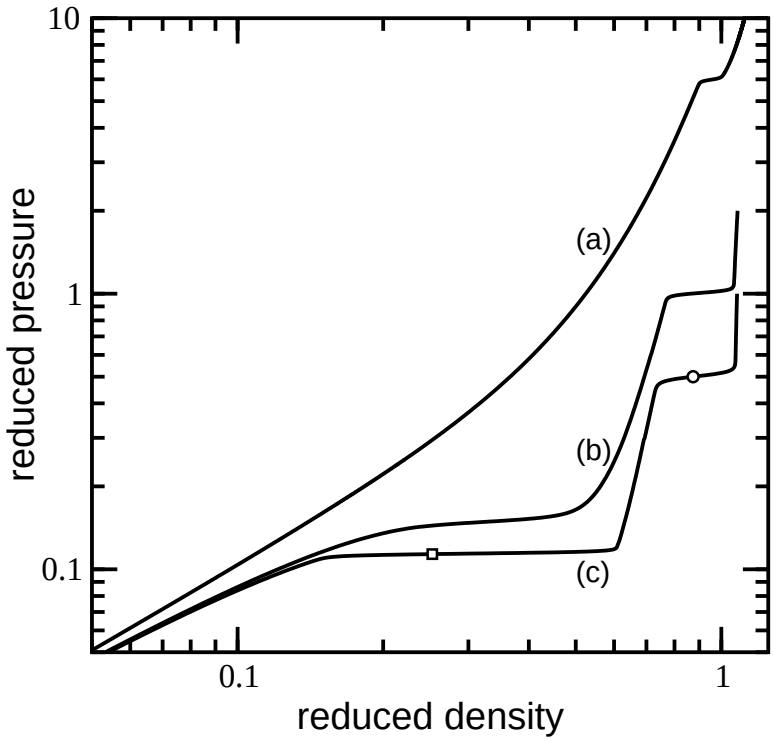


Figure 5.1: Reduced pressure, p^*, vs average reduced density, $\langle \rho^* \rangle$, for the solid phase (i.e., the constrained cell model) of a system of $N = 256$ particles. The three p–ρ isotherms shown in the figure correspond to reduced temperatures (a): $T^* = 2$, (b): $T^* = 0.75$, and (c): $T^* = 0.70$.

As the pressure is reduced from high values, there is a first-order phase transition between a compressed and an expanded solid phase. The volume distribution at such coexistence for a system of $N = 256$ particles at $T^* = 0.70$ is shown in Fig. 5.2(a). The specific value of the pressure ($p^* = 0.5$) in Fig. 5.2(a), found through histogram reweighting, yields a volume distribution with two peaks of equal area [12]. The coexistence point is marked with an open circle in the $T^* = 0.70$ p–ρ isotherm in Fig. 5.1. Note that the averaging is implemented over the whole distribution in Fig. 5.2(a) rather than the individual peaks. Averaging over the latter would yield a p–ρ isotherm with hysteresis loops. This phase transition is associated with the limit of mechanical stability of the solid. Such an anomaly was first detected by Hoover and Ree for the hard-sphere solid [56]. In that case, the anomaly appeared in the form of a kink or a cusp because the simulations were implemented at constant volume. Molecular dynamics simulations [128] for the hard-sphere solid suggest that this anomaly or instability is a weak first-order transition. Constant-pressure simulations for hard- and soft-spheres as well as for Lennard-Jones systems, obtained in the previous sections, support the weak, first-order transition scenario, although the anomaly appeared in the form of an inflection rather than a kink. For the present molecular model, the horizontal portions of the p–ρ isotherms in Fig. 5.1 and the shape of the volume distribution in Fig. 5.2(a) suggest a strong first-order transition. As first explained by Hoover and Ree [56], for densities that are higher than the densities associated with this transition, the solid phase (as modeled through the constrained cell model) can exist for substantial time periods without the presence of the cell walls. In contrast, at lower densities, the solid cannot

exist without the confinement imposed by the walls of the Wigner-Seitz cells and it rapidly melts to a disordered, fluid-like phase.

As the pressure is further reduced, the shape of the p–ρ isotherms associated with low temperatures ($T^* \lesssim 0.75$ in Fig. 5.1), indicates that there is another phase transition between a high- and a low-density solid phase. The volume distribution at coexistence for a system of $N = 256$ particles at $T^* = 0.70$, obtained through the equal-area rule [12], is shown in Fig. 5.2(b). The coexistence point is marked with an open square on the $T^* = 0.70$ p–ρ isotherm in Fig. 5.1. This phase transition is the analogue of the gas-liquid transition of the unconstrained system (i.e., the fluid phase) and it is caused by the attractive interactions of the pair potential, cf. Eq. (2.12). A behavior similar to that shown in Fig. 5.1 had also been seen by Hansen and Verlet [52] and in the previous sections for the Lennard-Jones solid. Inspection of the isotherms in Fig. 5.1 indicates that this phase transition terminates at a critical point at $T^* \cong 0.75$.

The previous observations regarding the phase transitions of the solid phase (as modeled through the constrained cell model) are summarized in the form of a T–ρ phase diagram in Fig. 5.3. The gas-like, liquid-like phase transition is shown as open circles ($N = 500$) and filled triangles ($N = 256$). As already commented, the volume distribution is bimodal; see Fig. 5.2(b), for instance. For a given temperature, the coexistence pressure was found via the equal-area rule [12]. The coexisting densities were obtained as the first moments of each peak. As is evident in Fig. 5.3, this phase transition terminates at a critical point, shown as a filled square. The critical values of the temperature, T_c, and of the density,

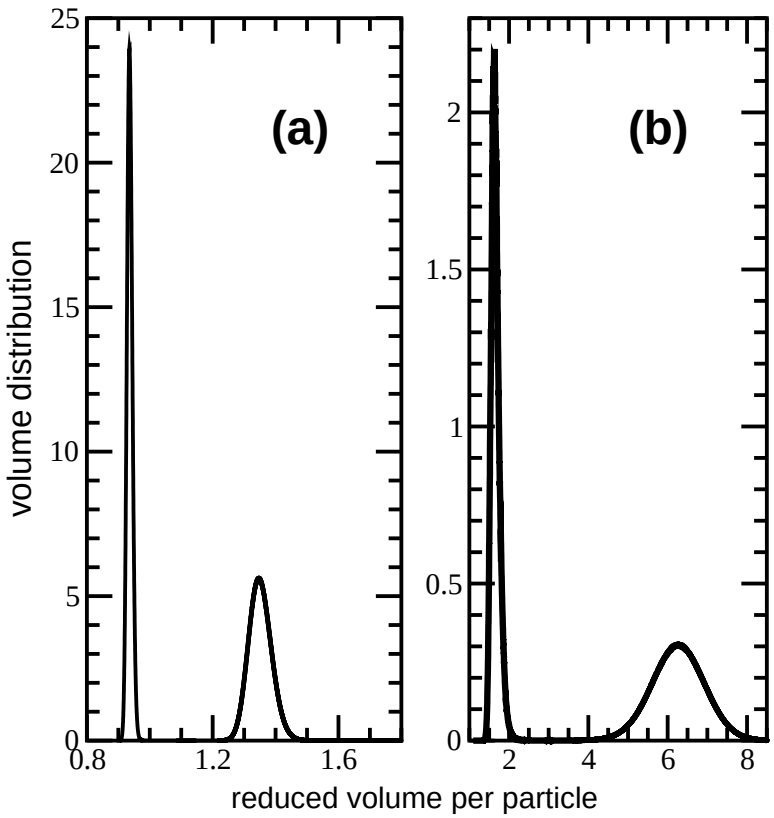


Figure 5.2: Volume distribution of the solid phase for $N = 256$ particles at $T^* = 0.70$ and (a): $p^* = 0.5$, (b): $p^* = 0.1135$. The value of the pressure has been obtained through the equal-area criterion. These distributions correspond to points shown as open circle ($\circ$) and open square ($\square$) of the $T^* = 0.70$ isotherm in Fig. 5.1.

ρ_c, were obtained by fitting the coexistence data for $N = 500$ to the leading power law for the density difference and the law of the rectilinear diameters, i.e., Eqs. 2.16 and 2.17. In Eqs. (2.16) and (2.17) it has again been assumed that $\beta = 0.326$, appropriate for the Ising universality class [73]. Application of this approach gives $T_c^* \cong 0.743$ and $\rho_c^* \cong 0.368$. In the same figure, the coexisting densities of the phase transition associated with the mechanical stability region are shown as filled circles. The form of this coexistence resembles that of fluid-solid coexistence in an ordinary system with one crucial difference, however. Ordinary liquid-solid coexistence involves phases of different symmetry (liquid and solid) and, hence, there is no critical point. In contrast, the coexistence associated with the filled circles in Fig. 5.3 involves phases of the same symmetry (expanded and compressed fcc structures) and, thus, the appearance of a critical point at a sufficiently high temperature cannot be excluded. As pointed out previously, this question merits its own separate investigation. These two phase boundaries intersect at $T^* \cong 0.661$. Phase coexistence at temperatures $T^* < 0.661$ involves a highly expanded and a highly compacted solid phase. As is evident in Fig. 5.3, this type of coexistence is reminiscent of sublimation in an ordinary system and $T^* = 0.661$ is, thus, the triple-point temperature. The phase diagram of the constrained cell model, shown in Fig. 5.3 is very similar to that of an ordinary system.

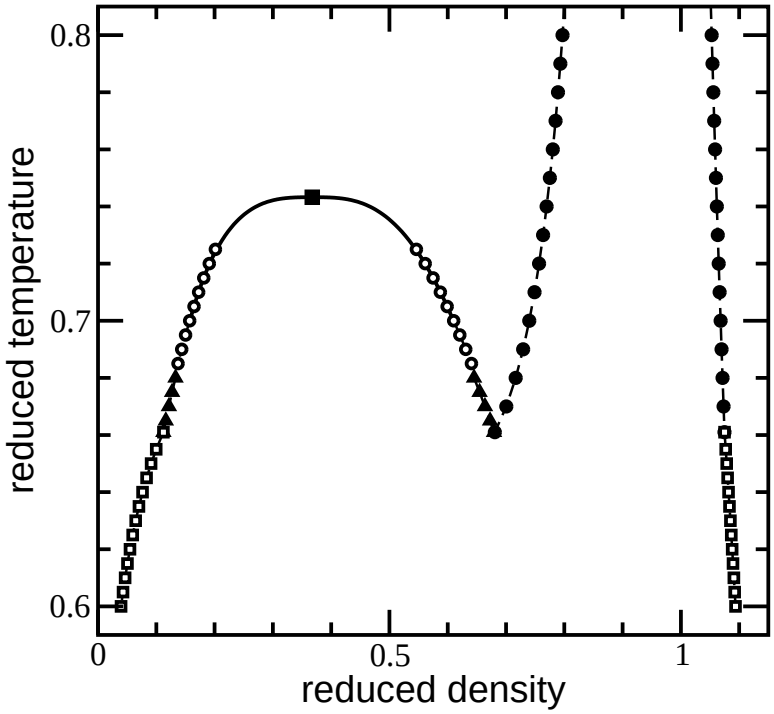


Figure 5.3: Phase diagram of the constrained cell model in the temperature-density plane. The filled circles, (•), correspond to the phase transition associated with the loss of mechanical stability. The filled triangles, (▲), and open circles, (○), correspond to the phase transition between a dilute and a dense solid phase. This transition terminates at a critical point which is shown as a filled square, (■). The solid line is a fit to the scaling laws for the density difference and the coexistence curve diameter; see Eqs. (2.16) and (2.17). These two phase transitions intersect at a triple point at $T^* \cong 0.661$. Phase coexistence for $T^* < 0.661$, shown as open squares (□), is reminiscent of sublimation equilibria in ordinary systems. The simulations have been performed for $N = 256$ particles with the exception of the points shown as open circles, (○), that have been obtained for $N = 500$ particles.

5.3 Determination of fluid-solid coexistence using generalized cell models

The present section is associated with the physical behavior of the generalized (or modified) cell model and its use in simulations of fluid-solid coexistence. As already emphasized previously, the partition function of this model, cf. Eq. (2.3), contains an external field, b, that controls the number of particles per Wigner-Seitz cell and, hence, the relative stability of the solid against the fluid phase. The unconstrained system, cf. Eq. (2.1), corresponds to $b = 0$, whereas the constrained cell model, cf. Eq. (2.2), is recovered in the limit in which $b \to \infty$. Hence by varying the field, b, from high values towards zero on a constant-pressure (or density) path, it is possible to link the solid (constrained cell model) with the fluid (unconstrained system) phase. In the following lines, the behavior of this model is analyzed on the $T^* = 2$ isotherm.

The effect of the external field variable b, on different constant-pressure paths for $N = 256$ particles at $T^* = 2$ is shown in Fig. 5.4. In that figure, the variation of the average fraction of the cells that contain a single particle, $-\langle M \rangle / N$, is plotted in terms of the field variable, b, for five isobars that correspond to the pressures $p^* = 6$, 6.5, 7, 7.5, and 8, respectively. As is evident in the figure, at high-field values, single occupancy configurations (i.e., $|M| = N$) dominate and the system is in the solid phase. As the field decreases from high values towards zero, the nature of the transformation of the solid to the fluid depends on the particular value of the pressure. Specifically, at low and moderate values of the pres-

sure, the passage from the solid to the fluid as b decreases is smooth and continuous, see, e.g., the $p^* = 6$ isobar in Fig. 5.4. At higher pressures ($p^* \gtrsim 6.5$ in Fig. 5.4), the shape of the isobars resembles that of a system undergoing a phase transition. The simulation results indicate that for $p^* \gtrsim 6.5$ there is a range of b values for which the distribution of volume, V, and of the number of singly occupied cells, $|M|$, both have a two-peak structure separated by a minimum. The inset of Fig. 5.4 contains the resulting volume distribution at $b = 0.358$ and $p^* = 7.5$. The specific value of b was found through histogram reweighting to yield two peaks of the same area. The low-volume peak, centered at $v^* \cong 0.94$, corresponds to an ordered phase with one particle per Wigner-Seitz cell. The high-volume peak, centered at $v^* \cong 1.07$, corresponds to a disordered phase for which the average fraction of singly occupied cells is $\sim 78\%$. Thus, at high pressures the transformation of the high-b, ordered phase, to the low-b, disordered phase as b decreases occurs discontinuously, i.e., through a phase transition. The value of b at this order-disorder transition of the generalized cell model depends on the pressure. As is evident from Fig. 5.4, as p^* increases, the value of b at this transition decreases.

The field-induced phase transition of the generalized cell model at $T^* = 2$, implied by the shape of the isobars in Fig. 5.4, has been studied by constant-pressure simulations for $N = 256$, 500, and 2048 particles. The simulation results for the coexisting densities are shown in Fig. 5.5 together with the relevant portion of the p–ρ isotherms of the fluid [unconstrained system, cf. Eq. (2.1)] and the solid [constrained cell model, cf. Eq. (2.2)]. The values of the coexisting densities of the ordered and the disordered phase of the generalized

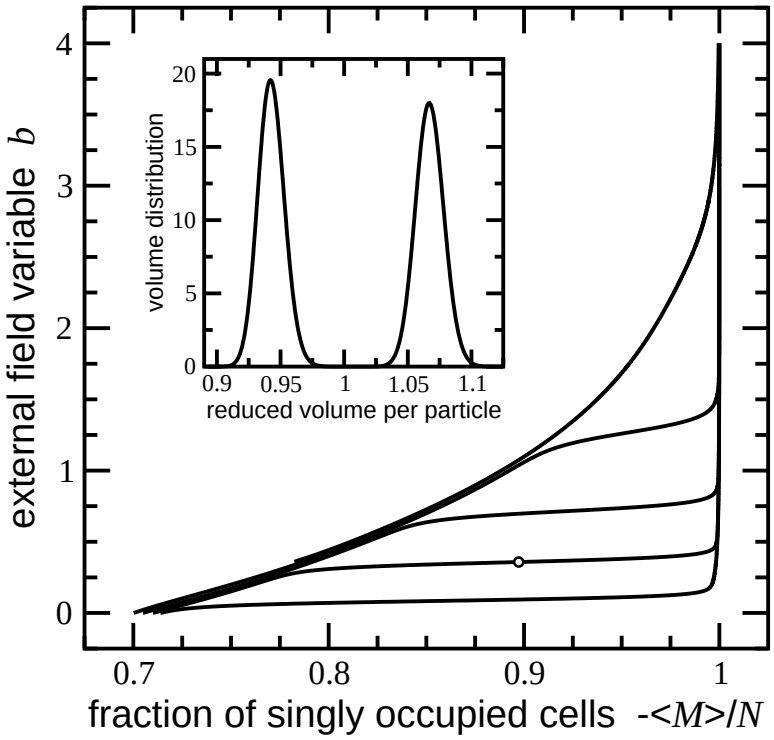


Figure 5.4: Variation of the external field variable, b with the average fraction of singly occupied cells, $-\langle M\rangle/N$, for $N = 256$ particles at $T^* = 2$. The curves from top to bottom correspond to reduced pressures $p^* = 6$, 6.5, 7, 7.5, and 8. The inset shows the volume distribution at $b = 0.358$ and $p^* = 7.5$. This state corresponds to the open circle, ($\circ$), on the $p^* = 7.5$ isobar in the main figure.

cell model lie in between the isotherms of the solid and the fluid phase. As the pressure increases, the value of the field, b at order-disorder coexistence decreases and the coexisting densities approach those of the fluid and the solid phase in the high-p, $b \to 0$ limit, as is evident in Fig. 5.5. As the pressure decreases, the value of the field, b at coexistence increases, and the coexisting densities of the ordered and the disordered phase appear to be merging towards a critical point in the neighborhood of the mechanical stability region of the solid phase. Detailed elucidation of such criticality requires scaling analysis for finite systems. Since the generalized cell model is an artificial model used for the sole purpose of connecting the fluid with the solid phase, such a scaling analysis was not implemented because it is of no practical value.

The previous findings suggest two equivalent ways to establish fluid-solid coexistence of the system of interest at temperature T using the generalized cell model. Referring to Fig. 5.5, fluid-solid coexistence may be established by analyzing the field-induced order-disorder transition of the generalized cell model in the high-p, $b \to 0$ limit, using standard flat-histogram techniques. Alternatively, a procedure that resembles thermodynamic integration may be employed. Specifically, both phases, fluid (unconstrained system) and solid (constrained cell model), are simulated for a series of pressures, starting from a reference pressure p_0. Then, the generalized cell model is used to link the free energies (i.e., chemical potentials) of the two phases. This task is achieved by performing a series of simulations on the constant-pressure path $p = p_0$ for progressively increasing values of the field b. The details of both methods have been documented in the previous chapters and they will

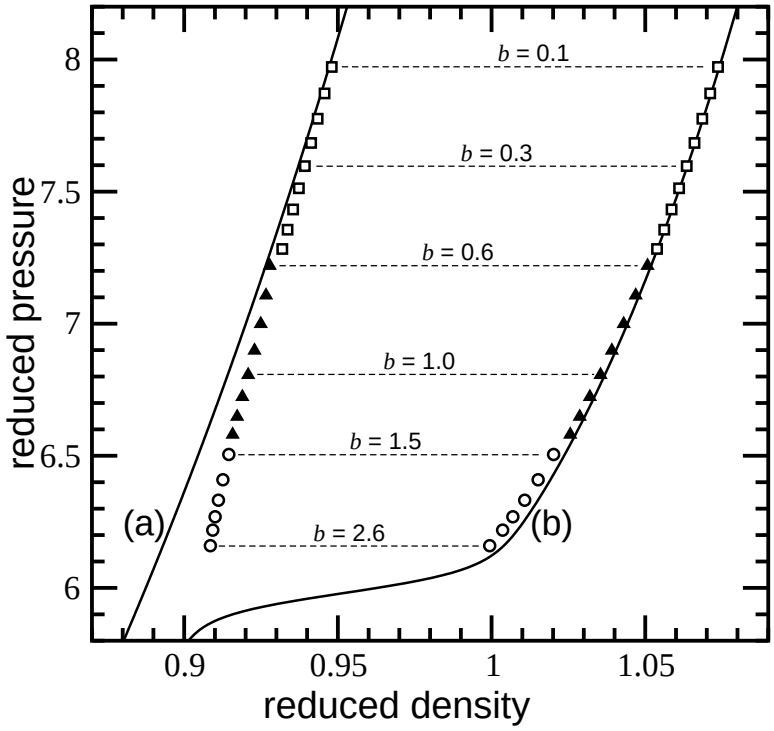


Figure 5.5: Phase diagram of the generalized cell model at $T^* = 2$ in the pressure-density plane. The solid lines correspond to the p–ρ isotherms of the fluid and the solid phase at $T^* = 2$ and $N = 256$. (a): fluid (unconstrained system); (b): solid (constrained cell model). The coexisting densities of the ordered and the disordered phases of the generalized cell model correspond to system sizes, N, of (○) 2048, (▲) 500, and (□) 256. The dashed horizontal lines are tie-lines and they are drawn for clarity.

not be repeated here. The volume distribution at fluid-solid (fcc) coexistence, obtained by analyzing the field-induced phase transition of the generalized cell model for $N = 256$ particles, is shown in Fig. 5.6. The estimate for the coexistence pressure, $p^* = 8.221 \pm 0.002$, was obtained through the equal-area rule [12]. The first moments of the peaks of the distribution shown in Fig. 5.6 gave $\rho^*_{\text{fcc}} = 1.0804 \pm 0.0004$ and $\rho^*_{\text{fluid}} = 0.9535 \pm 0.0003$, respectively. The reported uncertainties correspond to two standard deviations of ten block averages. Thermodynamic integration using a reference pressure $p^*_o = 5$ gave results within the quoted uncertainties.

5.4 Tracing fluid-solid phase boundaries

In the previous section, fluid-solid coexistence for a system of particles interacting with the pair potential given by Eq. (2.12), was established at a single temperature ($T^* = 2$). For cases for which the determination of the entire fluid-solid phase boundary is desired, the methodology of the previous section is still applicable provided that one stores the data in terms of three-dimensional histograms of the volume, the order parameter M, and the energy. The main drawbacks of such a task are increased memory and storage requirements. In the present section, a simple thermodynamic integration type of technique [95] is used to trace the entire fluid-solid phase boundary using the coexistence point at $T^* = 2$ as the reference point.

The most popular simulation method for tracing phase boundaries is the so-called

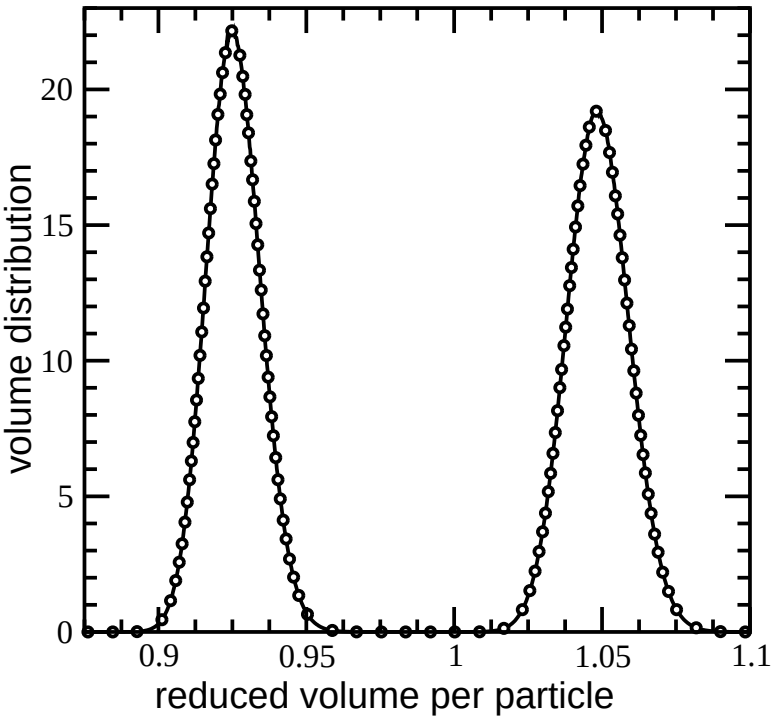


Figure 5.6: Distribution of reduced volume per particle, v^*, at fluid-solid (fcc) coexistence for $N = 256$ particles at $T^* = 2$. The value of the coexistence pressure ($p^* = 8.221$) was estimated through the equal-area rule.

Gibbs-Duhem integration method [66, 65]. According to this method, the two phases (i.e., fluid and solid) are simulated separately on a path that coincides with the phase boundary. Provided that a coexistence point is known by some other means, the method finds another coexistence point through numerical integration of the Clapeyron equation. The parameters that are required for numerical integration (such as enthalpies and entropies of the coexisting phases) are obtained from the simulations. The previous process is executed iteratively to yield the entire phase boundary. In the present section, a different tracing technique [95], that avoids numerical integration of Clapeyron equations, is used.

Suppose that fluid-solid coexistence of a system of N particles at a reference temperature T_r has been established by a procedure analogous to that described in the previous section and let p_r be the value of coexistence pressure. In order to determine coexistence at a temperature $T \neq T_r$, a series of constant-pressure simulations is performed for both phases, fluid (unconstrained system) and solid (constrained cell model), covering the range from T_r to T. The output of the simulations comprises histograms of the volume and the energy that are analyzed according to histogram reweighting. The net result is the density of states up to an unknown multiplicative constant which allows for calculation of partition function ratios and thus free energy (chemical potential) differences. The following expressions can then be written for the chemical potentials of the fluid [cf. Eq. (2.1)] and the solid [cf. Eq. (2.2)] at T and p:

$$\widetilde{\mu}_f(p,T) - \widetilde{\mu}_f(p_r,T_r) = -\frac{1}{N}\ln\left[\frac{\Delta_f(N,p,T)}{\Delta_f(N,p_r,T_r)}\right], \tag{5.1}$$

$$\widetilde{\mu}_s(p,T) - \widetilde{\mu}_s(p_r,T_r) = -\frac{1}{N}\ln\left[\frac{\Delta_s(N,p,T)}{\Delta_s(N,p_r,T_r)}\right]. \tag{5.2}$$

The partition functions that appear in Eqs. (5.1) and (5.2) are evaluated from the density of states which has been obtained from histogram reweighting. Since $\widetilde{\mu}_f(p_r, T_r) = \widetilde{\mu}_s(p_r, T_r)$, the chemical potentials of the fluid and the solid in Eqs. (5.1) and (5.2) both contain the same reference point. The coexistence pressure at temperature $T \neq T_r$ is found by locating the common intersection of the two μ–p curves at temperature T. Repetition of this procedure allows for the determination of the fluid-solid phase boundary for a range of temperatures.

The resulting phase diagram in the temperature-density plane is reported in Fig. 5.7. Fluid-solid coexistence was obtained according to the tracing technique described previously. The vapor-liquid phase diagram was obtained from constant-pressure simulations of the unconstrained system for $N = 256$ and 500 particles using tempering and histogram reweighting. The vapor-liquid critical point ($T_c^* \cong 0.566$, $\rho_c^* \cong 0.401$) was estimated by fitting the coexisting densities for $N = 500$ to the scaling law for the density difference, Eq. (2.16), and the law of the rectilinear diameters, Eq. (2.17). As anticipated, the vapor-liquid demixing transition lies "inside" the fluid-solid coexistence lines. Since the range of the attractive interactions is short, the liquid phase is not stable, and hence, fluid demixing is metastable against freezing. The phase diagram is very different from that of the constrained cell model for which there is a stable critical and a triple point, cf. Fig. 5.3. Furthermore, the critical temperature of the constrained cell model is higher than that of the unconstrained system. The main reason for the increase in the critical value of the temperature is the reduction in the entropy caused by the single occupancy constraint.

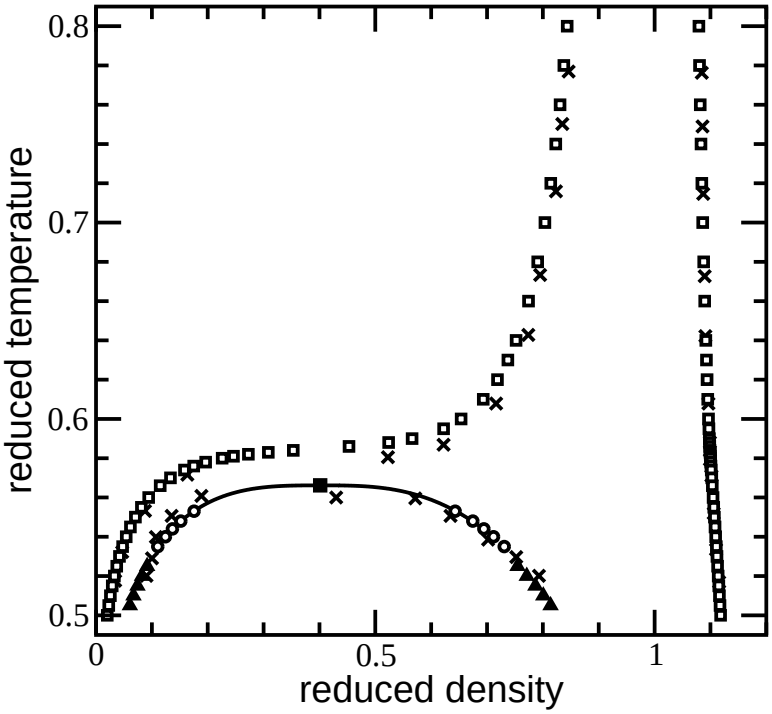


Figure 5.7: Phase diagram of the system of particles interacting through the pair potential given by Eq. (2.12) with $n = 12$, in the temperature-density plane. The open squares, (□), are the fluid-solid coexistence points. The open circles, (○), and the filled triangles, (▲), correspond to the vapor-liquid coexistence points. The filled square, (■), is the estimate of the vapor-liquid critical point. The solid line is a fit to the scaling laws for the density difference and the coexistence curve diameter; see Eqs. (2.16) and (2.17). Statistical uncertainties do not exceed the symbol sizes. All calculations have been performed for $N = 256$ particles, with the exception of the points shown as open circles, (○), that have been obtained for $N = 500$ particles. The points shown as crosses, (×), are the estimates of Ref. [125] for the same system.

The phase diagram shown in Fig. 5.7 is also compared to that of Vliegenthart *et al.* [125] obtained using Gibbs-Duhem integration [66, 65] and Gibbs ensemble Monte Carlo [101] simulations for the same system. The overall agreement is satisfactory. There are some disparities in the vapor-liquid critical region and, more importantly, in the fluid branch of the melting line above the metastable critical point. A likely explanation might be related to instabilities in the numerical integration algorithm. Note that in the region above the metastable critical point, the fluid phase is very compressible and, hence, the fluid branch of the melting curve is nearly horizontal in that region. The present tracing technique does not require numerical integration since it is based on histogram reweighting. With the exception of the fluid region above the metastable critical point, few simulations are required to obtain the fluid-solid phase boundary for wide temperatures ranges.

5.5 Conclusions

Accurate calculation of freezing/melting phase boundaries from numerical simulations still remains a challenging problem. The main advantage of the techniques used in the present chapter, compared to the previous direct and indirect approaches (see, e.g. [130] for a review), is their simplicity. The programming/bookkeeping requirements associated with the division of space into Wigner-Seitz cells is minor. Similar bookkeeping schemes (such as the Verlet neighbor list [37], for instance), abound in the simulation literature. Furthermore, the computational effort associated with handling the additional field variable, b, is also minor. The present techniques cannot find the equilibrium solid structure of the system

under consideration. The prevailing solid phase must be known a priori or found by some other means. This is also true in all numerical techniques based on cell models.

In the present chapter, these cell models were used to determine the entire phase diagram of a system of particles interacting through the pair potential given by Eq. (2.12) with $n = 12$. The vapor-liquid phase diagram was readily established by constant-pressure simulations, tempering, and histogram reweighting techniques. The fluid-solid phase boundary was established through a simple thermodynamic integration type of technique based on histogram reweighting. All these thermodynamic integration techniques require a reference point for which the free energy is either known by some other means or can be evaluated analytically. Most previous works assumed a reference system of no attractive interactions. Alternatively, the reference point was taken at $\beta = 0$ (i.e., $T \to \infty$). The main drawback of selecting such a reference point/system is the large number of intermediate states that must be simulated in order to obtain the phase boundary within the range of interest. In contrast, in the current chapter the reference fluid-solid coexistence point was determined through direct simulation of such coexistence using the generalized cell model. These methods utilizing generalized cell modes can be readily extended to cover a range of temperatures in a single simulation. Such a task was not attempted here, however, due to increased memory and storage requirements in handling multi-dimensional histograms.

The main reason for which the constrained cell model, in which each particle is confined within a single cell, has not been so popular in simulation studies of fluid-solid transitions, is the anomalous behavior associated with the appearance of phase transitions. For

systems with purely repulsive interactions, early simulations studies of the corresponding constrained cell models found evidence for a weak, first-order transition at densities that are approximately 60-70% of the density at close packing. As already emphasized above, this transition is associated with the mechanical stability limit of the solid. The work from this and previous chapters in this dissertation indicates that for systems with both repulsive and attractive interactions, the phase diagram of the corresponding cell model is much more complex than previously thought of. In addition to the previous transition, there is another phase transition between a highly expanded and a highly compressed solid phase that terminates at a critical point. The resulting phase diagram resembles that of an ordinary system with a critical and a triple point. The results of the present work as well as of previous work of the Lennard-Jones model, indicate that the range of attractive interactions has no effect on the topology of the phase diagram of the constrained cell model. The present (as well as previous) work indicates that the complexities associated with these phase transitions induce no complications in cell model-based simulations of fluid-solid phase transitions.

Chapter 6

Crystal shape modeling and control in protein crystal growth

6.1 Introduction

In the previous 3 chapters of this dissertation, cell models were introduced and then used to elucidate phase diagrams for systems interacting via a pair-wise additive potential given by 2.12. In this chapter, we move to the macroscale considering crystal shape modeling and control for lysozyme protein crystallization. Protein crystallization is imperative in the pharmaceutical industry, which is a major contributor to both scientific advancement and economic growth. It is believed that protein crystallization proceeds in three stages: nucleation, crystal growth, and cessation of growth [62]. The first part of the current chapter focuses on the second stage of protein crystallization, i.e., crystal growth, which will be

modeled via kinetic Monte Carlo (kMC) simulations. Several attempts aimed at modeling protein nucleation [38, 103] and growth [26, 34, 67] have also been made. These efforts make it possible to manipulate the size distribution and morphology of the protein crystal, which is a very critical variable for pharmaceutical products. For the tetragonal form of lysozyme experiments indicate that at low supersaturation growth depends on a lattice defect mechanism, whereas at high supersaturation growth proceeds via two-dimensional nucleation [25, 26, 123]. The present work focuses on crystals that already have been seeded via two-dimensional nucleation since a large supersaturation is normally required in a batch crystallization process [3, 124] to achieve reasonable growth rates [34]. It is noted that two-dimensional nucleation proceeds at supersaturation $\gtrsim 1.6$ [122]. Since crystal growth is a non-equilibrium process, kMC simulation methods are used to model the growth. Kinetic Monte Carlo algorithms, which form the basis for applying the Monte Carlo method to simulate dynamic processes [13, 21, 22, 43, 44, 45, 46, 47, 48, 109, 110, 119], are based on a dynamic interpretation of the Master equation [33, 80]. As is common practice in simulations of crystal growth, the solid-on-solid model [27] is used in this work to interpret the growth of protein crystals from supersaturated solutions. In the solid-on-solid approximation, particles are deposited on the growing crystal lattice without voids or overhangs, resulting in a highly compacted crystal.

The kMC methodology [20] proposed in this chapter uses rate equations originally developed by Durbin and Feher [27]. Durbin and Feher [27] found that different crystal faces produce different growth rates depending on the conditions of each independent face

of the crystal. Their results show a crossover type of behavior between the growth rates of the (110) and (101) faces, a fact that is consistent with experimental findings [26].

The simulations comprise three microscopic events, namely molecular attachment, detachment, and migration events on the (110) and (101) faces. All attachment events made in this work are implemented using monomer units. Assuming that all surface sites are available for attachment, the attachment rate is considered independent for each lattice site. Detachment and migration events, however, are dependent on their local environment. The local environment of a given lattice site is specified based on the number of nearest neighbors surrounding that site. The nearest neighbors of a lattice site are on the (N, S, E, W) directions which are of the same height or higher compared to the current lattice site. Another nearest neighbor is located directly below each surface particle, however no surface particle is without a nearest neighbor below itself due to the solid-on-solid model. Thus we will consider the nearest neighbor below each particle in the pre-exponential factor of both the desorption and migration rate equations presented in the next section. That being said, the number of nearest neighbors we explicitly consider will only be on the (N, S, E, W) directions ranging from zero to four giving a total of five classes. These classes will be used to lessen the computational cost when calculating the rates for the three microscopic events, described in the following section. Owing to the dependence of detachment rate on the surface configuration, kMC simulation is needed to compute the net crystal steady-state growth rate as a function of temperature and protein concentration in the continuous phase.

6.2 kMC methodology

As already emphasized, the solid-on-solid model, which is a simple way to look at the crystallization process, is used to model the growth of lysozyme. As noted by Ke *et al.* [61], the system size does not largely affect crystal growth. They report that no finite-size effects were found on systems of sizes 30×30, 60×60, and 120×120 sites. In the present work, a periodic square lattice of length and width of $N = 50$ sites is used. The height at a given location within the lattice is defined as the number of particles in the growth direction. Each simulation comprises 4 million events, or approximately 1600 events per lattice site on average. At the beginning of each simulation, the lattice is initialized to a flat surface. To ensure that the initial configuration does not have a noticeable impact on the results, the first 50 thousand events are discarded in order to allow the surface to roughen. Each event of our kMC simulation is chosen randomly based on the rates of the three microscopic phenomena, described below.

6.2.1 Surface Kinetics

The following description of the surface kinetics for the present model follows closely that of Ke *et al.* [61] which is based on the work by Durbin and Feher [27]. As emphasized earlier, since each surface site is available for attachment, the attachment rate is independent of each lattice site and defined as

$$r_a = K^+(\Delta\mu) = K_0^+ \exp\left(\frac{\Delta\mu}{k_B T}\right), \tag{6.1}$$

where K_0^+ is the attachment coefficient, k_B is the Boltzmann constant, T is the temperature in Kelvin, and $\Delta\mu = k_B T \ln(c/s)$, where c is the concentration and s is the protein solubility and this term is the crystallization driving force. It is noted that $r_a \propto c$. Since the total number of lattice sites is N^2, the total rate of adsorption, W_a, is defined as

$$W_a = N^2 r_a. \tag{6.2}$$

As already commented, the desorption rate of a surface particle depends on the local environment. Thus, the desorption rate of a lattice site with i nearest neighbors is given by

$$K^-(E_b) = K_0^- \exp\left(-\frac{E_b}{k_B T}\right) = K_0^- \exp\left(-i\frac{E_{pb}}{k_B T}\right), \tag{6.3}$$

where K_0^- is the desorption coefficient, i is the number of bonds, E_{pb} is the average binding energy per bond, and $E_b = iE_{pb}$ is the total binding energy. The bond-dependent desorption rate, $r_d(i)$, is thus defined as

$$r_d(i) = K^-(E_b) = K_0^- \exp\left(-i\frac{E_{pb}}{k_B T}\right). \tag{6.4}$$

It can be seen that with less nearest neighbors, the desorption rate becomes higher. The total rate of desorption is computed by

$$W_d = \sum_{i=0}^{4} W_{d_i} \quad \text{with} \quad W_{d_i} = M_i r_d(i), \tag{6.5}$$

where W_{d_i} is the total rate of desorption for each class and M_i is the number of lattice sites with i nearest neighbors. Similar to the desorption rate, Ke *et al.* [61] defined the migration rate the following way

$$r_m(i) = K_0^- \exp\left(-i\frac{E_{pb}}{k_B T} + \frac{E_{pb}}{2k_B T}\right). \tag{6.6}$$

The extra term added in Eq. (6.6) will cause migration events to be favored versus desorption events. The total migration rate is computed as

$$W_m = \sum_{i=0}^{4} W_{m_i} \quad \text{with} \quad W_{m_i} = M_i r_m(i). \tag{6.7}$$

The total rate, W_{tot}, is computed by summing over all rates of the three microscopic events, i.e., $W_{\text{tot}} = W_a + W_d + W_m$.

As was pointed out by Durbin and Feher [27] and Ke *et al.* [61], K_0^+ and K_0^- are not independent. At equilibrium, $\Delta\mu = 0$ and also the attachment and detachment rates are equal, i.e.,

$$K^+(\Delta\mu = 0) = K^-(\phi), \tag{6.8}$$

where ϕ is the binding energy per molecule of a fully occupied lattice. Thus, thermodynamic equilibrium provides the following connection between the adsorption, K_0^+, and desorption, K_0^-, coefficients

$$K_0^+ = K_0^- \exp\left(-\frac{\phi}{k_B T}\right), \tag{6.9}$$

or by rearranging,

$$K_0^- = K_0^+ \exp\left(\frac{\phi}{k_B T}\right). \tag{6.10}$$

Plugging this result back into the original equations yields the following expressions for the desorption and migration rates

$$r_d(i) = K_0^+ \exp\left(\frac{\phi}{k_B T} - i\frac{E_{pb}}{k_B T}\right), \tag{6.11}$$

$$r_m(i) = K_0^+ \exp\left(\frac{\phi}{k_B T} - i\frac{E_{pb}}{k_B T} + \frac{E_{pb}}{2k_B T}\right). \tag{6.12}$$

Since the binding energies cannot be accessed by experiments [27, 30], previous simulation work assigned a range of values to E_{pb} and ϕ, until satisfactory agreement between the calculated and the experimental growth rates was achieved [27]. The same approach is also followed in this work. Additionally we use $K_0^+ = 0.4704$ s^{-1} following from similar numbers of Refs. [27] and [50]. Additionally, we use $c = 7.8$ mg/mL, $E_{pb}/k_B = 218.99$ K and $\phi/k_B = 734.78$ K for the (110) face, and $E_{pb}/k_B = 259.34$ K and $\phi/k_B = 564.77$ K for the (101) face. These values can also be found in Table 6.1 and the corresponding kMC simulation model can be seen in Fig. 6.1 along with the experimental data extracted from [26]. It is noted that E_{pb}/k_BT and ϕ/k_BT are required in Eqs. (6.11) and (6.12) and these quantities must be divided by the temperature in Kelvin. Finally, it is important to emphasize that the computation of the growth rate for each face requires the use of kMC simulations owing to dependence of the detachment rate on the surface micro-configuration and it cannot be computed by simply subtracting the attachment and detachment rates.

Face	E_{pb}/k_B	ϕ/k_B
(110)	218.99	734.78
(101)	259.34	564.77

Table 6.1: Parameters for face (110) and (101) at 7.8 mg/mL NaCl and pH= 4.5%. Additionally, $K_o^+ = 0.4704$ seconds^{-1}.

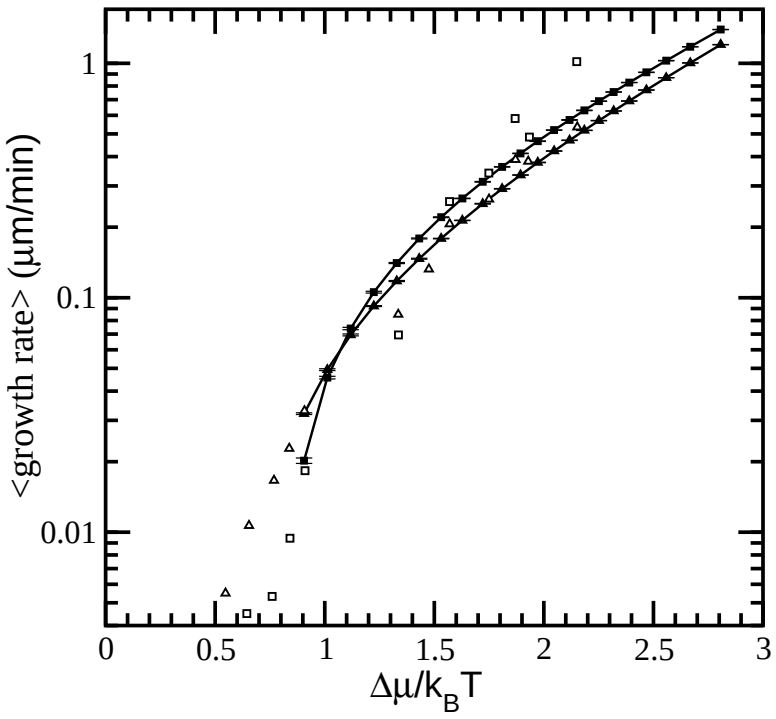


Figure 6.1: The expected growth rate versus the degree of supersaturation at $c = 7.8$ mg/mL. The (■) represents the (110) face with $E_{pb}/k_B = 218.99$ K and $\phi/k_B = 734.78$ K. The (▲) represents the (101) face with $E_{pb}/k_B = 259.34$ K and $\phi/k_B = 564.77$ K. The error bars represent two standard deviations of the growth rate. The (□) and (△) represent the (110) and (101) faces respectively, obtained from Ref [26].

6.2.2 Event execution

In order to execute an event, a uniform random number, $\zeta_1 \in [0,1)$, is generated. If $\zeta_1 \leq W_a/W_{\text{tot}}$, then an adsorption event is executed. If $W_a/W_{\text{tot}} < \zeta_1 \leq (W_a + W_d)/W_{\text{tot}}$, then a desorption event is executed. Lastly, if $\zeta_1 > (W_a + W_d)/W_{\text{tot}}$, then a migration event is executed. For the case of adsorption, a lattice site is chosen at random for the adsorption event to take place. For desorption and migration, the specific class needs to be determined. In the case of a desorption event, the k^{th} class is determined to be an integer from $[0,4]$ such that,

$$\frac{W_a + \sum_{i=0}^{k-1} W_{d_i}}{W_{\text{tot}}} < \zeta_1 \leq \frac{W_a + \sum_{i=0}^{k} W_{d_i}}{W_{\text{tot}}}. \tag{6.13}$$

Once the class is determined, a second random number, ζ_2, is generated to select a random lattice site within class k to execute the desorption event. Migration events work in an analogous way to desorption events with a minor modification to Eq. (6.13) as follows,

$$\frac{W_a + W_d + \sum_{i=0}^{k-1} W_{m_i}}{W_{\text{tot}}} < \zeta_1 \leq \frac{W_a + W_d + \sum_{i=0}^{k} W_{m_i}}{W_{\text{tot}}}. \tag{6.14}$$

In Durbin and Feher's work [27] it was found that only half the lattice sites were available for adsorption on the (101) face, whereas every lattice site is available on the (110) face. This is due to the fact that only half the molecules on the (101) face have dangling bonds (i.e., the points of attachment for incoming molecules), whereas every molecule on the (110) face has dangling bonds [27]. In the present work, this behavior is modeled by accepting 50% of adsorption events on the (101) face, compared to 100% on the (110)

face. In the case of desorption, the events are always accepted for both faces. Similar to desorption events, migration events are always accepted as long as there exists at least one available migration site. For this work, an available migration site is a nearest neighboring site which is lower in height than the current lattice site where the migration event is taking place. Similar to Gilmer and Brennema [49], each available migration site is given equal probability to accept the displaced particle.

After each event is executed, a time increment, Δt, is computed based on the total rate of the microscopic events as follows

$$\Delta t = -\ln(1-\zeta)/W_{\text{tot}}, \tag{6.15}$$

where ζ is a uniform random number, $\zeta \in [0,1)$. Events will continue to take place until the conclusion of the simulation.

6.3 Linking growth rate ratio with concentration and temperature

The data generated from the kMC simulations was used to construct a nonlinear algebraic equation (visualized by a 3-D plot). This equation was utilized in the model predictive control (MPC) formulation to relate the crystal growth rate ratio to the temperature and protein concentration in the continuous phase. The shape of the crystal is determined by the two independent faces and their relative growth rates. Therefore, it is possible to control the evolution of crystal shapes by controlling the ratio between the growth rates of the

two faces. The growth rate ratio curve (discussed in Section 6.4) and the measurement of current temperature as well as of the concentration are assumed to be available to the controller.

6.3.1 Modeling steady-state growth rate ratio dependence on temperature and solution concentration

The three variables (i.e., temperature, growth rate ratio, and protein concentration) used for the model construction are obtained from the open-loop simulations using the kMC model in Section 6.2. The protein solution concentration (temperature) is fixed during each open-loop simulation to observe the dependence of the crystal growth rate ratio on temperature (protein solution concentration).

A nonlinear algebraic equation relating the relative growth rate ratio versus protein concentration and temperature is derived to quantify the evolution of the crystal growth accounting for the effect of protein concentration variations and temperature changes in supersaturated protein solutions. The concentration variation results from the mixing problem of the batch process as well as the technical difficulty in the measurement of protein concentration. Therefore, a Gaussian noise with the following mean and covariance is added to the concentration measurements, $c(t)$, in simulations:

$$\langle c(t)\rangle = c_n, \qquad \langle c(t)c(t')\rangle = \sigma_g^2 c_n^2, \tag{6.16}$$

where c_n is the nominal concentration of the system and σ_g^2 shows how far a set of measured

concentrations deviates from its nominal value. In order to suppress the uncertainty in the protein concentration, the kMC simulations are repeated to compute expected values [i.e., averages denoted as $\langle\cdot\rangle$ see, e.g., Eq. (6.16)]. Also, it is important to note that the concentration variation results in a change of the attachment rate [cf. Eq. (6.1)].

Different operating conditions strongly affect the crystallization process and, consequently, the behavior of the growth rate ratio profile. Fig. 6.2 shows the ratio data collected from open-loop simulations. In this work, the temperature, T, is chosen as the manipulated input, and the pH and NaCl concentration are maintained fixed at 4.6 and 5.0%, respectively. The protein concentration as well as the temperature dependence of the growth rate ratio is obtained by generating a 3-D plot for a variety of protein concentrations ranging from 7.2 to 8.4 mg/mL, and temperatures ranging from 4 to 25°C. It should be noted that 25°C is the maximum temperature at which tetragonal crystals can be obtained [5]. There is a maximum in the growth rate ratio located at $\sim$ 15°C. Thus the ratio tends to decrease as the temperature increases or decreases from 15°C and the protein concentration decreases. Fig. 6.2 is the 3-D surface plot of $r = f(T,C)$, where f is a nonlinear function of the growth rate ratio, to better illustrate the effects of changing temperatures and protein concentrations. To this end, in Fig. 6.2, the data was interpolated to fit a regularly spaced grid. It can be clearly seen that the growth rate ratio is a strong function of the temperature and protein concentration, and this dependence is the basis for using temperature, T, to control protein crystal shape.

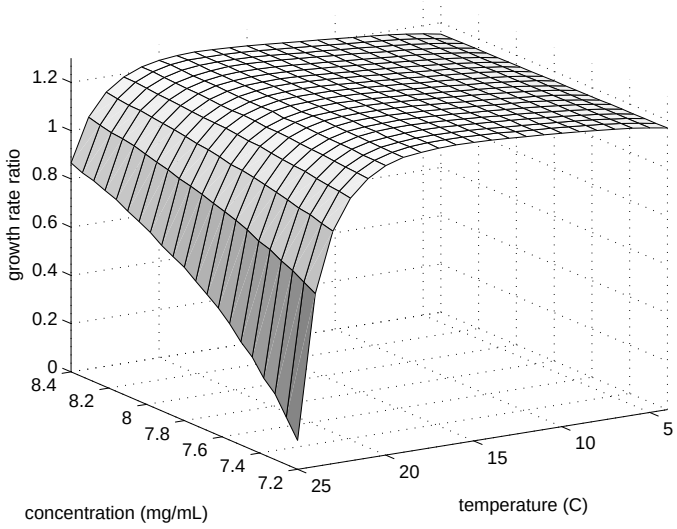


Figure 6.2: Surface plot of the growth rate ratio data for tetragonal lysozyme at pH 4.6 and 5% NaCl. The data from the open-loop kMC simulation are plotted to demonstrate the effect of temperature and concentration variations on this ratio. Protein concentration and temperature range from 7.2 to 8.4 mg/mL and 4 to 25°C, respectively.

6.4 Model predictive control of crystal shape

As emphasized previously, in the kMC simulations molecular attachment, detachment, and migration events are considered. Depending on the relative attachment energy of the crystal faces and assuming that the independent crystal faces that appear during the crystal growth are the (110) and (101) faces, kMC simulations reproduce the experimentally observed crossover behavior in the crystal growth rates between the two faces [cf. Fig. 6.1]. In this section, a model predictive controller is designed based on the 3-D plot (i.e., the non-linear growth rate ratio equation), Fig. 6.2, to suppress the concentration variation and achieve the desired set-point values by manipulating the temperature. A desired set-point value of the growth rate ratio is included in the cost function in the MPC formulation. Regarding the choice of MPC for the controller design, it is noted that classical control schemes like proportional (P) control cannot be employed to explicitly account for input/state constraints, optimality considerations, and the nature of the attachment, detachment, and migration processes. Also, dynamic open-loop optimization may be used but it does not provide robustness against model inaccuracies and fluctuation in the protein concentration.

6.4.1 Model predictive control formulation

To this end, consider the control problem of the growth rate ratio by using an MPC design. The expected value of the growth rate ratio, $\langle \alpha \rangle$, is chosen as the control objective, where α is defined as a ratio G_{110}/G_{101}. Here G_{110} and G_{101} signify the growth rate on the (110) and

(101) faces, respectively. The temperature is used as the manipulated input. The pH and NaCl concentration are fixed during all closed-loop simulations. The proposed modeling and control methods do not depend on the specific number of the manipulated variables and can be easily extended to the case of multiple inputs. To account for a number of practical considerations, several constraints are added to the control problem. First, there is a constraint on the range of variation of the temperature such that $4 \leq T \leq 25^{\circ}$C. This constraint ensures validity for the kMC model by imposing a range that will not damage the protein crystal. Another constraint is imposed on the rate of change of the temperature to account for actuator limitations. The control action at time t is obtained by solving a finite-horizon optimal control problem. The cost function in the optimal control problem includes a penalty on the deviation of $\langle \alpha \rangle$ from its set-point value which is determined based on the desired crystal shape. Since the protein crystallization process is a batch process, a desired minimum thickness (i.e., minimum amount of growth on each face of the crystal) may be required to ensure that the crystal has the properties necessary for its desired application at the end of the crystallization process. The thickness may be obtained by adjusting the time that growth takes place since the growth rates can be estimated. However, in this work to simplify the development and focus on crystal shape control, imposing a minimal crystal size of the two surfaces is disregarded in this MPC formulation. The optimal temperature profile is calculated by solving a finite-dimensional optimization problem in a receding

horizon fashion. The MPC problem is formulated as follows:

$$\begin{aligned} \underset{T_1,\ldots,T_i,\ldots,T_p}{\text{minimize}} \quad & \sum_{i=1}^{p} F_i \\ \text{subject to} \quad & F_i = (\langle\alpha\rangle - \alpha_{\text{set}})^2 \\ & \alpha = f(T,C) \\ & T_{\min} \leq T_i \leq T_{\max} \\ & \left|\frac{T_{i+1} - T_i}{\Delta}\right| \leq R_T \\ & i = 1,2,\ldots,p \end{aligned} \tag{6.17}$$

where t is the current time, F_i is the cost function expressing the deviation of $\langle\alpha\rangle$ from its set-point ratio, α_{set}, Δ is the sampling time, p is the number of prediction steps, $p\Delta$ is the specified prediction horizon, t_i, $i = 1,2,\ldots,p$, is the time of the i^{th} prediction step, $t_i = t + i\Delta$, respectively, $T_i, i = 1,2,\ldots,p$, is the temperature at the i^{th} step, $T_i = T(t + i\Delta)$, respectively, $T_{\min}$ and $T_{\max}$ are the lower and upper bounds on the temperature, respectively, and R_T is the limit on the rate of change of the temperature. The optimal set of control actions $(T_1, T_2, \ldots, T_p)$, is obtained from the solution of the multi-variable optimization problem of Eq. (6.17), and only the first value of the manipulated input trajectory, T_1, is applied to the protein crystallization process from time t until the next sampling time, when a new measurement of protein concentration in the continuous phase is received from the kMC simulation and the MPC problem of Eq. (6.17) is re-solved for the computation of the next optimal input trajectory. The physical properties of the system (i.e., protein solubility and so on) were obtained from experimental data for the lysozyme protein solution [14]. In a previous work, empirical expressions were obtained for the growth rates by

fitting algebraic expressions to the available experimental data [117]. In the present work, the growth rates are computed following the kMC methodology from Section 6.2. Furthermore, the stochastic nature of the system and the model uncertainty will be accounted for in the protein concentration variations; see, e.g., Shi *et al.* [116] and Chiu and Christofides [19], for results on model predictive control and robust-control of crystallization systems, respectively.

6.5 Single crystal results

6.5.1 Open-loop simulations

For a given set of the simulation conditions comprised of temperature, pH, salt and protein solution concentrations, the method described in the Section 6.2 results in averaged lysozyme face growth rates at various values of supersaturations.

In Fig. 6.1, crystal growth is modeled at supersaturation $0.9 \lesssim \ln(c/s) \lesssim 2.8$, where c (mg/mL) is the solution protein concentration and s (mg/mL) is the solubility. The growth rates produced are the average growth rate for each set of conditions over 10 independent kMC runs and are compared against data from Durbin and Feher [26]. For Fig. 6.1, $c =$ 7.8 mg/mL. The solubility is determined by Eq. (6.18) using a third order polynomial of solubility in terms of temperature T (°C) at pH = 4.6 and 5%(w/v) NaCl [14, 15].

$$\begin{aligned} s(T) &= 3.506 \times 10^{-4} T^3 - 9.046 \times 10^{-3} T^2 \\ &\quad + 1.303 \times 10^{-1} T + 7.209 \times 10^{-2}. \end{aligned} \tag{6.18}$$

The above equation allows for accurate modeling of the solubility in terms of temperature at the selected pH and salt concentration with an error of 5.4% [15]. As is evident in Fig. 6.1, crossover behavior between the (110) and (101) faces does, indeed, occur. Specifically, the crossover value of the growth rate from Fig. 6.1 is ~ 0.06 μm/min and Durbin and Feher [26] show ~ 0.1 μm/min. The same growth rates are shown in terms of temperature in Fig. 6.3. This figure directly shows the relationship between the growth rates and the temperature. As shown in Fig. 6.3, as temperature increases, the growth rates on both faces decrease. The growth rate for the (110) face decreases at a higher rate than the growth rate for the (101) face. These results follow previous experimental work for pH and salt concentration near our given values [81, 82].

Furthermore, the kMC simulations were tested over a varying range of protein concentrations in the liquid solution. Concentrations were taken at finite values of 6.8, 7.8, 8.8, and 9.8 mg/mL. As anticipated, higher protein concentrations and lower temperatures yield higher growth rates. However, as temperature rises, the difference between the growth rate of the constant concentration curves becomes smaller. Each of the points in Fig. 6.4 and Fig. 6.5 is taken from averaging over 10 kMC simulations to compute accurate expected values, where the error bars represent two standard deviations. Since the rates for desorption and migration change after the execution of each microscopic event, the steady-state growth rate must be computed by averaging over the individual growth rates obtained from several kMC processes. Fig. 6.5 displays the 4 finite values of concentration at $T = 14^{\circ}$C located in the rectangular box from Fig. 6.4. There are minimal fluctuations in Fig. 6.5,

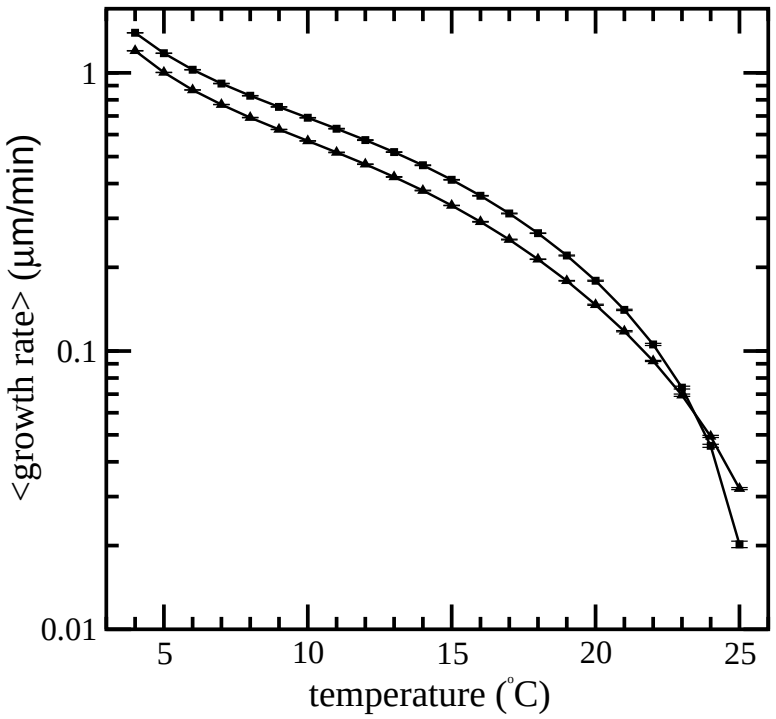


Figure 6.3: The expected growth rate versus temperature of $c = 7.8$ mg/mL. The (■) represents the (110) face with $E_{pb}/k_B = 218.99$ K and $\phi/k_B = 734.78$ K. The (▲) represents the (101) face with $E_{pb}/k_B = 259.34$ K and $\phi/k_B = 564.77$ K. The error bars represent two standard deviations of the growth rate.

that allow one to deduce reliable estimates for the growth rate.

6.5.2 Closed-loop simulations

In this section, the proposed model predictive controller of Eq. (6.17) is applied to the kMC model described in Section 6.2. These closed-loop simulations will be performed to test the ability of the model predictive controller to drive the growth rate ratio to desired set-point values in the presence of significant variation / disturbance in the operating conditions. At each sampling time (1 second), the optimal temperature, obtained by solving the optimization problem of Eq. (6.17), is applied to the closed-loop system until the next sampling time. The optimization problem is solved via a local constrained minimization algorithm using the nonlinear algebraic model described previously to predict the dependence of the crystal growth rate ratio on temperature and protein concentration.

The desired values (set-points) in the closed-loop simulations are $\langle\alpha\rangle = 1.19$ and $\langle\alpha\rangle = 0.67$. The protein concentration randomly varies following the Gaussian distribution of Eq. (6.16) and the pH and NaCl concentration are fixed at 4.6 and 5.0%, respectively, and the initial temperature is 15°C. The nominal concentration is 7.8 mg/mL with $\sigma_g = 2.5\%$ ($\sigma_g \times c_n = 0.195$ mg/mL). We would like to note that no material balance was included for the solute concentration since the focus of this work is on modeling and control of the crystal shape of an individual crystal over a time period which has minimal concentration drop. The maximum rate of change of the temperature is 2°C/min, and the minimum and maximum temperature allowed by the controller is 4°C and 25°C, respectively. Since the

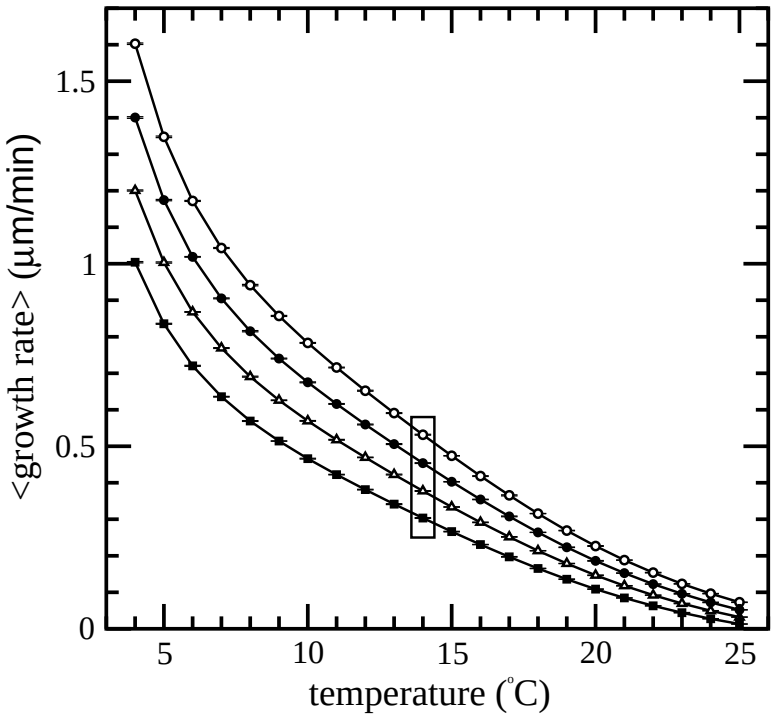


Figure 6.4: The expected growth rate on the (101) face of pH 4.6, 5% NaCl, and various values of lysozyme concentration. The labeling is as follows, (○): $c = 9.8$ mg/mL, (●): $c = 8.8$ mg/mL, (△): $c = 7.8$ mg/mL, and (■): $c = 6.8$ mg/mL. For these simulations, $E_{pb}/k_B = 259.34$ K and $\phi/k_B = 564.77$ K. The error bars represent two standard deviations of the growth rate.

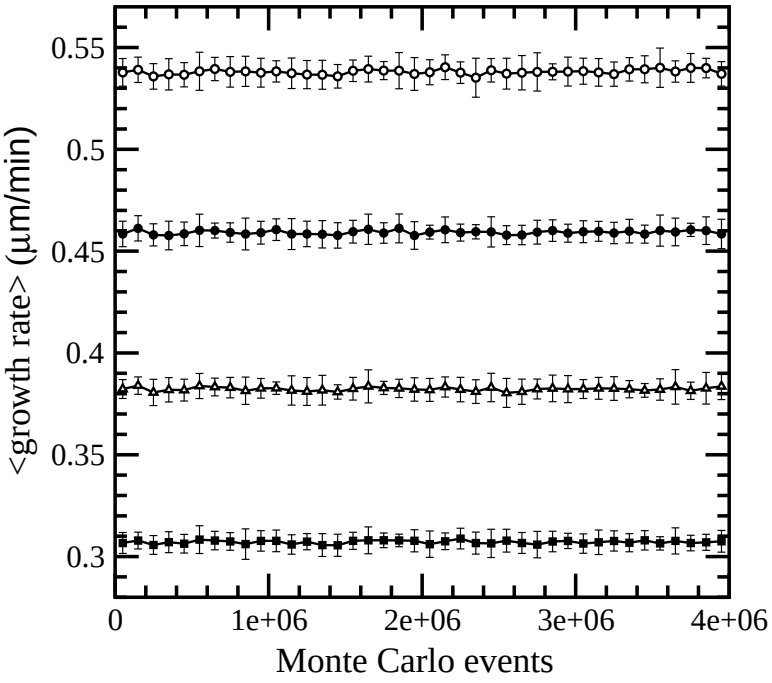


Figure 6.5: The expected growth rate on the (101) face at pH 4.6, 5% NaCl, $T = 14^{\circ}$C, and various values of lysozyme concentration. The labeling is as follows, (○): $c = 9.8$ mg/mL, (●): $c = 8.8$ mg/mL, (△): $c = 7.8$ mg/mL, and (■): $c = 6.8$ mg/mL. For these simulations, $E_{pb}/k_BT = 0.903$ and $\phi/k_BT = 1.967$. The error bars represent two standard deviations of the growth rate.

MPC formulation uses the steady-state growth rates, the number of prediction steps is set to be $p = 1$. The time interval between two sampling times is 1 second. The prediction horizon of each step is fixed at $p\Delta = 1$ second. The concentration varies every 0.333 seconds. The computational time that is used to solve the optimization problem with the current available computing power is negligible with respect to the sampling time interval. The closed-loop simulation duration is 3000 seconds.

In the closed-loop simulations associated with controlling the growth rate ratio to the desired set-point values, the control objective is to separately regulate the expected ratio to the desired values, $\langle\alpha\rangle = 1.19$ and $\langle\alpha\rangle = 0.67$ respectively. Thus, the cost function of this problem contains penalty on the deviation of the expected growth rate ratio from the set-point value.

The results of the closed-loop simulations are shown in Figs. 6.6-6.9. From Fig. 6.6, for the low set-point ratio $\langle\alpha\rangle = 0.67$, it can be seen that soon after the initial rise, the growth rate ratio decreases constantly, then fluctuates for the rest of the time towards the end of the simulation. Although disturbance results in the fluctuation of the concentration, the MPC can successfully drive the growth rate ratio to the desired set-point. Also, as is shown in Fig. 6.8 and 6.9, the use of expected values (i.e., averages obtained from 100 independent simulations for the same set of conditions) in the control formulation, suppresses fluctuations. Again, the error bars displayed are two standard deviations in Figs. 6.8 and 6.9. The shape of the protein crystal can be estimated from Fig. 6.6 and it is slightly elongated along the (101) direction for the lower set-point ratio while it is

more equidimensional to slightly elongated along the (110) direction for the higher set-point ratio. In Fig. 6.7 and Fig. 6.9, temperature profiles are displayed to show how the optimal input changes over time. The temperature reaches steady-state with minor fluctuation once the growth rate ratio settles onto the desired ratio. The expected growth over 3000 seconds for the set-point ratio $\langle\alpha\rangle = 1.19$ on the (110) and (101) faces were approximately 45.69 and 38.22 μm/min, respectively giving a ratio of 1.195. In contrast, the expected growth for the set-point ratio $\langle\alpha\rangle = 0.67$ on the (110) and (101) faces were approximately 2.07 and 2.42 μm/min, respectively giving a ratio of 0.855. The reason this ratio is much further away from the set-point value is due to the starting temperature at 15°C compared to the steady state value of approximately 25°C causing the growth rate at the beginning of the run to be much higher than at the end. If we remove the first 300 seconds from the growth, the expected growth for the set-point ratio $\langle\alpha\rangle = 0.67$ on the (110) and (101) faces were approximately 1.07 and 1.58 μm/min, respectively giving a ratio of 0.677, which is in good accord with the set-point. This shows the importance of the starting temperature in the batch process to control the crystal shape. The larger error bars are due to the higher sensitivity of the crystal growth rate ratio in the presence of solute concentration disturbance at higher temperatures. We note that the magnitude of the solute concentration disturbance is independent of the nominal solute concentration value and thus the disturbance has a more pronounced effect at low solute concentrations resulting in larger error bars given the gaussian nature of the disturbance.

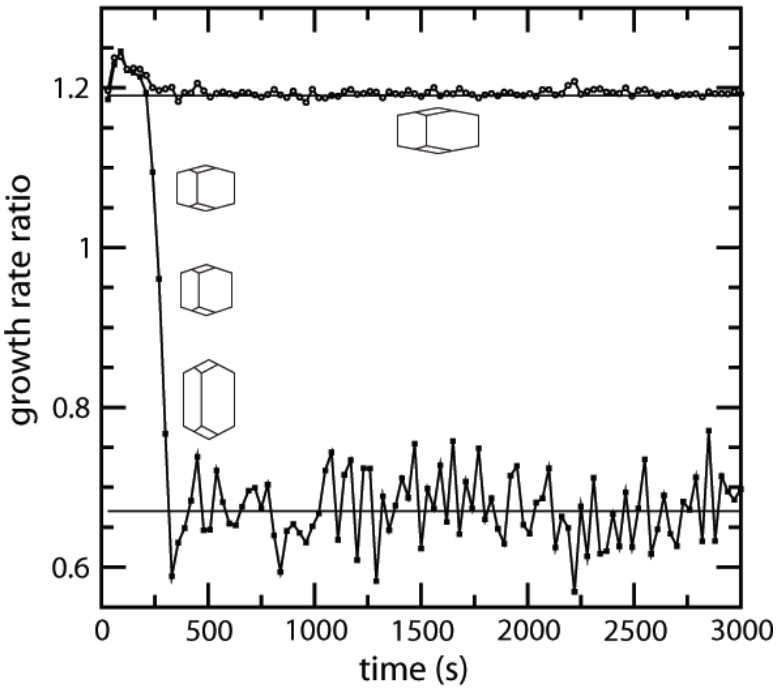


Figure 6.6: Profiles of growth rate ratio under closed-loop operation. The growth rate ratio set-point values are (○): $\langle\alpha\rangle = 1.19$ and (■): $\langle\alpha\rangle = 0.67$, respectively.

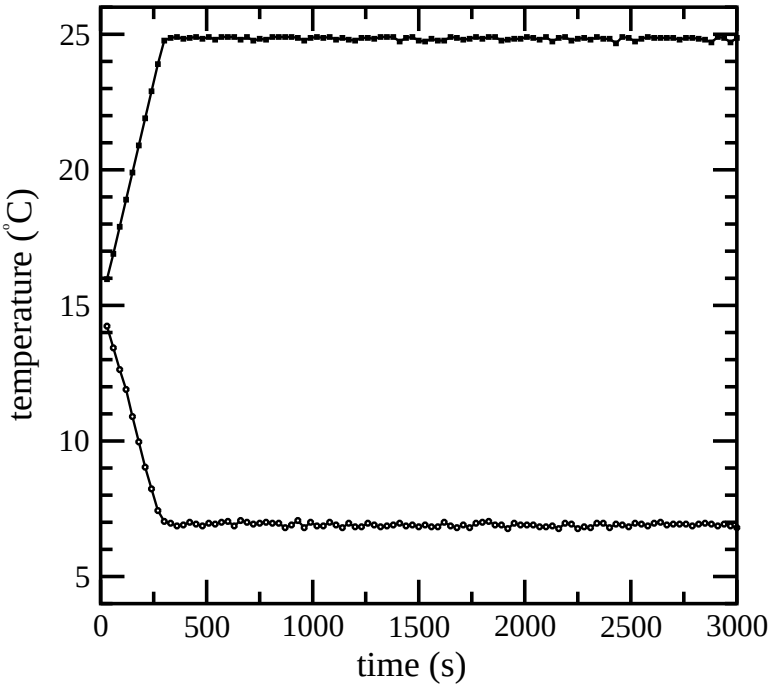


Figure 6.7: Profiles of temperature under closed-loop operation. The growth rate ratio set-point values are (○): $\langle \alpha \rangle = 1.19$ and (■): $\langle \alpha \rangle = 0.67$, respectively.

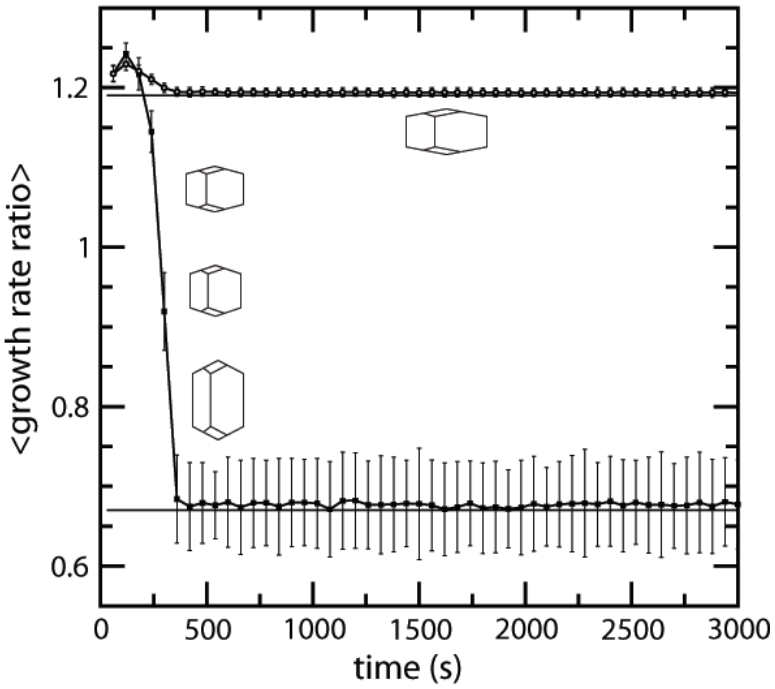


Figure 6.8: Profiles of the expected values of growth rate ratio under closed-loop operation. The growth rate ratio set-point values are (○): $\langle\alpha\rangle = 1.19$ and (■): $\langle\alpha\rangle = 0.67$, respectively. The error bars represent two standard deviations of the growth rate ratio.

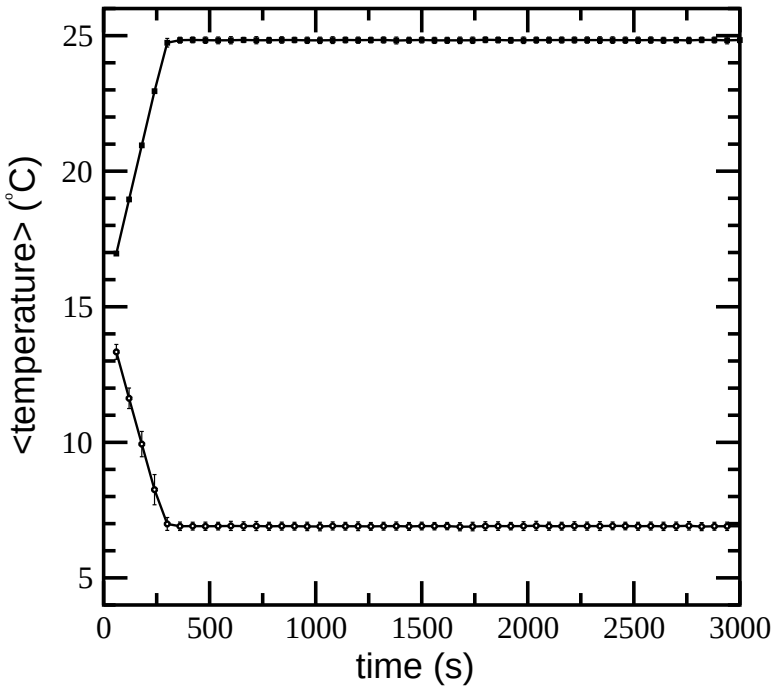


Figure 6.9: Profiles of the expected values of temperature under closed-loop operation. The growth rate ratio set-point values are (○): $\langle\alpha\rangle = 1.19$ and (■): $\langle\alpha\rangle = 0.67$, respectively. The error bars represent two standard deviations of the temperature.

6.6 Batch Crystallization Process

Up to this point, a microscopic model has been considered for the crystal growth process of a single lysozyme crystal. In order to consider a batch crystallization process, we must reformulate our overall setup to allow for nucleation which will then be followed by kMC simulations for each crystal representing the crystal growth step. Many other additional considerations must be made in order to account for the population of crystals, and will be presented in the following sections and subsections.

6.6.1 Nucleation

The following equation for the nucleation rate, $J(0,0,t)(\sigma)$, at time t (with units $[cm^{-3} \cdot \text{sec}]$), was extracted from Ref. [39] at pH $= 4.5$ and 4%(w/v) NaCl:

$$J(0,0,t)(\sigma) = \begin{cases} 0.041\sigma + 0.063 & \text{if} \quad \sigma \geq 3.11 \\ 8.0 \times 10^{-8} \exp(4.725\sigma) & \text{if} \quad \sigma < 3.11 \end{cases} \tag{6.19}$$

where σ is the supersaturation defined as $\ln(c/s)$. Since the working range has been changed to account for available nucleation rate data, the kMC growth model given by Eqs. (6.1) to (6.15) must be fit to experimental growth rate data again and is described in the next subsection. Additionally, the protein solubility, which is dependent on temperature (°C), is redefined for pH $= 4.5$ and 4%(w/v) NaCl by [14, 15] with the following

third-order polynomial:

$$\begin{aligned} s(T) &= 2.88\times10^{-4}T^3 - 1.65\times10^{-3}T^2 \\ &\quad +4.619\times10^{-2}T + 6.008\times10^{-1}. \end{aligned} \tag{6.20}$$

It is reported that Eq. (6.20) gives an error of 6.8% for the selected pH and salt concentration [15].

6.6.2 Fitting experimental data via open-loop simulations

For a given set of the simulation conditions comprised of temperature, pH, salt, and protein solute concentrations, the method described earlier in Section 6.2 results in averaged lysozyme face growth rates at various values of supersaturation.

In Fig. 6.10, crystals have been grown at supersaturation $2.1 \lesssim \ln(c/s) \lesssim 3.95$, where c (mg/mL) is the protein solute concentration and s (mg/mL) is the solubility, and the estimated growth rate at 4.0% NaCl is compared with the experimentally measured data at 3.5% and 5.0% NaCl from Ref. [25], respectively. The parameters for the kMC simulation in Figs. 6.10 to Fig. 6.18 are listed in Table 6.2. The original parameters were taken from [27] and modifications were performed by making small changes to these parameters to ensure both validity in the crossover behavior of the growth rates of the (110) and (101) faces, and that the data was in the correct regime for the 4% NaCl to make use of available nucleation rate data [39] given previously. The growth rates produced are the average growth rate for each set of conditions over 10 independent kMC runs. For Fig. 6.10,

$c = 45.0$ mg/mL. The solubility is determined by Eq. (6.20) in terms of temperature (°C) at pH $= 4.5$ and 4%(w/v) NaCl. As is evident in Figs. 6.10 and 6.11, cross-over behavior between the (110) and (101) faces does, indeed, occur. It is also apparent in Fig. 6.10 that the kMC simulation models the experimental result by Ref. [27]. We also note that with a different set of experimental data, the parameters can be fit again to achieve a satisfactory agreement between simulation and experiment.

Face	ϕ/k_B	E_{pb}/k_B
(110)	1077.26	227.10
(101)	800.66	241.65

Table 6.2: Parameters for face (110) and (101) at 45 mg/mL NaCl and pH= 4.5 at $T =$ 18°C. Additionally, $K_o^+ = 0.211$ seconds^{-1}.

6.7 MPC of crystal size and shape for batch crystallization

In the kMC simulations, crystal nucleation is considered alongside crystal growth via molecular attachment, detachment, and migration events. Since the nucleation rate depends on the supersaturation, it has been manipulated by changing the temperature for a given concentration. Parameters of crystal growth have also been appropriately chosen for kMC simulations to capture the experimentally observed cross-over behavior in the crystal growth rates between the (110) and (101) faces [26, 27], which are given in Table 6.2. In this section, a model predictive controller is designed based on the non-linear growth rate

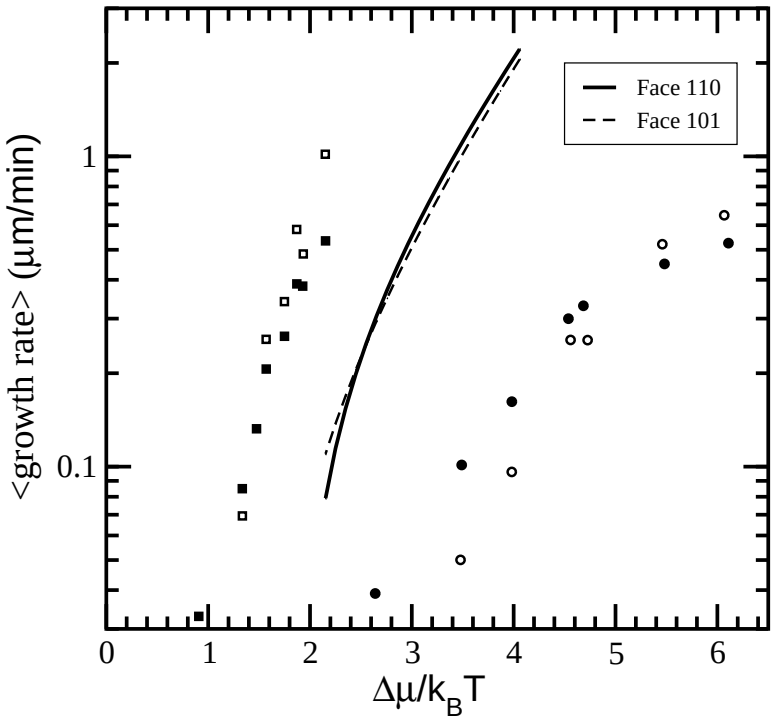


Figure 6.10: The expected growth rate versus the degree of supersaturation at $c = 45$ mg/mL and 4% NaCl are shown as the solid (110 face) and the dashed (101 face) lines. The (■) and (□) represent the measured experimental data for 110 and 101 faces with 5% NaCl; (●)/(○) represent the measured experimental data with 3.5% NaCl; extracted from Ref. [26] at pH= 4.6%.

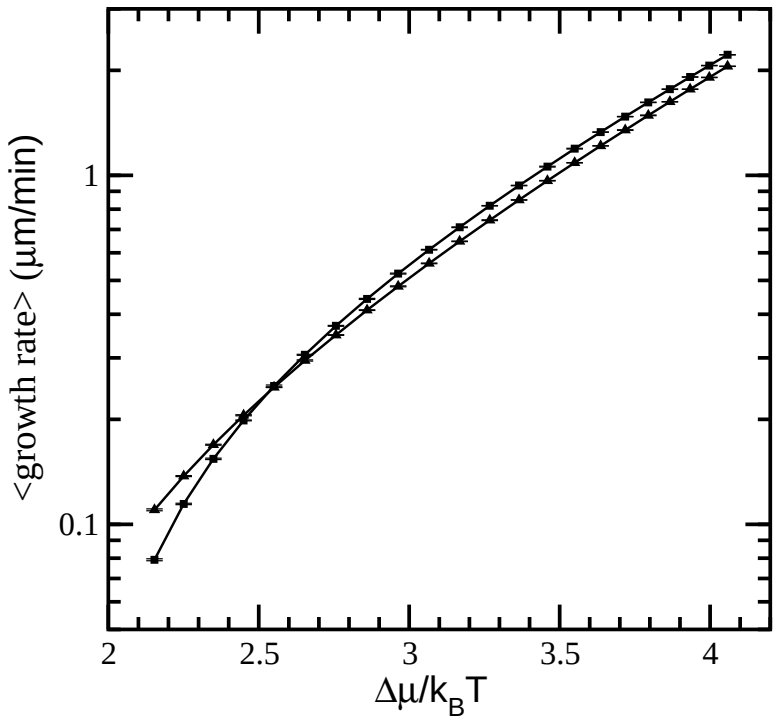
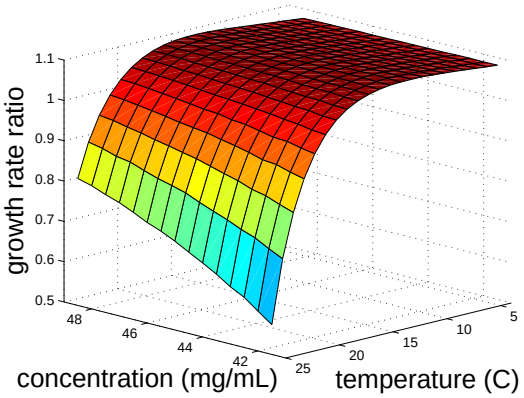


Figure 6.11: The expected growth rate versus the degree of the supersaturation at $c = 45$ mg/mL with $E_{pb}/k_BT = (0.78/0.83)$ and $\phi/k_BT = (3.7/2.75)$ for 110 (■) and 101 faces (▲), respectively. Error bars represent two standard deviations of ten simulations for each point.

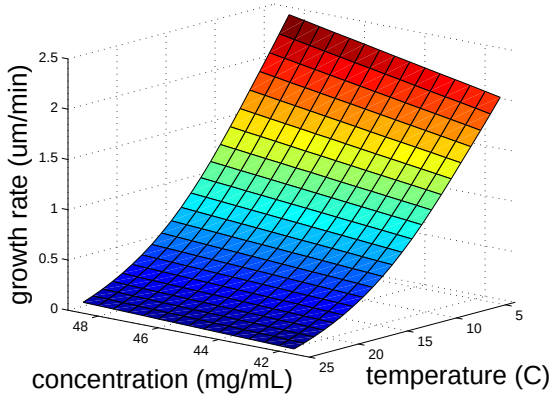
equations (i.e., 3-D plots) for both faces and the growth rate ratio which has been shown in Figs. 6.12(a), 6.12(b), and 6.12(c) to suppress the uncertainty in the solute concentration, and achieve the desired set-point values by manipulating the temperature. A desired set-point value of the growth rate ratio and minimum crystal sizes for the (110) and (101) faces are included in the cost function in the MPC formulation. MPC resolves the drawbacks of the classical control schemes like proportional integral (PI) control, which cannot explicitly take into account input/state constraints, optimality considerations, and the nature of the nucleation, attachment, detachment, and migration processes. Furthermore, dynamic open-loop optimization may be used. However, open-loop optimizations does not provide robustness against model inaccuracies and stochastic variation in the protein concentration and the operating environment.

6.7.1 The population balance equation for protein crystallization

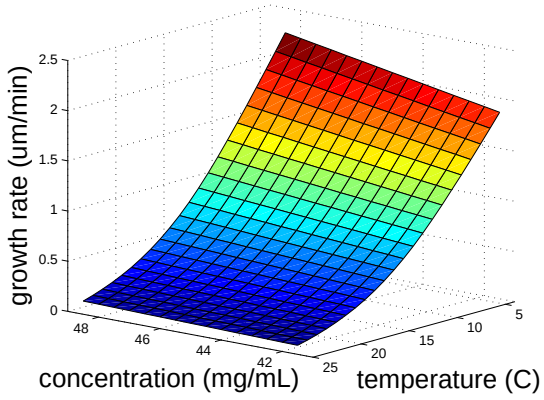
The evolution of the particle size and shape distribution in a protein crystallization process is typically described by a population balance equation (PBE). More specifically, for the batch crystallizer considered in this work, a population balance model can be used to describe the evolution of the crystal size distribution (CSD), $n(h_{110}, h_{101}, t)$, which represents the number of crystals at time t with height h_{110} and h_{101} on each face. To describe the behavior of the CSD in a crystallization process, it is necessary to know the nucleation rate as well as the growth rate. In Liu's work [74], they only considered crystal growth from seeds. However, in the present work we consider the nucleation process as well. We as-



(a) Growth rate ratio



(b) Growth rate of the (110) face



(c) Growth rate of the (101) face

Figure 6.12: Plots of the growth rate data obtained for the (110) face, the (101) face, and the growth rate ratio between (110) and (101) faces for tetragonal lysozyme protein crystals at pH $= 4.5$. It is noted that the growth rate ratio, G, is equal to $\frac{G_{110}(t)}{G_{101}(t)}$ and is dimensionless. The data from the open-loop kMC simulation are plotted to demonstrate the effect of temperature and concentration variations on growth rates. Protein concentration and temperature range from 41.5 to 48.5 mg/mL and 4 to 25°C, respectively.

sume that the height of nuclei for the (110) and (101) faces, h_{110} and h_{101}, is negligible, and the number of nuclei newly formed at time t is denoted as $n(0,0,t)$. We have two reasons supporting this assumption. First, a nucleus consists of 3 (to 4) HEW lysozyme molecules, whose size is negligible relative to the final crystal size [83]. In addition to this, nuclei can not be observed until they reach the resolution limit of the optical microscope, $\sim 0.5\mu$m [121]. The resulting population balance equation has the following form:

$$\begin{aligned}
&\frac{\partial n(h_{110},h_{101},t)}{\partial t} + \partial\left(\frac{(G_{110}(h_{110},t)n(h_{110},h_{101},t))}{\partial h_{110}}\right) + \partial\left(\frac{(G_{101}(h_{101},t)n(h_{110},h_{101},t))}{\partial h_{101}}\right) = 0 \\
&\alpha = f_{\alpha}(T,C) \\
&G_{110}(h_{110},t) = f_{110}(T,C,t) \\
&G_{101}(h_{101},t) = f_{101}(T,C,t) \\
&n(h_{110},h_{101},t) = \sum_{k=0}^{t-1} \frac{\Delta}{J(0,0,k)\cdot V} \\
&J(0,0,t) = f_{nucleation}(\sigma(t))
\end{aligned} \tag{6.21}$$

where $G_{110}(h_{110},t)$ and $G_{101}(h_{101},t)$ are the growth rates for (110) and (101) faces, Δ is the sampling time, V is the volume of the crystallizer, $\alpha = \frac{G_{110}(t)}{G_{101}(t)}$, is the growth rate ratio. The nonlinear equations, f_{α}, f_{110}, and f_{101}, show their dependencies on temperature, solute concentration and time, respectively. The number of crystals nucleated at time, t, is obtained from Ref. [39], where the homogeneous nucleation rate is determined by the supersaturation and NaCl concentration, as well as pH level. In this work, we assume that heterogeneous nucleation, i.e., nucleation on a surface, is negligible. The nucleation rate

was given previously by Eq. (6.19). Instead of solving Eq. (6.21) directly, we execute several kMC simulations with nuclei formed at different times. This is considered to be equivalent to solving Eq. (6.21). Also, for a given supersaturation, the nucleation rate is obtained from the experimental results by Galkin *et al.* [39]. In the next subsection, the controller design method given in Section 6.4 is further generalized to the case with nucleation of crystals.

6.7.2 Model Predictive Control (MPC) formulation

We consider the control of the crystal size and shape for a crystal population nucleated at different times as the batch crystallization process proceedes by using a Model Predictive Control (MPC) design. The expected value of the growth rate ratio, $\langle \alpha \rangle$, is chosen as the control objective, where α is defined as the ratio G_{110}/G_{101}. A desired minimum crystal size is also included in the cost function of the MPC formulation and the temperature is used as the manipulated input. Although there are various factors that affect the protein solubility and crystal morphology during the crystallization process [2, 79], we assume only temperature and solute concentration vary, while all other parameters remain constant for the closed-loop simulations (e.g., pH, NaCl concentration, buffer concentration, etc.). As stated in Section 6.4, the proposed modeling and control methods can be easily extended to the case of multiple inputs since they do not depend on the specific number of the manipulated variables. To account for a number of practical considerations, several constraints are added to the control problem. First, there is a constraint on the range of variation of

the temperature to ensure validity for the kMC model by imposing a temperature range that will not damage the protein crystal. Specifically, $4^{\circ}\text{C} \leq T \leq 25^{\circ}\text{C}$ is the working range for temperature values. Second, another constraint is imposed on the rate of change of the temperature to account for actuator limitations. The controller can change temperature at a maximum of 2°C/min. The other constraint restricts the number of crystals nucleated during, for example, the second half of the batch run, which limits the nucleation of very small crystal fines. The control action (optimal temperature) at time t is obtained by solving a finite-dimensional optimization problem in a receding horizon fashion. The cost function in the optimal control problem includes a penalty on the deviation of $\langle \alpha \rangle$ from its desired crystal shape. It also includes penalty costs on the negative deviation of the crystal size when the crystal size is less than the minimum. This helps to prevent crystal fines at the end of the crystallization process. For a given solute concentration and temperature, crystal growth during the time interval is estimated by using Fig. 6.12, which was previously obtained from the open-loop simulations. The MPC problem is formulated as follows:

$$
\begin{aligned}
&\underset{T_1,\ldots,T_i,\ldots,T_p}{\textit{minimize}} \quad \sum_{i=1}^{p} F_{\langle\alpha\rangle,i} + F_{h_{110},i} + F_{h_{101},i} \\
&\text{subject to} \quad F_{\langle\alpha\rangle,i} = \left(\langle\alpha\rangle - \alpha_{\text{set}}\right)^2 \\
&\qquad F_{h_j,i} = \begin{cases} \frac{h_{j,min} - \langle h_j(t_i)\rangle}{h_{j,min}} & \text{for} \quad \langle h_j(t_i)\rangle < h_{j,min} \\ 0 & \text{for} \quad \langle h_j(t_i)\rangle \geq h_{j,min} \end{cases} \\
&\qquad \alpha_i = f_\alpha(T, C, t_i) \\
&\qquad G_j(t_i) = f_j(T, C, t_i) \\
&\qquad T_{\min} \leq T_i \leq T_{\max} \\
&\qquad \left|\frac{T_{i+1} - T_i}{\Delta}\right| \leq R_T \\
&\qquad n(0,0,t) \leq n_{\text{limit}} \qquad \forall t \geq t_f/2 \\
&\qquad \langle h_j(t_i)\rangle = \frac{\langle h_j(t_{i-1})\rangle n(h_{110}, h_{101}, t_{i-1})}{n(h_{110}, h_{101}, t_i)} + R_j(t_{i-1})\Delta \\
&\qquad n(h_{110}, h_{101}, t_i) = n(h_{110}, h_{101}, t_{i-1}) + n(0,0,t_{i-1}) \qquad \forall i \\
&\qquad i = 1, 2, \ldots, p \\
&\qquad j \in \{110, 101\}
\end{aligned}
\tag{6.22}
$$

where t is the current time, t_i, $i = 1, 2, \ldots, p$, is the time of the i^{th} prediction step, $t_i = t + i\Delta$, respectively, t_f is the length of the batch simulation, F_i is the cost function expressing the deviation of $\langle\alpha\rangle$ from its set-point ratio, α_{set}, $F_{h_{110},i}$ and $F_{h_{101},i}$ are the cost functions expressing the penalty on the negative deviation of $\langle h_{110}\rangle$ and $\langle h_{101}\rangle$ from its minimum crystal size, $h_{110,min}$ and $h_{101,min}$, at time t_i, p is the number of prediction steps, $p\Delta$ is the speci-

fied prediction horizon, $T_i, i = 1, 2, \ldots, p$, is the temperature at the i^{th} step, $T_i = T(t + i\Delta)$, respectively, $T_{\min}$ and $T_{\max}$ are the lower and upper bounds on the temperature, respectively, R_T is the limit on the rate of change of the temperature, n_{limit} limits the number of crystals nucleated during the latter half of the simulation time. The number of crystals, $n(h_{110}, h_{101}, t_i)$, and average height of the crystal face j, $\langle h_j(t_i) \rangle$, at time t_i are updated at every sampling time through the recursive equations [cf. Eq. (6.22)], respectively. The optimal set of control actions $(T_1, T_2, \ldots, T_p)$, is obtained from the solution of the multivariable optimization problem of Eq. (6.22), and only the first value of the manipulated input trajectory, T_1, is applied to the protein crystallization process from time t until the next sampling time. At this point a new measurement of protein concentration in the continuous phase is received from the kMC simulation and the MPC problem of Eq. (6.22) is re-solved for the computation of the next optimal input trajectory. The physical properties of the system (i.e., protein solubility and so on) were obtained from experimental data for the lysozyme protein solution [14]. In a previous work [117], empirical expressions were obtained for the growth rates by fitting algebraic expressions to the available experimental data. In the present work, the growth rates are computed following the kMC methodology described previously. Furthermore, the stochastic nature of the system and the model uncertainty will be accounted for in the protein concentration variations. For results on robust-control of crystallization systems and model predictive control see Ref. [116] and Ref. [19], respectively.

6.8 Batch crystallization under closed-loop operation

In this section, the proposed model predictive controller of Eq. (6.22) is applied to the kMC model described previously. At each sampling time (1 second), the optimal temperature, obtained by solving the optimization problem of Eq. (6.22), is applied to the closed-loop system until the next sampling time. The optimization problem is solved via a local constrained minimization algorithm using the nonlinear algebraic models described previously (cf. Fig. 6.12) to predict the dependence of the crystal growth rate and growth rate ratio on temperature and solute concentration, respectively.

The solute concentration randomly varies following the Gaussian distribution of Eq. (6.16) from Section 6.3. Also, it is important to note that the concentration variation results in a change of the attachment rate [cf. Eq. (6.1)]. For all closed-loop simulations the nominal concentration is 45 mg/mL with $\sigma_g = 2.5\%$ of c_n. The maximum rate of change of the temperature is 2°C/min. The volume of the crystallization batch is 1.0L. The limit on the number of crystals nucleated during the latter half of the simulation time, $n_{\text{limit}} = 500$. The following heuristic is also taken into account: It starts from a high enough supersaturation region in the beginning that many crystals are nucleated. Then it moves on to a low enough supersaturation region to keep the rate of nucleation low in order to avoid the nucleation of the small crystal fines in the end of the batch run. This optimal strategy is taken into consideration in this work as one of the constraints restricting the number of crystals nucleated during the latter half of the simulation time. See e.g., Figs. 6.13(a) and 6.15(a). Since the MPC formulation uses steady-state growth rates (cf. Fig. 6.12(a) and 6.12(b)), the number

of prediction steps is set to be $p = 1$. The time interval between the two sampling times is 1 second. The prediction horizon of each step is fixed at $p\Delta = 1$ second. The solute concentration varies every 1 second. The computational time that is used to solve the optimization problem with the current available computing power is negligible with respect to the sampling time interval. The closed-loop simulation duration $t_f = 4000$ seconds.

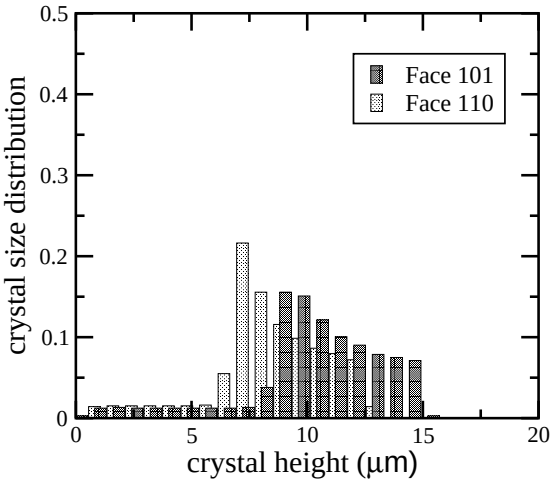
In the closed-loop simulations associated with controlling the growth rate ratio to the desired set-point values, the control objective is to separately regulate the expected ratio to the desired values, $\langle\alpha\rangle = 0.82$ and $\langle\alpha\rangle = 1.10$, respectively. Thus, the cost function of this problem contains a penalty on the deviation of the expected growth rate ratio from the set-point value. Also, the minimum thickness constraints described previously were imposed on each face to avoid undesirable crystal sizes.

The results of the closed-loop simulations are shown in Figs. 6.13-6.18. Specifically, Figs. 6.13 and 6.14 show results for six simulated batch runs under the different set-point growth rate ratios ($\langle\alpha\rangle = 0.82$ and 1.10) and different starting temperatures ($T_o = 5^\circ$C, $T_o = 15^\circ$C, and $T_o = 23^\circ$C). We note we chose three starting temperatures, one at both extremes and one in the middle, in addition to two growth rate ratios at opposite ends to show a wide range of testing that can help guide practitioners while looking to achieve a desired shape and size distribution. In Fig. 6.17, the lower desired crystal growth rate ratio, $\langle\alpha\rangle = 0.82$, has shown that final crystal shape distribution is very narrow when the initial temperature is close to the optimal temperature. Depending on the growth rate ratio set-point value and the initial temperature of the crystallizer, it takes a different amount of

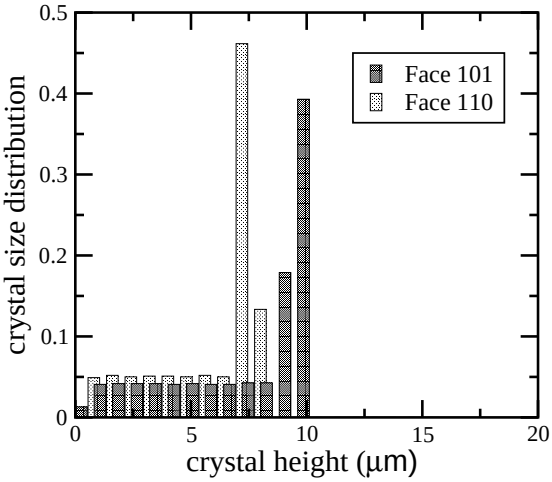
time until the temperature reaches the desired set-point value as it is seen in Figs. 6.13 and 6.15. In this case, the optimal temperature is $\sim$23°C. Figs. 6.13(a) and 6.15(a) display that starting from a low initial temperature leads to nucleation of most of the crystals within the first 600 seconds of the batch simulation and thus a narrow size distribution but, in Fig. 6.17, a broad shape distribution is achieved. In Figs. 6.13(b) and 6.15(b), since the initial temperature is closer to the optimal temperature, the system reaches its optimal temperature faster resulting in a less narrow size distribution and, in Fig. 6.17, a less broad shape distribution. Lastly, 6.13(c) and 6.15(c) display that the system reaches its optimal temperature very early and hence the size distribution is broader but, Fig. 6.17, the shape distribution becomes very narrow.

In contrast to $\langle\alpha\rangle = 0.82$, the case of the higher desired crystal growth rate ratio, $\langle\alpha\rangle = 1.10$, has shown that the sensitivity of the system to the initial temperature is much less and the final crystal shape distribution is extremely narrow when the growth rate ratio set-point is 1.10 as it is seen in Fig. 6.18. The reason is that the system reaches its optimal temperature faster, and thus the crystals uniformly nucleated along the batch go through the optimal temperature from the beginning. As stated earlier, a higher growth rate ratio as well as a higher nucleation rate is obtained at a low temperature region.

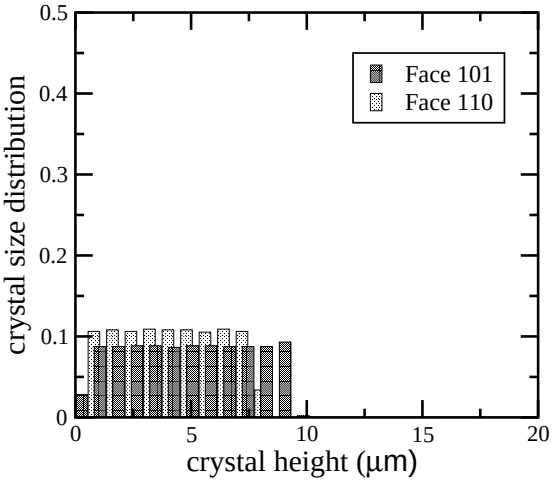
In particular for Figs. 6.14 and 6.16 the optimal temperature is $\sim$13°C. The system reaches its optimal temperature faster, and crystals are uniformly nucleated along the batch and thus a broad size distribution but, in Fig. 6.18, a narrow shape distribution is achieved. We note that if the temperature reaches its optimal state earlier in the batch simulation,



(a) $\langle \alpha \rangle = 0.82 \quad T_o = 5.0°\text{C}$

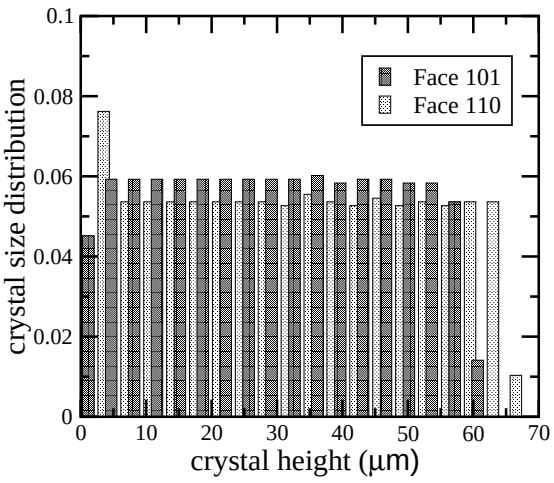


(b) $\langle \alpha \rangle = 0.82 \quad T_o = 15.0°\text{C}$

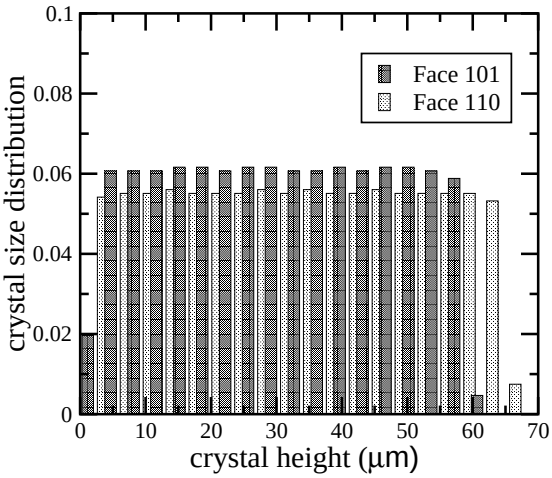


(c) $\langle \alpha \rangle = 0.82 \quad T_o = 23.0°\text{C}$

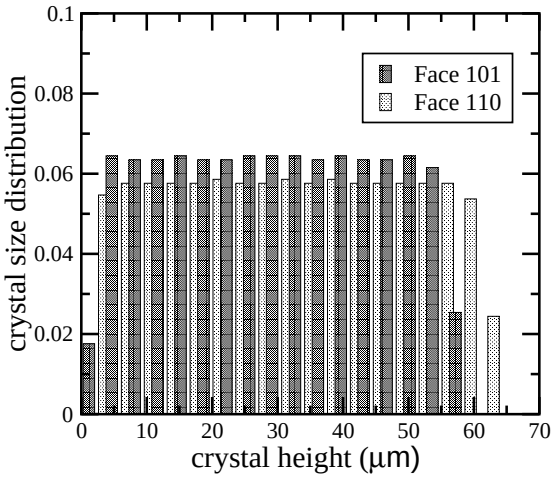
Figure 6.13: The final crystal size distribution at the end of the batch simulation for each face under closed-loop operation starting from different initial temperature and aiming at growth rate ratio set-point value, $\langle \alpha \rangle = 0.82$. It is noted that the crystal size distribution is a dimensionless variable and is normalized over the entire crystal population so that summing over all histogram bars for each face will add up to 1, for each different set of growth rate ratio and starting temperature.



(a) $\langle \alpha \rangle = 1.10 \quad T_o = 5.0°\text{C}$

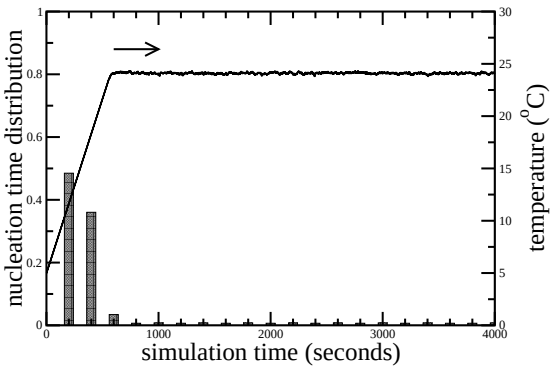


(b) $\langle \alpha \rangle = 1.10 \quad T_o = 15.0°\text{C}$

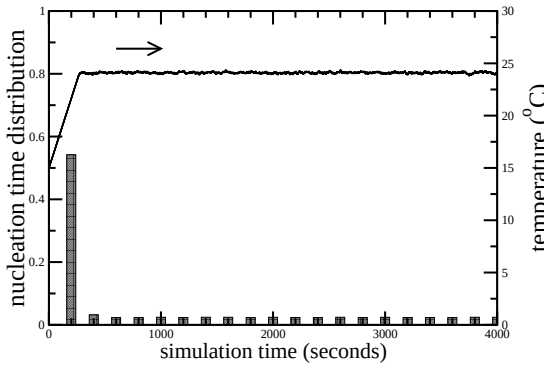


(c) $\langle \alpha \rangle = 1.10 \quad T_o = 23.0°\text{C}$

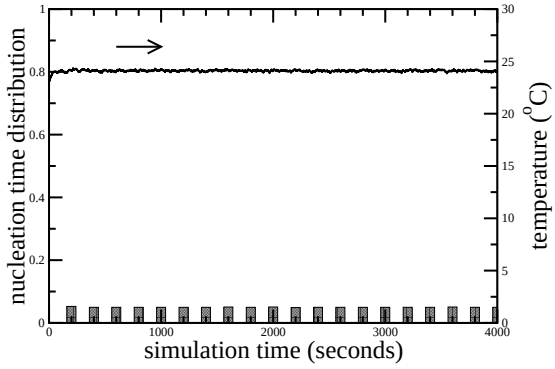
Figure 6.14: The final crystal size distribution at the end of the batch simulation for each face under closed-loop operation starting from different initial temperature and aiming at growth rate ratio set-point value, $\langle \alpha \rangle = 1.10$. It is noted that the crystal size distribution is a dimensionless variable and is normalized over the entire crystal population so that summing over all histogram bars for each face will add up to 1, for each different set of growth rate ratio and starting temperature.



(a) $\langle \alpha \rangle = 0.82 \quad T_o = 5.0°\text{C}$

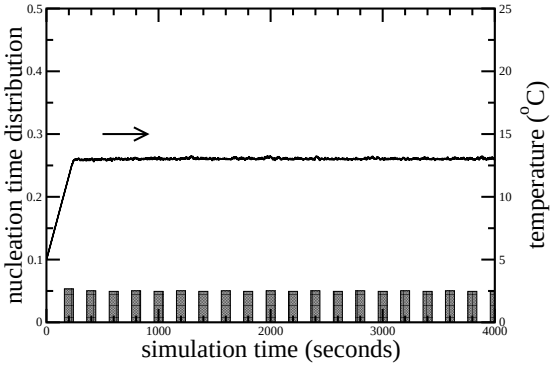


(b) $\langle \alpha \rangle = 0.82 \quad T_o = 15.0°\text{C}$

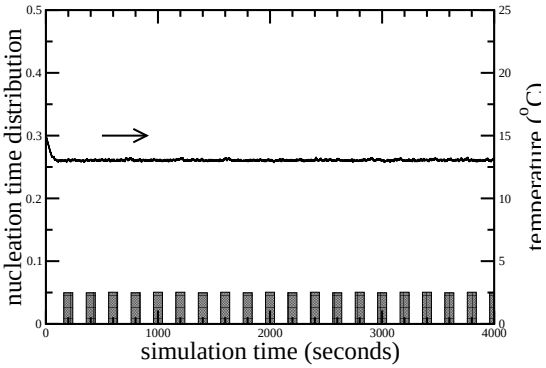


(c) $\langle \alpha \rangle = 0.82 \quad T_o = 23.0°\text{C}$

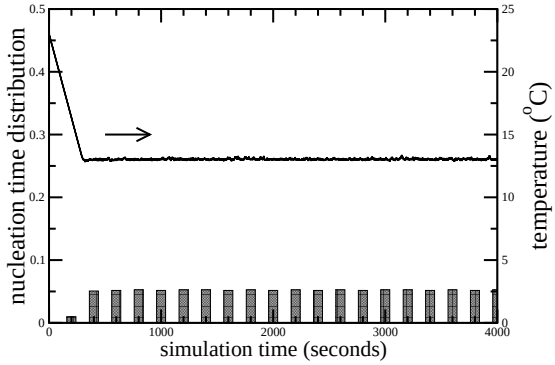
Figure 6.15: Profiles of nucleated crystals and temperature with time during the batch run under closed-loop operation for varying initial temperature and for the growth rate ratio set-point value, $\langle \alpha \rangle = 0.82$. It is noted that the nucleation time distribution is a dimensionless variable and is normalized over the entire crystal population so that summing over all histogram bars, for each different set of growth rate ratio and starting temperature, will add up to 1.



(a) $\langle \alpha \rangle = 1.10 \quad T_o = 5.0°\text{C}$



(b) $\langle \alpha \rangle = 1.10 \quad T_o = 15.0°\text{C}$



(c) $\langle \alpha \rangle = 1.10 \quad T_o = 23.0°\text{C}$

Figure 6.16: Profiles of nucleated crystals and temperature with time during the batch run under closed-loop operation for varying initial temperature and for the growth rate ratio set-point value, $\langle \alpha \rangle = 1.10$. It is noted that the nucleation time distribution is a dimensionless variable and is normalized over the entire crystal population so that summing over all histogram bars, for each different set of growth rate ratio and starting temperature, will add up to 1.

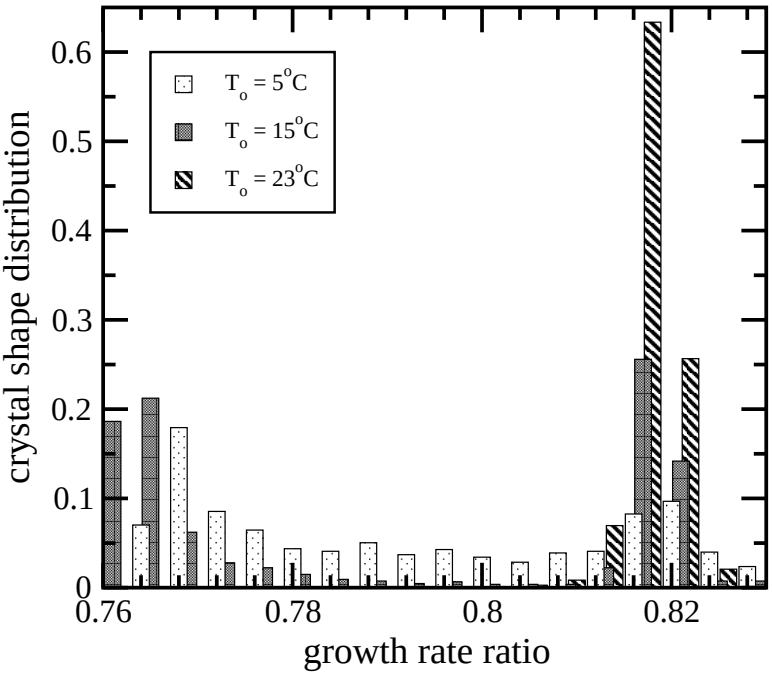


Figure 6.17: The final crystal shape distribution at the end of the batch simulation under closed-loop operation for varying initial temperature and for the growth rate ratio set-point value, $\langle \alpha \rangle = 0.82$. It is noted that the crystal shape distribution is a dimensionless variable and is normalized over the entire population so that summing over all histograms will add up to 1 for each starting temperature.

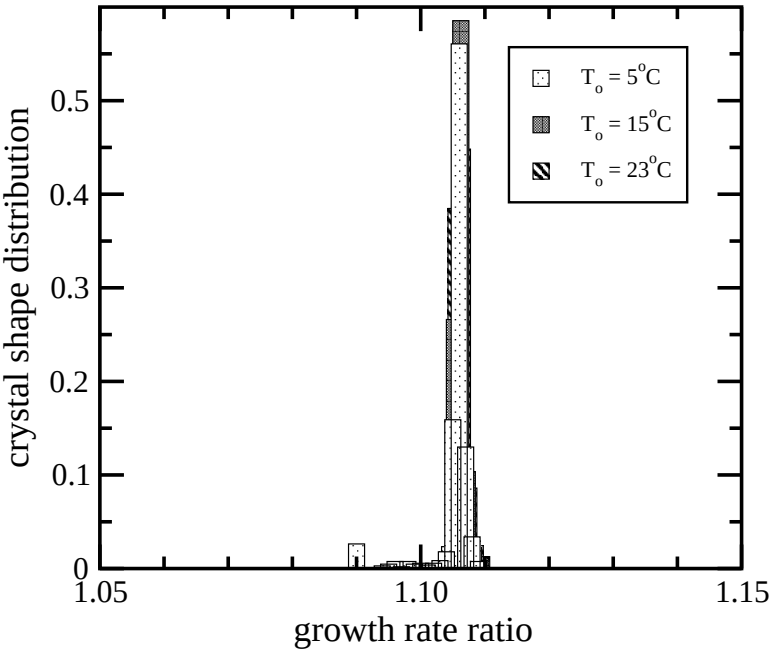


Figure 6.18: The final crystal shape distribution at the end of the batch simulation under closed-loop operation for varying initial temperature and for the growth rate ratio set-point value, $\langle \alpha \rangle = 1.10$. It is noted that the crystal shape distribution is a dimensionless variable and is normalized over the entire crystal population so that summing over all histogram bars for each different starting temperature will add up to 1.

undesirable effects of the nucleation and growth rates are minimized. In addition to this, as it is shown in Fig. 6.12, variation in the solute concentration has less effect on the growth rates as well as the growth rate ratio for $\langle \alpha \rangle = 1.10$, and thus the system reaches its steady-state much faster, which is shown in Figs. 6.15 and 6.16.

Therefore, qualitatively speaking, MPC successfully drives the final crystal shape distribution to the desired set-point value, and a narrow size distribution can be achieved depending on the desired crystal morphology. For instance, for $\langle \alpha \rangle = 1.10$, it is much more likely to obtain a broad crystal size distribution, because the system reaches its optimal temperature faster regardless of the starting temperature. Although disturbance results in the fluctuation of the solute concentration, the model predictive controller can successfully drive the growth rate ratio to the desired set-point. The desired distribution of the protein crystal shape can be estimated from Fig. 6.17 and it is slightly elongated along the (101) direction for the lower set-point ratio while it is more equidimensional for the higher set-point ratio, shown in Fig. 6.18.

6.9 Conclusions

The present work focuses on the application of modeling, simulation, and control to a batch protein crystallization process in order to first model a single lysozyme protein crystal, and then model multiple lysozyme protein crystals with the consideration of crystals nucleated at different times in the batch simulation. Based on the assumption that the two indepen-

dent crystal faces are the (110) and (101) faces, dependence of the crystal growth of the two faces on temperature and protein solute concentration, obtained from kMC open-loop simulations, was observed. The crystal shape of the resulting lysozyme crystals as well as the nucleation rate was controlled through temperature variations. This was achieved via a nonlinear steady-state model generated from open-loop kMC simulations. Moreover, the nucleation rate data was obtained from experiments and is a function of supersaturation. The nonlinear model captures the protein solute concentration and temperature dependence of the growth rate ratio, thereby describing the key elements of the protein crystallization process. The other key achievement of the present work is that it takes nucleation into account resulting in different crystals nucleated at different times throughout the batch simulation.

An MPC strategy, which uses the steady-state model, was then designed to drive the final CSD to the desired set-point value while satisfying constraints on the magnitude and rate of change of temperature. As mentioned previously, the temperature is chosen as the manipulated input and this is in accordance with standard batch crystallization practice. The minimum thickness for each face is also imposed to avoid crystal fines. Simulation results demonstrated that the proposed controller was able to control the final crystal shape and size distribution by appropriately manipulating the temperature. The present methodology shows that more than 90% of the final crystals can be produced at the desired shape, either cubic or elongated crystals, as shown in Figs. 6.17 and 6.18. Moreover, in order to implement this strategy in practice only protein solute concentration and temperature

measurements are necessary. Once these measurements are available, the controller uses the algebraic model connecting the protein solute concentration and the temperature to the crystal growth rate ratio to compute the control action (i.e., temperature change); no crystal size/shape measurements are needed in the controller. These temperature changes are feasible due to the constraints put onto the controller.

Furthermore, the proposed kMC simulation offers a number of advantages compared to previous work. In previous simulations, such as Ke *et al.* [61], they only considered the crystal seeds to initiate the crystallization. In contrast, in the present simulation, the nucleation rate was extracted from experiments and an appropriate region was picked and tested to model the experimental data. In the present work, a controller is designed to drive existing crystals as well as newly formed nuclei to desired crystal shape and size.

Chapter 7

Modeling and Control of Ibuprofen Crystal Growth and Size Distribution

7.1 Introduction

In Chapter 6, crystal shape modeling and control for lysozyme was investigated. Similar methodology is extended in this chapter to model and control the batch crystallization of ibuprofen. Once again it is emphasized that crystallization is a key separation process in the pharmaceutical industry which is estimated to be over a $1 trillion per year industry. It is used for drug purification, separation, and pre-formulation. A key consideration in crystallization is that in order to achieve desired crystal product quality, the process environment must be controlled appropriately. Otherwise, the target drug could lose purity, stability, and bio-availability.

Simulation techniques are valuable tools that can be used in crystal growth modeling which usually consist of equilibrium Monte Carlo (MC) and kinetic Monte Carlo (kMC) simulations, as well as molecular dynamics (MD) simulations [75]. A well-written book by Frenkel and Smit [37], in addition to a review by Rohl [111], go into detail about the development of these simulation techniques. In regards to crystallization, MD simulations are quite helpful when looking at how molecules move and how they are incorporated into the crystal, however, the length and time scales of MD simulations make them difficult to use for process modeling [75]. On the other hand, kMC simulations allow for more realistic length and time scales by using rate equations that describe different microscopic phenomena. To this end, kMC simulation methods have been widely used to simulate molecular-level phenomena like crystal nucleation, growth, and aggregation [13, 21, 22, 43, 48, 109, 110, 119, 49, 68, 70, 69]. Furthermore, kMC simulation methods have been successfully applied to compute the net crystal steady-state growth rate accounting for the dependence of the desorption and migration rates on the local surface microconfiguration. For that reason, we look to investigate the batch crystal growth process of ibuprofen, one of the most widely used non-steroidal anti-inflammatory drugs (NSAID), via kMC simulations in this work. Due to the lack of availability of primary nucleation rate data, we will consider a seeded batch crystallization process and keep the supersaturation at low enough levels that the impact of nucleation and crystal fines formation will be minimal compared to the amount of crystals seeded to the system.

RS-ibuprofen (2-(4-isobutyl-phenyl) propionic acid), $C_{13}H_{18}O_2$, is used to treat pain,

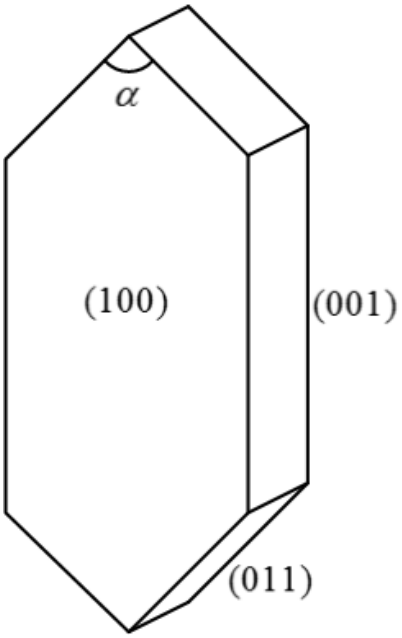


Figure 7.1: Geometry of the ibuprofen crystal. Labels show the (100), (001), and (011) faces, as well as the interfacial angle, α.

inflammation, cold, and fever. The geometry for an individual ibuprofen crystal is shown in Fig. 7.1. It works by reducing prostaglandins, which are the hormones causing inflammation and pain in the body. These are usually referred to as local hormones since they only act close to the location where they are produced. Although they are helpful initially since swelling will restrict injured areas and increased blood flow will assist in healing, longer term pain is undesirable. Thus, many different types of painkillers are used, where ibuprofen is one of the most common and widely available choices. In the US, ibuprofen brand Advil was the top over the counter (OTC) brand by revenue in 2013 with just over $490 million according to statista [58].

In this chapter, we first model the ibuprofen crystal growth process. In order to do this,

we investigate the growth rates of the (001) and (011) faces via a kMC simulation model. To account for variability in experimental crystal growth rate data, we develop a model for growth rate dispersion (GRD) since this phenomenon is known to affect ibuprofen crystal growth rates and this model is applied at the individual crystal level. After that, a seeded batch crystallizer will be considered, requiring the development of mass and energy balances for the modeling of the continuous-phase variables and this macroscopic model is combined with the microscopic crystal growth model. Finally, the crystal size distribution will be controlled by a model predictive controller (MPC) and compared against classical control strategies used in industry.

7.2 Ibuprofen Crystal Growth

7.2.1 Kinetic Monte Carlo Modeling and Simulation

In the present work, we will use kinetic Monte Carlo (kMC) simulations in order to model the growth rates of ibuprofen crystal faces since crystal growth is a non-equilibrium process. Unlike equilibrium Monte Carlo simulations, kMC simulations add an element of time by using rate equations representing different microscopic phenomena. Furthermore, this allows modeling the dependency of the crystal growth rates on the surface microconfiguration, in addition to the ability to consider individual crystals, thereby allowing for a more realistic model for growth rate dispersion. Ibuprofen has unit cell parameters of $a = 14.397$Å, $b = 7.818$Å, $c = 10.506$Å, and $\beta = 99.70^{\circ}$ with four molecules per unit

cell [114, 115]. For this work, we will consider an $N \times N$ lattice with one molecule per lattice site and periodic boundary conditions. The types of microscopic events we consider in our kMC simulations in order to model the crystal growth are adsorption, desorption and migration. Nearest neighbor lists will be used to aide the computational efficiency when calculating the total rates for each of the microscopic phenomena [20]. The rate equations considered in this work are set up similar to that of Durbin and Feher for lysozyme [27], however, they have been modified to account for available growth rate data of ibuprofen on the (001) and (011) faces [90]. Cano et al. [16] reported data for all three faces (i.e., (001), (011), and (100)), however they conducted their experiments at very low supersaturation ($\sigma = 0.013$). If more data becomes available in the future for the (100) face, then the dynamics of the (100) face can easily be integrated into the present kMC simulation model.

7.2.2 Rate Equations

The per-site adsorption rate is defined as:

$$r_a = K_a \cdot \sigma, \tag{7.1}$$

where K_a is the adsorption coefficient and σ is the relative supersaturation of the system defined by Eq. (7.2) below:

$$\sigma = \frac{\frac{I}{E} - \frac{I^*}{E}}{\frac{I^*}{E}}, \tag{7.2}$$

where I is the ibuprofen content, E is the ethanol content, and $\frac{I^*}{E}$ is the solubility. The solubility will be taken from [106, 107] and is defined as

$$\frac{I^*}{E} = 0.497 + 0.001026 \cdot T^2, \tag{7.3}$$

with temperature T defined in degrees Celsius. Since we consider an $N \times N$ lattice model, the total rate of adsorption is simply

$$W_a = N^2 \cdot r_a. \tag{7.4}$$

Unlike adsorption, the rates of desorption and migration will be dependent on the local environment at each lattice site (i.e., number of nearest neighbors to this site). When a particle has a high number of nearest neighbors, a lower desorption/migration rate will be associated with this site due to the fact that the particle is more stable in its current location. Likewise, when a particle has very few or no nearest neighbors, that particle will have a higher desorption/migration rate. Thus, we will use an Arrhenius type equation for the per-site rate of desorption, r_d, which is defined as follows:

$$r_d(i) = K_d \cdot \exp\left(-i\frac{E_{pb}}{k_B T}\right), \tag{7.5}$$

where K_d is the desorption coefficient, i is the number of nearest neighbors for the current lattice site ranging from zero to four (N, S, E, W directions), E_{pb} is the binding energy per bond, k_B is the Boltzmann constant, and T is the temperature in Kelvin. In order to find the total rate of desorption, we sum over the entire lattice. This can be done in a simple way by taking advantage of the fact that there are five different types of local environments, rather

than checking each individual lattice site requiring an $\mathrm{O}\left(N^2\right)$ calculation. Thus, the total rate of desorption is

$$W_d = \sum_{i=0}^{4} W_{d_i}; \quad W_{d_i} = M_i \cdot r_d(i), \tag{7.6}$$

where W_{d_i} is the total rate of desorption for lattice sites with i nearest neighbors and M_i is the number of lattice sites with i nearest neighbors. Migration works in an analogous way and is defined as follows:

$$r_m(i) = K_m \cdot \exp\left(-i\frac{E_{pb}}{k_B T}\right), \tag{7.7}$$

$$W_m = \sum_{i=0}^{4} W_{m_i}; \quad W_{m_i} = M_i \cdot r_m(i), \tag{7.8}$$

where r_m is the per-site rate of migration, K_m is the migration coefficient, W_m is the total rate of migration, and W_{m_i} is the total rate of migration for lattice sites with i nearest neighbors. Lastly, the amount of time elapsed when an event occurs is defined in the following way:

$$\Delta t = -\ln(1-\zeta)/W_{\mathrm{tot}}, \tag{7.9}$$

where ζ is a uniform random number, i.e., $\zeta = [0,1)$, and $W_{\mathrm{tot}} = W_a + W_d + W_m$.

7.2.3 Growth Rate Dispersion

Growth rate dispersion (GRD) is a well-known phenomenon where crystals of the same type, undergoing seemingly the same conditions, grow at different rates [40, 129, 60, 104, 42, 94]. Previous models that describe this process include the constant crystal growth (CCG) model [40, 72], the random fluctuation (RF) model [104], and the fast growers, slow growers (FGSG) model [23, 113, 41]. In the CCG model, a distribution of crystals has a

distribution of growth rates and individual crystals adhere to a specific growth rate from that distribution during the entire period of growth [72]. The RF model requires individual crystal growth rates to fluctuate around an average value. Lastly, the FGSG model states that small crystals will grow at lower growth rates compared to the larger ones. In the present work, we account for GRD in a way that is similar to the CCG model by randomly giving each crystal a uniform random number, ζ_{GRD}, at the start of the simulation which will be used to calculate the GRD factor, GRD_f in the following way:

$$\mathrm{GRD}_f = 2\frac{C_{\mathrm{GRD}}}{\sigma} \cdot \zeta_{\mathrm{GRD}} + \left(1 - \frac{C_{\mathrm{GRD}}}{\sigma}\right), \tag{7.10}$$

where C_{GRD} is the GRD constant and will be calculated to fit experimental data of ibuprofen crystal growth rate dispersion. It is noted that GRD_f is dependent on σ due to the fact that error bars became too small at lower supersaturation values and too large at higher supersaturation values when fitting to experimental data without having this dependence. The results of this fit is presented in the next subsection. The GRD factor will affect the rate of adsorption (i.e., each crystal will have a slightly different rate of adsorption depending on the ζ_{GRD} assigned to that crystal at the start of the simulation). This will allow for variation in the growth rates in a manner consistent with the experimentally computed values and it will be explicitly defined in the following way in this work:

$$\sigma_{\mathrm{GRD}} = \sigma \cdot \mathrm{GRD}_f. \tag{7.11}$$

It is noted that in order for this change to take place, σ_{GRD} will replace σ in Eq. (7.1) to give

$$r_a = K_a \cdot \sigma_{\mathrm{GRD}}. \tag{7.12}$$

7.2.4 Fitting the kMC Model Parameters to Experimental Results

For a given set of simulation conditions comprised of temperature, ibuprofen content, ethanol content, and water content, the kMC simulation methodology and GRD model described earlier in this section result in growth rate values of ibuprofen for the (001) and (011) faces over a range of supersaturations. In Fig. 7.2, ibuprofen crystal growth rates are modeled at 95% ethanol, $\frac{I}{E} = 2$, and a relative supersaturation range of $0.68 \lesssim \sigma \lesssim 1.20$. The growth rates at each point in the kMC are produced by averaging 640 independent crystal runs with the error bars representing the standard deviation. Results are compared to experimental growth rates at 95% ethanol from [90], as well as a best fit line given by [108] which has the equation $G = k_G \cdot s$, where $k_G = 15$ and $s = \frac{I}{E} - \frac{I^*}{E}$.

The model parameters used for the kMC simulations are listed in Table 7.1. Additionally, C_{GRD} was found to be 0.07 resulting in an average coefficient of variation (CV) for the kMC simulation data to be 0.14, compared to 0.12 given in [90]. Also, the kMC growth rate data for the (001) and (011) faces were fit using a least squares linear regression model which will be used later in the model predictive controller. The result of this fit are

$$G_{001} = 24.843 \cdot \sigma - 15.564, \tag{7.13}$$

and

$$G_{011} = 24.412 \cdot \sigma - 7.2772. \tag{7.14}$$

It is noted that size effects of the lattice were considered and results from $N = 15$ and larger were consistent and showed no change in results. So, for this work N was set to 20 in order

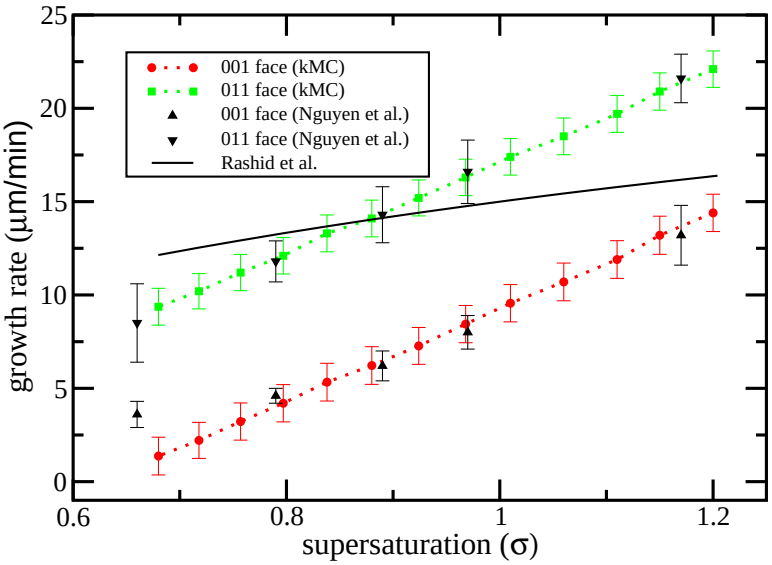


Figure 7.2: Growth rate versus supersaturation for the (001) and (011) faces for the kMC model. Additionally, the experimental results from Nguyen et al. [90] and the trendline from Rashid et al. [108] are shown.

to ensure consistency without being too large, thereby causing an exponential increase in required simulation time.

Parameter	Value	Units
E_{pb}/k_B (001)	17.47	K
E_{pb}/k_B (011)	125.20	K
K_a	380	sec^{-1}
K_d	270	sec^{-1}
K_m	300	sec^{-1}

Table 7.1: Parameters for faces (001) and (011) at $\frac{I}{E} = 2$.

7.3 Batch Crystallization

7.3.1 Energy and Mass Balance Equations

The energy and mass balance equations which calculate the change in temperature, T, and ibuprofen content, I, are given by the following ordinary differential equations:

$$\frac{dT}{dt} = -\frac{\rho_c \Delta H_c}{\rho_{\text{slurry}} C_p V_{\text{slurry}}} \cdot \frac{dV_c}{dt} - \frac{U_j A_j}{\rho_{\text{slurry}} C_p V_{\text{slurry}}} \cdot \left(T - T_j\right), \quad T(0) = T_0, \tag{7.15}$$

$$\frac{dI}{dt} = -\rho_c \frac{dV_c}{dt}, \quad I(0) = I_0, \tag{7.16}$$

where ρ_c is the density of the crystal phase, ΔH_c is the enthalpy of crystallization, ρ_{slurry} is the density of the slurry phase, C_p is the specific heat capacity, V_{slurry} is the volume of

the slurry phase, V_c is the total volume of all the crystals, t is the time, U_j and A_j are the overall heat transfer coefficient and area between the jacket stream and the crystallizer, respectively, and T_j is the temperature of the jacket stream. Additionally, T_0 and I_0 are the starting temperature and ibuprofen content of the batch system, respectively. The values for these parameters are given in Table 7.2.

Parameter	Description	Value	Units
ρ_c	crystal density	1030	mg/cm^3
ΔH_c	enthalpy of crystallization	-112.95 [90]	kJ/kg
ρ_{slurry}	slurry density	485-510	mg/cm^3
C_p	specific heat capacity	1.85-2.0	J/g·K
A_j	surface area of crystallizer wall	0.25 [116]	m^2
U_j	heat transfer coefficient of crystallizer wall	1800 [116]	kJ/m^2· h·K

Table 7.2: Parameters for faces (001) and (011) at $\frac{I}{E} = 2$. Please note that ranges are given for the slurry density and specific heat capacity since they are calculated by composition of the slurry throughout the entire simulation.

Volume Calculation

In order to properly calculate the mass and energy balance terms that require volume change information, we first need to accurately estimate the volume of the ibuprofen crystals. In order to do this, we need to know the height for all three faces [i.e., (001), (011), and (100)],

along with the interfacial angle α. Since we explicitly model the growth rates for the (001) and (011) faces, we can easily determine the heights of the (001) and (011) faces. On the other hand, for the (100) face, we will use visual approximation from [105] to estimate its relative height. The results of this approximation show that the (100) face is roughly 4 to 8 times slower growing than the (001) face. Thus, we will assume:

$$h_{100} \cong \frac{h_{001}}{6}, \tag{7.17}$$

where h_{100} and h_{001} are the heights of the (100) and (001) faces, respectively. Second, we will use the images provided in [90] in order to measure the interfacial angle, α, as a function of supersaturation. Using these images, we found the following relationship:

$$\alpha = -14.368^{\circ} \cdot \sigma + 105.41^{\circ}. \tag{7.18}$$

With the use of Eqs. (7.17) and (7.18), the volume of each crystal (see e.g., Fig. 7.1) can now be calculated in the following way:

$$V_{\text{crystal}} = \frac{4h_{001}}{\sin\left(\frac{\alpha}{2}\right)} \left(2h_{011} - h_{001} \cdot \cos\left(\frac{\alpha}{2}\right)\right) \cdot h_{100}. \tag{7.19}$$

7.4 Model Predictive Control

In the seeded batch crystallization process of ibuprofen, kMC simulations are considered for the crystal growth process via adsorption, desorption, and migration type microscopic surface events. The growth rates produced by these simulations are directly related to the supersaturation of the system, which can be modified by changing the temperature of

the jacket which is in contact with the batch reactor. In this section, a model predictive controller (MPC) is presented for seeded batch ibuprofen crystallization control. MPC is used in order to provide optimality, robustness, and constraint handling in the batch crystallization process [116, 117, 70]. In particular, the objective of the MPC will focus on minimizing the crystal size distribution by computing a set of optimal jacket temperatures over the length of the prediction horizon. The main reason shape control is not directly considered in this work is due to the fact that the shape of ibuprofen crystals is more dependent on the solvent choice rather than the batch temperature. Additionally, an actuator constraint on the rate of change of the jacket temperature is imposed, as well as a constraint on the temperature and supersaturation of the system so that crystallization will take place in an appropriate environment to avoid damaging the crystal. Furthermore, the growth rates will be modeled via Eqs. (7.13) and (7.14) in the MPC. Lastly, the energy and mass balance equations are considered [i.e., Eqs. (7.15) and (7.16)]. The formulation for the MPC

developed in this work is as follows:

$$
\begin{aligned}
\underset{T_{j,1},\ldots,T_{j,i},\ldots,T_{j,p}}{\text{minimize}} \quad & \sum_{i=1}^{p}\left(\left(V_{\text{set}}-\frac{M_1}{M_0}\right)/V_{\text{set}}\right)^2 \\
\text{subject to} \quad & \frac{dM_1}{dt}=G_{\text{vol}}\cdot M_0, \quad M_0=5\cdot 10^6 \\
& G_{001}=24.843\cdot\sigma-15.564, \quad G_{011}=24.412\cdot\sigma-7.2772 \\
& G_{100}=\frac{G_{001}}{6} \\
& \sigma=\frac{\frac{I}{E}-\frac{I^*}{E}}{\frac{I^*}{E}}, \quad \frac{I^*}{E}=0.497+0.001026\cdot T^2 \\
& \langle V_{\text{crystal}}\rangle=\frac{4\langle h_{001}\rangle}{\sin\left(\frac{\alpha}{2}\right)}\left(2\langle h_{011}\rangle-\langle h_{001}\rangle\cdot\cos\left(\frac{\alpha}{2}\right)\right)\cdot\langle h_{100}\rangle \\
& \langle h_k\rangle=\langle h_k\left(t_{i-1}\right)\rangle+G_k\Delta \\
& \frac{dT}{dt}=-\frac{\rho_c\Delta H_c}{\rho_{\text{slurry}}C_pV_{\text{slurry}}}\cdot\frac{dM_1}{dt}-\frac{U_jA_j}{\rho_{\text{slurry}}C_pV_{\text{slurry}}}\cdot\left(T-T_j\right), \quad T\left(0\right)=T_0 \\
& \frac{dI}{dt}=-\rho_c\frac{dM_1}{dt}, \quad I\left(0\right)=I_0 \\
& T_{\min}\le T\le T_{\max}, \quad \left|\frac{T_{j,i+1}-T_{j,i}}{\Delta}\right|\le 2.0^{\circ}\text{C/min} \\
& \sigma_{\min}\le\sigma\le\sigma_{\max} \\
& i=1,2,\ldots,p, \quad k\in\{001,011,100\}
\end{aligned}
\tag{7.20}
$$

where $p=10$ is the length of the prediction horizon, $\Delta=40$ is the sampling time in seconds, V_{set} is the desired average volume set point, $\langle V_{\text{crystal}}\rangle$ is the average volume of the crystal distribution, T_j is the jacket temperature, $T_{j,i}$ is the jacket temperature at the ith prediction step, $\langle h_k\rangle$ is the average height on face k, and M_0 and M_1 are the zeroth and first moments of the crystal size distribution, respectively. M_0 represents the total number of crystals and M_1 represents the total volume of the crystals. It is noted that since we consider a seeded batch

crystallizer without nucleation, M_0 will be constant for this work. If nucleation data was available, then it follows that M_0 would need to be a variable in the control problem given by Eq. (7.20) and more considerations would be taken to attempt to minimize the presence of crystal fines. Additionally, G_{vol} is the volumetric growth rate and is calculated by finding the change in average crystal volume. Finally, values of $\sigma_{\min} = 0.6$, $\sigma_{\max} = 1.3$, $T_{\min} = 10^{\circ}\text{C}$, and $T_{\max} = 40^{\circ}\text{C}$ are used for this work. The set of optimal jacket temperatures along the prediction horizon is obtained by solving Eq. (7.20) in a receding horizon fashion with IPOPT, an open source software package for large-scale nonlinear optimization. The first value, $T_{j,1}$, is then applied to the system until the next sampling time when a new set of optimal jacket temperatures is calculated.

7.5 Closed Loop Simulations

For the seeded batch crystallization simulations, we investigate the crystal size and shape distributions. The same initial conditions, other than starting temperature, are used in every simulation to ensure consistency. For this work, the initial conditions of the seeded batch reactor are $V_{\text{solvent}} = 400\text{mL}$ (95% ethanol), $\frac{I}{E} = 2$, $I_{\text{start}} = 6 \times 10^5\text{mg}$, and $M_0 = 5 \times 10^6$. Each of the simulations is completed when the average crystal volume reaches the set point, $V_{\text{set}} = (150\mu\text{m})^3 = 3.375 \times 10^6 \mu\text{m}^3$. Due to the nature of the batch process and the dependence of the growth rate on the supersaturation and temperature trajectories, the time to finish each simulation will vary. To deal with this, we will consider a normalized time to compare the different simulations, i.e., 0 at the start of the simulation and 1 when the batch

has reached V_{set}. Also, it is noted that the kMC simulations are run with constant batch parameters (i.e., temperature, ibuprofen content, and supersaturation) for 0.333 seconds. At that point, Eqs. (7.15) and (7.16) are calculated, all system parameters are updated, and this process is repeated until the end of the simulation.

7.5.1 MPC Performance

In this subsection, we investigate the closed-loop performance of the proposed MPC scheme to regulate the volume and shape distributions of ibuprofen crystals produced from a seeded batch crystallization process. Specifically, we look at the effect of different initial temperatures and growth rate dispersion on the size and shape distributions of ibuprofen crystals. We consider starting temperatures ranging from 15°C to 30°C with a step size of 5°C. The crystal volume distribution for each of the cases is shown in Fig. 7.3. What can be noticed is that the lower starting temperatures lead to a slightly more narrow size distribution. This is due to the fact that lower temperatures correspond to higher supersaturation values, and at these higher supersaturation values the relative effect of the GRD is less compared to the effect of GRD on lower supersaturation values (see e.g., Fig. 7.2). The differences in each of the starting conditions becomes much more noticeable in Fig. 7.4 which shows the crystal shape distribution. We define the crystal shape to be the relative average height of the (011) face to the (001) face since the (100) face is determined by Eq. (7.17). The crystal shape distribution not only becomes wider as the starting temperature gets higher, but also it shifts to the right meaning that the crystals become more elongated. Again, when looking

at Fig. 7.2, it is evident that the ratio between the (011) and (001) faces is greater at lower values of supersaturation (i.e., higher values of temperature) which results in an elongated crystal shape for the higher starting temperatures in Fig. 7.4. Looking at Fig. 7.5, we can infer a more detailed view of the dynamics of the batch crystallizer conditions. What is important to notice is that MPC is able to successfully deal with the constraints of the system (e.g., $T_{\text{start}} = 15°\text{C}$ or $30°\text{C}$ where the supersaturation starts outside of the supersaturation constraint region). Furthermore, after the MPC has changed the batch temperature from the initial starting temperature, each of the different simulations follows a path that resembles crystallizer cooling. This is done since as the crystallization progresses, ibuprofen content will go from the slurry phase to the crystal phase causing a decrease in concentration (i.e., $\frac{I}{E}$). In order to balance this effect and keep the supersaturation from falling to very low values, the temperature is lowered in order to keep the crystal growth progressing.

7.5.2 Comparison of MPC Performance With Other Control Strategies

In order to compare the performance of the proposed MPC, we performed additional simulations using constant temperature control (CTC) and constant supersaturation control (CSC) strategies. For these simulations, we chose the starting temperature $T_{\text{start}} = 20°\text{C}$ for CTC, CSC, and MPC which corresponds to a starting supersaturation $\sigma \cong 1.2$. This starting point was chosen to ensure both CTC and CSC would be in a valid operating region accounting for the desired supersaturation and temperature ranges since these control

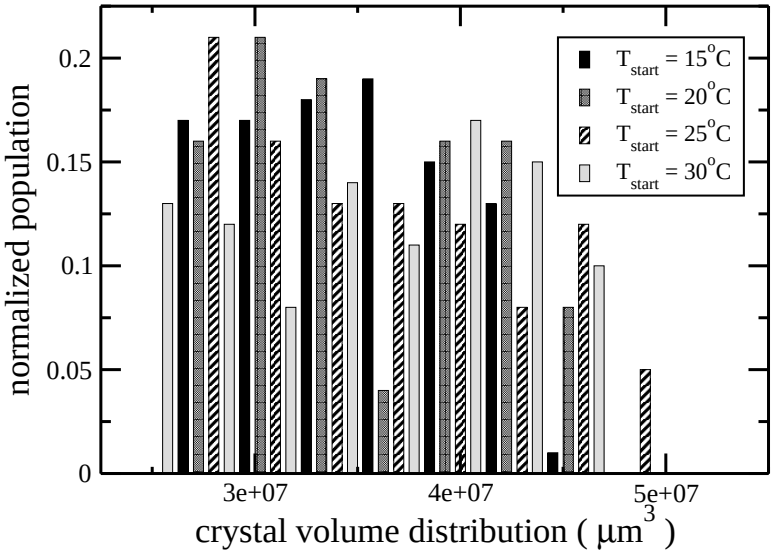


Figure 7.3: Crystal volume distribution for MPC showing results for starting temperatures 15°C, 20°C, 25°C, and 30°C.

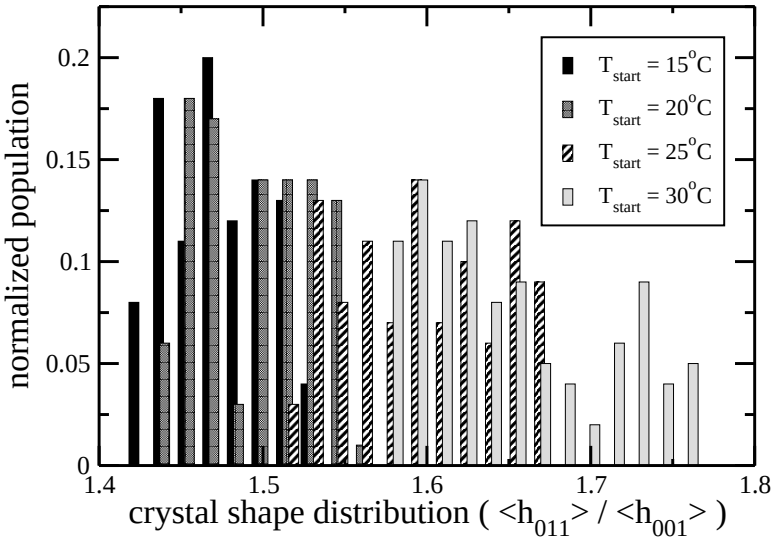


Figure 7.4: Crystal shape distribution for MPC showing results for starting temperatures 15°C, 20°C, 25°C, and 30°C.

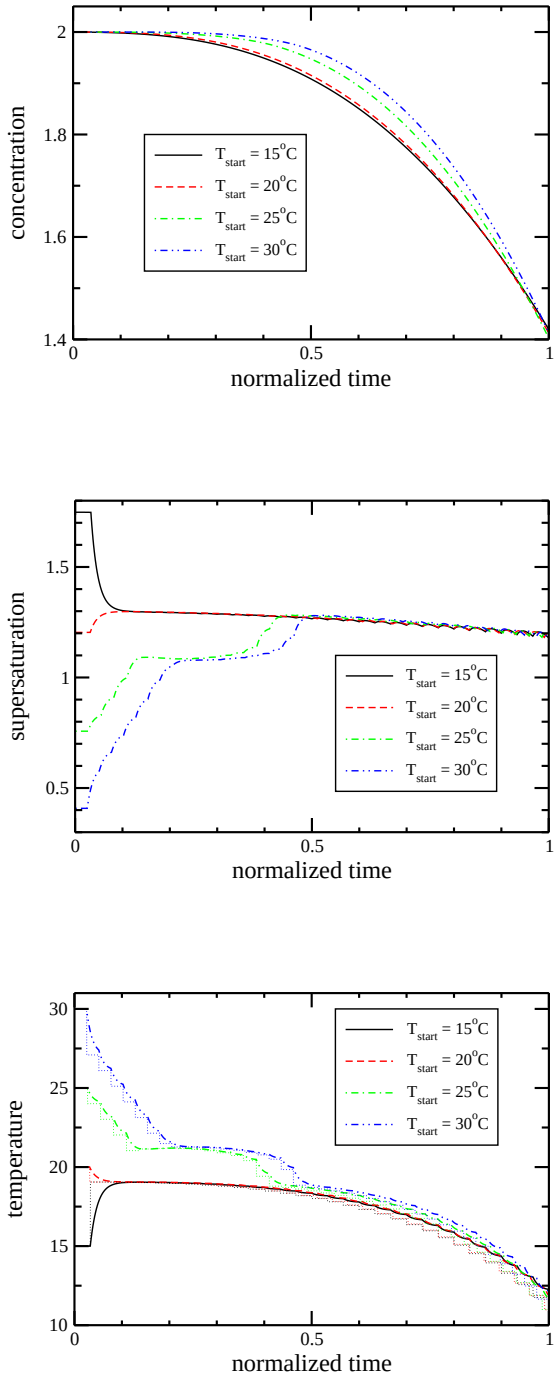


Figure 7.5: Concentration, supersaturation, and temperature versus normalized time for MPC showing results for starting temperatures 15°C, 20°C, 25°C, and 30°C. For the temperature plot, the dotted lines represent the jacket temperature, T_j, for each of the runs.

methods are unable to deal with constraints. The crystal volume distribution can be seen in Fig. 7.6. It is clear that CTC leads to the most broad crystal size distribution and it can be seen that MPC gives a slightly more narrow distribution than CSC. Similar behavior is seen in Fig. 7.7 for the shape distribution where CSC and CTC shift the crystal shape distribution to the right compared to the MPC. MPC produces the most narrow crystal size distribution due to the jacket temperature trajectory it chooses and due to its ability to work within a constrained region. Additionally, the way the MPC goes about minimizing the volume distribution also happens to produce the most narrow crystal shape distribution.

The differences in each of these policies are highlighted when looking at the dynamics of the batch reactor in Fig. 7.8. As expected, CTC holds the jacket temperature at 20°C throughout the entire simulation, however, it is noted that the supersaturation drops significantly below 1.2 in the CSC policy. This is due to the actuator constraint on T_j thereby limiting the maximum rate of change and causing the supersaturation to drop. It is also interesting to note that MPC and CSC take nearly identical pathways in terms of concentration to reach the desired set-point. Overall, MPC is able to outperform the other techniques since it is able to "plan ahead" and predict what will happen next which is especially important when there is significant concentration drop in the system.

7.5.3 Computational Performance and Scaling

To close out this section, it is important to note the computational performance and scaling for this work. In order to make this comparison, we ran the same seeded batch crystalliza-

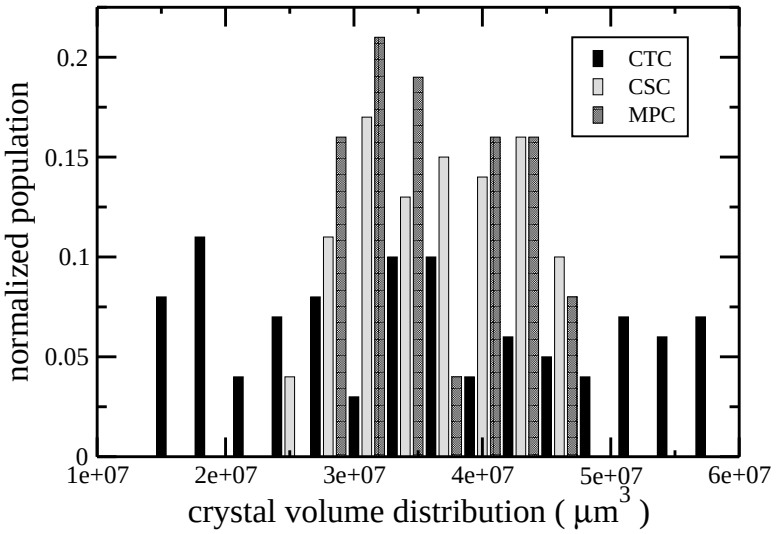


Figure 7.6: Crystal volume distribution for CTC, CSC, and MPC at the end of the batch.

tion simulation and initial conditions with different random seeds on the Texas Advanced Computing Center's Stampede cluster. The code was optimized using Message Passing Interface (MPI) over the crystal growth stage since it was determined to be the bottleneck of this simulation. Specifically, at the start of the simulation, crystals are assigned to one of the available cores. Next, the the growth process runs while the batch system parameters remain constant until it is time to update the crystallizer conditions. After these parameters are updated, the crystals will go back into the growth stage on their assigned core. This process is repeated until the end of the simulation. The results of these simulations for varying number of cores are shown in Fig. 7.9 and the data points are given in Table 7.3. What can be seen from Fig. 7.9 is that there is a significant decrease in time required to complete the batch simulation as the number of cores are increased. Looking at Table 7.3,

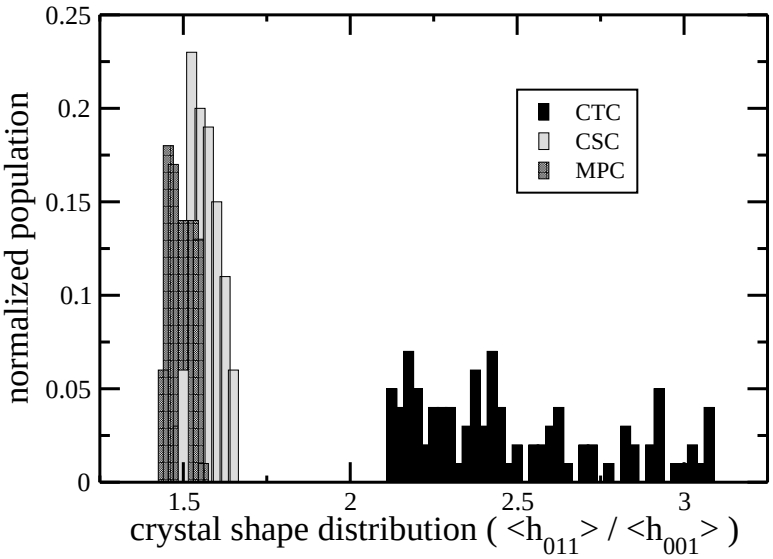


Figure 7.7: Crystal shape distribution for CTC, CSC, and MPC at the end of the batch. Please note that one bar from CSC has been placed in front of MPC due to the fact that it was completely covered by the MPC bars.

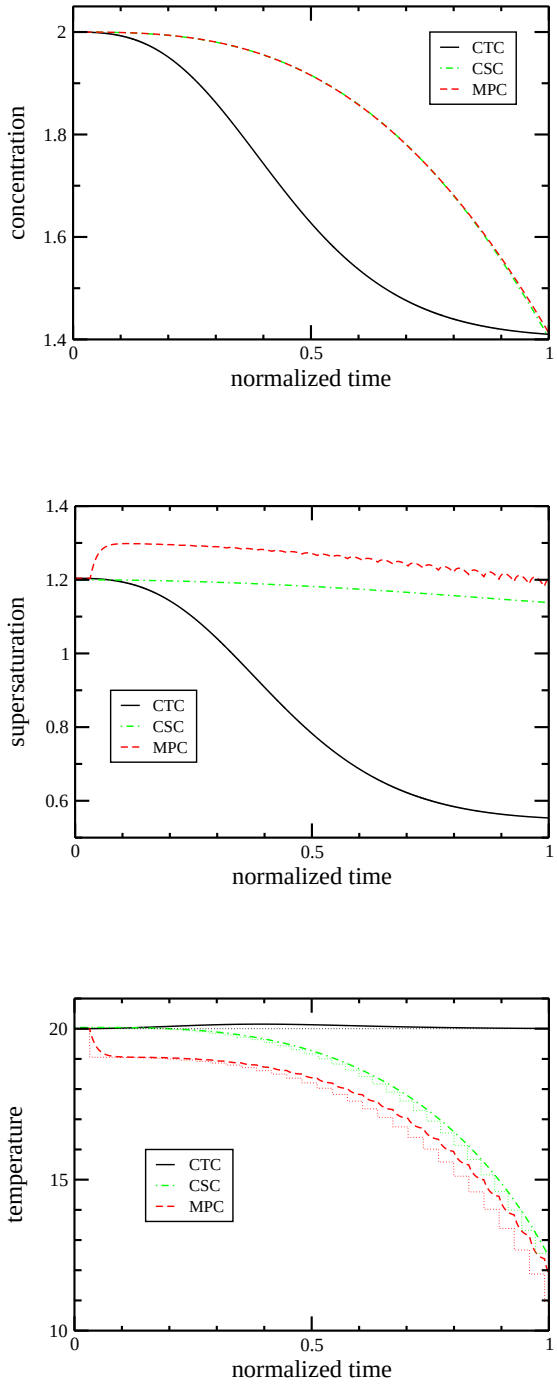


Figure 7.8: Concentration, supersaturation, and temperature versus normalized time for CTC, CSC, and MPC. For the temperature plot, the dotted lines represent the jacket temperature, T_j, for each of the runs. Additionally, it is noted that both CSC and MPC follow a very similar path in the concentration plot until the very end.

it is evident that as the number of cores is doubled, the simulation time goes down by about half. In order to further analyze the scalability of this parallel process, it is useful to analyze the strong scaling behavior, which is defined as:

$$S_{\text{strong}} = \frac{t_1}{n_{\text{cores}} \cdot t_n}, \tag{7.21}$$

where t_1 is the time the process takes on 1 core, n_{cores} is the number of cores, and t_n is the time the process takes on n cores. Strong scaling is good for analyzing systems like this one that are CPU bound, showing how well the process can be parallelized without adding too much wasted time in overhead costs. From Table 7.3, it can be seen that the strong scaling stays above 90% when using 16 or fewer cores and drops down afterwards. This is likely due to the fact that simulations were run on compute nodes which had 16 cores per node (two 8-core CPUs) and when going over 16 cores, communication must then take place between multiple nodes, thus adding overhead costs. Overall, it is clear from both Fig. 7.9 and Table 7.3 that the batch crystallization process of ibuprofen is greatly benefiting from the use of MPI for the kMC process.

7.6 Conclusions

In this chapter, we studied the seeded batch crystallization process of ibuprofen. First we used kMC simulations to develop a growth rate model which also accounts for GRD. Next, we proposed an MPC strategy in order to control the crystal size distribution. Lastly, we compared the proposed MPC strategy to CTC and CSC policies. We found that the

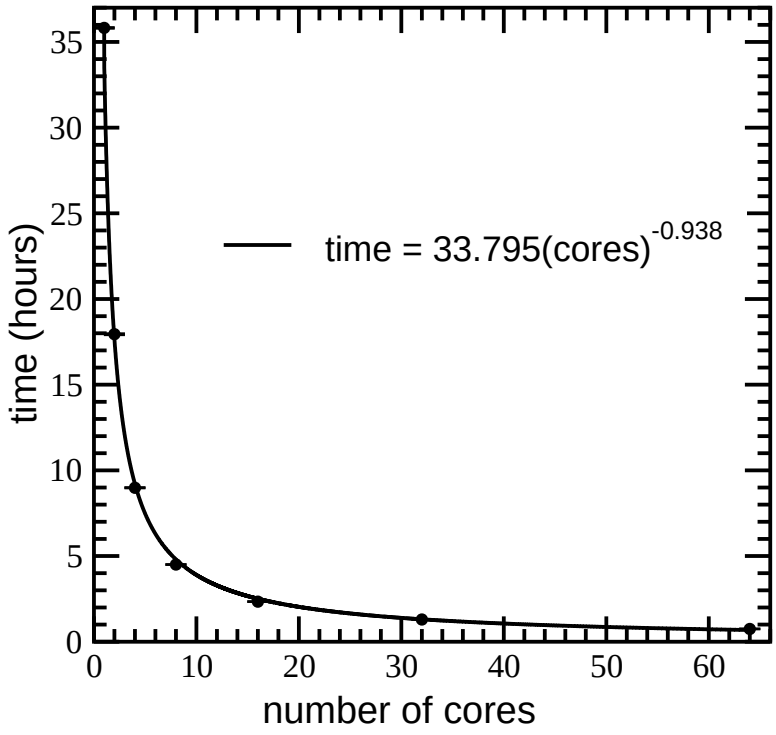


Figure 7.9: The number of cores versus the average amount of time required to finish the batch simulation. Error bars are shown as one standard deviation over 10 simulations for each batch run. The best fit line has equation: $\text{time} = 33.795 \cdot \text{cores}^{-0.938}$ with an $R^2 = 0.9982$.

cores	time (h)	speedup (%)	Strong Scaling (%)
1	35.82	0.0	100.0
2	17.95	49.9	99.8
4	8.98	74.9	99.7
8	4.50	87.4	99.4
16	2.34	93.5	95.6
32	1.30	96.4	86.2
64	0.75	97.9	74.8

Table 7.3: The time to finish each simulation for varying number of cores and the corresponding speedup and strong scaling. Please note that the speedup is defined as $\frac{t_1 - t_n}{t_1}$, where t_1 is the time the process takes on 1 core and t_n is the time the process takes on n cores.

MPC is able to deal with constraints and a wide variety of staring conditions for ibuprofen crystal growth. Additionally it was found that MPC produced more narrow volume and shape distributions compared to the other control strategies which is important because the product quality is directly determined by the final crystal size and shape distributions. It is important to note that the growth rate dispersion is mainly responsible for the wide distribution ranges seen in this work. Lastly, we found an extreme benefit in the use of MPI for this work due to high CPU time requirements.

Chapter 8

Conclusions

This dissertation presented a practical framework which uses multiscale modeling in order to model, simulate, and control crystallization processes. Following the introduction, cell models were used to find equilibrium phase diagram properties via coarse-graining techniques, equilibrium Monte Carlo simulations, and finite-size scaling theory in both systems of hard spheres and Lennard-Jones particles. Then, kinetic Monte Carlo simulations were used to deduce microscopically consistent rate laws for crystal growth. After that, these crystal growth rate laws were used in a batch crystallization process to develop a macroscale model. Finally, model predictive control was used to regulate the crystal size and shape distributions in the batch crystallizer.

Specifically in Chapter 2, the precise simulation of freezing transitions was presented. Cell models from Hoover and Ree were introduced and a generalized cell model facilitated fluid-solid transitions, thus simplifying the overall process of thermodynamic integration

by greatly reducing the number of simulated states. Additionally, coexistence pressure and densities were determined through finite-size scaling techniques for first-order phase transitions which are based on analyzing the size-dependent behavior of susceptibilities and dimensionless moment ratios of the order parameter.

Subsequently, in Chapter 3, phase transitions, criticality, and three phase coexistence using constrained cell models were investigated. In particular, the phase diagram of the constrained cell model was determined for a system of Lennard-Jones particles. It is important to understand the phase diagram of the constrained cell model in order to fully explain its behavior in relation to the unconstrained system. The pressure-density isotherms exhibited inflection points which were interpreted as the mechanical stability limit of the solid phase. The phase diagram of the constrained system contained a critical and a triple point. It is believed that entropy was reduced by the single occupancy constraint of the constrained cell model causing the temperature and pressure at the critical and triple point to be higher than the corresponding unconstrained system.

Chapter 4 looked at the direct determination of triple-point coexistence through cell model simulations. Once again, the generalized cell model was used in order to investigate coexistence. However, this time it was used to provide direct estimates of triple-point equilibria. In previous work, the triple-point was estimated by indirect methods, i.e., by locating the common intersection of the various p–T curves through extrapolations. Furthermore, we know from the earlier chapters that there will be a size-dependence for the triple-point parameters which can be resolved through finite-size scaling analysis, however

this will be a computationally demanding task due to the amount of statistical samples that are required in a single simulation as system sizes get larger.

In Chapter 5, simulations of phase boundaries using constrained cell models were investigated for a system of Lennard-Jones type particles that resembles that of protein or colloidal suspensions. The main advantage of the techniques used in this chapter, compared to the previous direct and indirect approaches, is their simplicity. The results also indicated that the range of attractive interactions had no effect on the topology of the phase diagram of the constrained cell model.

Finally, in Chapter 6 and Chapter 7, modeling and control of crystal shape in batch crystallization processes was explored for both lysozyme and ibuprofen crystals. First, kinetic Monte Carlo simulations were developed in order to model the growth rates of each crystal. The parameters of the growth rate model were assigned a range of values until satisfactory agreement with experiments was obtained for the system under consideration. Next, a batch crystallization process was studied which considered many different types of processes (e.g., nucleation, growth rate dispersion, conservation equations, etc.). The results of the batch crystallizer were then connected back to the microscopic simulations via model predictive control (MPC). Additionally, the MPC closed-loop performance for ibuprofen was compared against constant temperature control (CTC) and constant supersaturation control (CSC) policies. The proposed MPC was able to deal with the constraints of the control problem, in addition to minimizing the spread of the crystal shape and size distributions in a superior fashion compared to the other control methodologies, which im-

proves the overall crystal product quality at the end of the batch. Finally, extreme benefits were found in the use of MPI for this work due to heightened CPU time requirements.

Bibliography

[1] R. Agrawal and D. A. Kofke. Thermodynamic and structural properties of model systems at solid-fluid coexistence. 2. Melting and sublimation of the Lennard-Jones system. *Mol. Phys.*, 85:43–59, 1995.

[2] N. Aldabaibeh, M. J. Jones, A. S. Myerson, and J. Ulrich. The solubility of orthorhombic lysozyme chloride crystals obtained at high pH. *Cryst. Growth Des.*, 9:3313–3317, 2009.

[3] N. Asherie. Protein crystallization and phase diagrams. *Methods*, 34:266–272, 2004.

[4] B. A. Berg and T. Neuhaus. Multicanonical ensemble: A new approach to simulate first-order phase transitions. *Phys. Rev. Lett.*, 68:9–12, 1992.

[5] J. Berthou and P. Jolles. A phase transition in a protein crystal: The example of hen lysozyme. *Biochem. Biophys. Acta*, 336:222–227, 1974.

[6] K. Binder. Theory of first-order phase transitions. *Rep. Prog. Phys.*, 50:783–859, 1987.

[7] K. Binder. Applications of Monte Carlo methods to statistical physics. *Rep. Prog. Phys.*, 60:487–559, 1997.

[8] K. Binder, M. S. S. Challa, and D. P. Landau. Finite-size effects at temperature-driven first-order phase transitions. *Phys. Rev. B*, 34:1841–1852, 1986.

[9] K. Binder and D. P. Landau. Finite-size scaling at first-order phase transitions. *Phys. Rev. B*, 30:1477–1485, 1984.

[10] H. W. J. Blöte, E. Luijten, and J. R. Heringa. Ising universality in three dimensions: a Monte Carlo study. *J. Phys. A*, 28:6289, 1995.

[11] C. Borgs and S. Kappler. Equal weight versus equal height: a numerical study of an asymmetric first-order transition. *Phys. Lett. A*, 171:37–42, 1992.

[12] C. Borgs and R. Kotecky. Finite-size effects at asymmetric first-order phase transitions. *Phys. Rev. Lett.*, 68:1734–1737, 1992.

[13] A. B. Bortz, M. H. Kalos, and J. L. Lebowitz. New algorithm for Monte Carlo simulation of Ising spin systems. *J. Comput. Phys.*, 17:10–18, 1975.

[14] E. Cacioppo, S. Munson, and M. L. Pusey. Protein solubilities determined by a rapid technique and modification of that technique to a micro-method. *J. Gryst. Growth*, 110:66–71, 1991.

[15] E. Cacioppo and M. L. Pusey. The solubility of the tetragonal form of hen egg white lysozyme from pH 4.0 to 5.4. *J. Gryst. Growth*, 114:286–292, 1991.

[16] H. Cano, N. Gabas, and J. P. Canselier. Experimental study on the ibuprofen crystal growth morphology in solution. *Journal of Crystal Growth*, 224:335–341, 2001.

[17] J. Chang and S. I. Sandler. Determination of liquid-solid transition using histogram reweighting method and expanded ensemble simulations. *J. Chem. Phys.*, 118:8390, 2003.

[18] B. Chen, J. I. Siepmann, and M. L. Klein. Direct Gibbs ensemble Monte Carlo simulations for solid-vapor phase equilibria: Applications to Lennard-Jonesium and carbon dioxide. *J. Phys. Chem. B*, 105:9840–9848, 2001.

[19] T. Chiu and P. D. Christofides. Robust control of particulate processes using uncertain population balances. *AIChE J.*, 46:266–280, 2000.

[20] P. D. Christofides, A. Armaou, Y. Lou, and A. Varshney. *Control and Optimization of Multiscale Process Systems*. Birkhäuser, Boston, 2008.

[21] J. Dai, J. M. Kanter, S. S. Kapur, W. D. Seider, and T. Sinno. On-lattice kinetic Monte Carlo simulations of point defect aggregation in entropically influenced crystalline systems. *Phys. Rev. B*, 72:134102, 2005.

[22] J. Dai, W. D. Seider, and T. Sinno. Coarse-grained lattice kinetic Monte Carlo simulation of systems of strongly interacting particles. *J. Chem. Phys.*, 128:194705, 2008.

[23] P. J. Daudey. Crystallization of ammonium sulfate, Ph. D. diss. *Delft University of Technology, The Netherlands*, 1987.

[24] B. Dünweg and D. P. Landau. Phase diagram and critical behavior of the Si-Ge unmixing transition: A Monte Carlo study of a model with elastic degrees of freedom. *Phys. Rev. B*, 48:14182–14197, 1993.

[25] S. D. Durbin and W. E. Carlson. Lysozyme crystal growth studied by atomic force microscopy. *J. Cryst. Growth*, 122:71–79, 1992.

[26] S. D. Durbin and G. Feher. Crystal growth studies of lysozyme as a model for protein crystallization. *J. Cryst. Growth*, 76:583–592, 1986.

[27] S. D. Durbin and G. Feher. Simulation of lysozyme crystal growth by the Monte Carlo method. *J. Cryst. Growth*, 110:41–51, 1991.

[28] D. M. Elke, J. F. Brennecke, and E. J. Maginn. Toward a robust and general molecular simulation method for computing solid-liquid coexistence. *J. Chem. Phys.*, 122:014115, 2005.

[29] J. R. Errington. Solid-liquid phase coexistence of the Lennard-Jones system through phase-switch Monte Carlo simulation. *J. Chem. Phys.*, 120:3130, 2004.

[30] G. Feher and Z. Kam. Nucleation and Growth of Protein Crystals: General Principles and Assays. *Methods in Enzymology*, 114:77–1112, 1985.

[31] A. M. Ferrenberg and R. H. Swendsen. New Monte Carlo technique for studying phase transitions. *Phys. Rev. Lett.*, 61:2635–2638, 1988.

[32] A. M. Ferrenberg and R. H. Swendsen. Optimized Monte Carlo data analysis. *Phys. Rev. Lett.*, 63:1195–1198, 1989.

[33] K. A. Fichthorn and W. H. Weinberg. Theoretical foundations of dynamical Monte Carlo simulations. *J. Chem. Phys.*, 95:1090–1096, 1991.

[34] E. L. Forsythe, A. Nadarajah, and M. L. Pusey. Growth of (101) faces of tetragonal lysozyme crystals: measured growth-rate trends. *Acta Cryst. D*, 55:1005–1011, 1999.

[35] D. Frenkel. Entropy-driven phase transitions. *Physica A*, 263:26–38, 1999.

[36] D. Frenkel and 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. *J. Chem. Phys.*, 81:3188, 1984.

[37] D. Frenkel and B. Smit. *Understanding Molecular Simulation*. Academic Press, New York, 2002.

[38] O. Galkin and P. G. Vekilov. Direct determination of the nucleation rates of protein crystals. *J. Phys. Chem. B*, 103:10965–10971, 1999.

[39] O. Galkin and P. G. Vekilov. Nulceation of protein crystals: critical nuclei, phase behavior, and control pathways. *Journal of Crystal Growth*, 232:63–76, 2001.

[40] J. Garside. Industrial crystallization from solution. *Chem. Eng. Sci.*, 40:3–26, 1985.

[41] J. Garside and R. I. Ristic. Direct observation of secondary nuceli production. *J. Cryst. Growth*, 43:694–704, 1978.

[42] J. Garside and R. I. Ristic. Growth rate dispersion among ADP crystals formed by primary nucleation. *Journal of Crystal Growth*, 61:215–220, 1983.

[43] D. T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.*, 22:403–434, 1976.

[44] D. T. Gillespie. Exact stochastic simulation of chemical reactions. *J. Phys. Chem.*, 81:2340–2361, 1977.

[45] D. T. Gillespie. Monte Carlo simulation of random walks with residence time dependent transition probability rates. *J. Comput. Phys.*, 28:395–407, 1978.

[46] D. T. Gillespie. A rigorous derivation of the chemical master equation. *Physica A*, 188:404–425, 1992.

[47] D. T. Gillespie. Approximate accelerated stochastic simulation of chemically reacting systems. *J. Chem. Phys.*, 115:1716–1733, 2001.

[48] D. T. Gillespie. Stochastic simulation of chemical kinetics. *Annu. Rev. Phys. Chem.*, 58:35–55, 2007.

[49] G. H. Gilmer and P. Bennema. Simulation of crystal growth with surface diffusion. *J. Appl. Phys.*, 43:1347, 1972.

[50] R. F. P. Grimbergen, E. S. Boek, H. Meekes, and P. Bennema. Explanation for the supersaturation dependence of the morphology of lysozyme crystals. *J. Cryst. Growth*, 207:112–121, 1999.

[51] G. Grochola. Constrained fluid λ-integration: Constructing a reversible thermodynamic path between the solid and liquid state. *J. Chem. Phys.*, 120:2122, 2004.

[52] J.-P. Hansen and L. Verlet. Phase transitions of the Lennard-Jones system. *Phys. Rev.*, 184:151–161, 1969.

[53] M. Hasegawa and K. Ohno. The dependence of the phase diagram on the range of the attractive intermolecular forces. *J. Phys.: Condens. Matter*, 9:3361–3370, 1997.

[54] V. R. Heng, M. Nayhouse, M. Crose, A. Tran, and G. Orkoulas. Direct determination of triple-point coexistence through cell model simulation. *J. Chem. Phys.*, 137:141101, 2012.

[55] W. G. Hoover and F. H. Ree. Use of computer experiments to locate the melting transition and calculate the entropy in the solid phase. *J. Chem. Phys.*, 47:4873, 1967.

[56] W. G. Hoover and F. H. Ree. Melting transition and communal entropy for hard spheres. *J. Chem. Phys.*, 49:3609, 1968.

[57] W. G. Hoover, M. Ross, K. W. Johnson, D. Henderson, J. A. Barker, and B. C. Brown. Soft-sphere equation of state. *J. Chem. Phys.*, 52:4931, 1970.

[58] http://www.statista.com/statistics/296116/top-10-over-the-counter-brands-by-us revenue/.

[59] Hansen J.-P. Phase transition of the Lennard-Jones system. ii. High-temperature limit. *Phys. Rev. A*, 2:221, 1970.

[60] R. A. Judge, E. L. Forsythe, and M. L. Pusey. Growth rate dispersion in protein crystal growth. *Cryst. Growth and Design*, 10:3164–3168, 2010.

[61] S. C. Ke, L. J. DeLucas, and J. G. Harrison. Computer simulation of protein crystal growth using aggregates as the growth unit. *J. Phys. D: Appl. Phys.*, 31:1064–1070, 1998.

[62] A. M. Kierzek and P. Zielenkiewicz. Models of protein crystal growth. *Biophys. Chem.*, 91:1–20, 2001.

[63] Y. C. Kim. Yang-Yang anomalies and coexistence diameters: Simulation of asymmetric fluids. *Phys. Rev, E*, 71:051501, 2005.

[64] Y. C. Kim and M. E. Fisher. Fluid coexistence close to criticality: scaling algorithms for precise simulation. *Comput. Phys. Commun.*, 169:295–300, 2005.

[65] D. A. Kofke. Direct evaluation of phase coexistence by molecular simulation via integration along the saturation line. *J. Chem. Phys.*, 98:4149–4162, 1993.

[66] D. A. Kofke. Gibbs-Duhem integration: a new method for direct evaluation of phase coexistence by molecular simulation. *Mol. Phys.*, 78:1331, 1993.

[67] K. Kurihara, S. Miyashita, G. Sazaki, T. Nakada, Y. Suzuki, and H. Komatsu. Interferometric study on the crystal growth of tetragonal lysozyme crystal. *J. Cryst. Growth*, 166:904–908, 1996.

[68] J. S. Kwon, M. Nayhouse, P. D. Christofides, and G. Orkoulas. Modeling and control of protein crystal shape and size in batch crystallization. *AIChE J.*, 59:2317–2327, 2013.

[69] J. S. Kwon, M. Nayhouse, P. D. Christofides, and G. Orkoulas. Modeling and control of shape distribution of protein crystal aggregates. *Chem. Eng. Sci.*, 104:484–497, 2013.

[70] J. S. Kwon, M. Nayhouse, P. D. Christofides, and G. Orkoulas. Protein cyrstal shape and size control in batch crystallization: Comparing model predictive control with conventional operating policies. *Ind. & Eng. Chem. Res.*, 53:5002–5014, 2014.

[71] A. J. C. Ladd and L. V. Woodcock. Interfacial and co-existence properties of the Lennard-Jones system at the triple point. *Molecular Physics*, 36:611–619, 1978.

[72] M. A. Larson, E. T. White, K. A. Ramanarayanan, and K. A. Berglund. Growth rate dispersion in MSMPR crystallizers. *AIChE J.*, 31:90–94, 1985.

[73] A. J. Liu and M. E. Fisher. The three-dimensional Ising model revisited numerically. *Physica A*, 156:35, 1989.

[74] J. J. Liu, C. Y. Ma, Y. D. Hu, and X. Z. Wang. Modelling protein crystallisation using morphological population balance models. *Chemical Engineering Research and Design*, 88:437–446, 2010.

[75] M. A. Lovette, A. R. Browning, D. W. Griffin, J. P. Sizemore, R. C. Snyder, and M. F. Doherty. Crystal shape engineering. *Ind. Eng. Chem. Res.*, 47:8912–9833, 2008.

[76] A. P. Lybartsev, A. A. Martsinovski, S. V. Shevkunov, and P. N. Vorontsov-Velyaminov. New approach to Monte Carlo calculation of the free energy: Method of Expanded Ensembles. *J. Chem. Phys.*, 96:1776–1783, 1992.

[77] E. Marinari and G. Parisi. Simulated Tempering: a New Monte Carlo Scheme. *Europhys. Lett.*, 19:451–458, 1992.

[78] G. C. McNeil-Watson and N. B. Wilding. Freezing line of the Lennard-Jones fluid: A phase switch Monte Carlo study. *J. Chem. Phys.*, 124:064504, 2006.

[79] C. M*ü*ller and J. Ulrich. A more clear insight of the lysozyme crystal composition. *Cryst. Res. Technol.*, 46:646–650, 2011.

[80] H. Müller-Krumbhaar and K. Binder. Dynamic properties of the Monte Carlo method in statistical mechanics. *J. Stat. Phys.*, 8:1–24, 1973.

[81] A. Nadarajah, E. L. Forsythe, and M. L. Pusey. The averaged face growth rates of lysozyme crystals: the effect of temperature. *J. Cryst. Growth*, 151:163–172, 1995.

[82] A. Nadarajah, M. Li, and M. L. Pusey. Growth mechanism of the (110) face on tetragonal lysozyme crystals. *Acta Cryst.*, D53:524–534, 1997.

[83] C. N. Nanev and D. Tsekova. Heterogeneous nucleation of hen-egg-white lysozyme-molecular approach. *Crystal research and technology*, 35:189–195, 2000.

[84] M. Nayhouse, A. M. Amlani, V. R. Heng, and G. Orkoulas. Simulation of fluid-solid coexistence via thermodynamic integration using a modified cell model. *J. Phys.: Condens. Matter*, 24:155101, 2012.

[85] M. Nayhouse, A. M. Amlani, and G. Orkoulas. A Monte Carlo study of the freezing transition of hard spheres. *J. Phys.: Condens. Matter*, 23:325106, 2011.

[86] M. Nayhouse, A. M. Amlani, and G. Orkoulas. Precise simulation of the freezing transition of supercritical Lennard-Jones. *J. Chem. Phys.*, 135:154103, 2011.

[87] M. Nayhouse, V. R. Heng, A. M. Amlani, and G. Orkoulas. Precise simulation of subcritical freezing using constrained cell models. *Journal of Physics A: Mathematical and Theoretical*, 45:155002, 2012.

[88] M. Nayhouse, V. R. Heng, A. M. Amlani, and G. Orkoulas. Simulation of phase boundaries using constrained cell models. *J. Phys.: Condens. Matter*, 24:375105, 2012.

[89] M. Nayhouse, J. S. Kwon, and G. Orkoulas. Phase transitions, criticality and three-phase coexistence in constrained cell models. *J. Chem. Phys.*, 136:201101, 2012.

[90] T. T. H. Nguyen, R. B. Hammond, I. Marziano, and G. Nichols. Precision measurement of the growth rate and mechanism of ibuprofen 001 and 011 as a function of crystallization environment. *CrystEngComm*, 16:4568–4586, 2014.

[91] J. J. Nicolas, K. E. Gubbins, W. B. Streett, and D. J. Tildesley. Equation of state for the Lennard-Jones fluid. *Mol. Phys.*, 37:1429–1454, 1979.

[92] D. P. Noon and G. Orkoulas. Simulation of phase transitions via spatial updating and tempering. *Molecular Simulation*, 36:535–543, 2010.

[93] M. G. Noro, N. Kern, and D. Frenkel. The role of long-range forces in the phase behavior of colloids and proteins. *Europhys. Lett.*, 48:332–338, 1999.

[94] R. ÓMeadhra and G. M. van Rosmalen. Modelling of the growth of ammonium sulphate crystals in a DTB crystallizer. *Chem. Eng. Sci.*, 51:3919–3929, 1996.

[95] G. Orkoulas. Communication: Tracing phase boundaries via molecular simulation: An alternative to the Gibbs-Duhem integration method. *J. Chem. Phys.*, 133:111104, 2010.

[96] G. Orkoulas. Spatial updating Monte Carlo algorithms in particle simulations. *Molecular Simulation*, 36:516–525, 2010.

[97] G. Orkoulas, M. E. Fisher, and A. Z. Panagiotopoulos. Precise simulation of criticality in asymmetric fluids. *Phys. Rev. E*, 63:051507, 2001.

[98] G. Orkoulas and M. Nayhouse. Communication: A simple method for simulation of freezing transitions. *J. Chem. Phys.*, 134:171104, 2011.

[99] G. Orkoulas, A. Z. Panagiotopoulos, and M. E. Fisher. Precise simulation of criticality in asymmetric fluids. *Phys. Rev. E*, 63:051507, 2001.

[100] A. Z. Panagiotopoulos. Molecular simulation of phase coexistence: Finite-size effects and determination of critical parameters for two- and three-dimensional Lennard-Jones fluids. *Int. J. Thermophys.*, 15:1057–1072, 1994.

[101] A.Z. Panagiotopoulos. Direct determination of phase coexistence properties of fluids by Monte Carlo simulation in a new ensemble. *Mol. Phys.*, 61:813–826, 1987.

[102] J. M. Polson, E. Trizac, and D. Frenkel. Finite-size corrections to the free energies of crystalline solids. *J. Chem. Phys.*, 112:5339, 2000.

[103] M. L. Pusey and A. Nadarajah. A model for tetragonal lysozyme crystal nucleation and growth. *Crystal Growth & Design*, 2:475–483, 2002.

[104] A. D. Randolph and E. T. White. Modeling size dispersion in the prediction of crystal-size distribution. *Chem. Eng. Sci.*, 32:1067–1076, 1977.

[105] N. Rasenack and B. W. Müller. Ibuprofen crystals with optimized properties. *International Journal of Pharmaceutics*, 244:45–57, 2002.

[106] A. Rashid, E. T. White, T. Howes, J. D. Litster, and I. Marziano. Racemic ibuprofen solubility in ethanol and aqueous ethanolic mixtures. In *Proceedings of Chemica 2008*, pages 1393–1401, Newcastle, Australia, 2008.

[107] A. Rashid, E. T. White, T. Howes, J. D. Litster, and I. Marziano. Crystallization kinetics of ibuprofen from ethanol and aqueous ethanol. *Chemical Engineering Transactions*, 24:631–636, 2011.

[108] A. Rashid, E. T. White, T. Howes, J. D. Litster, and I. Marziano. Growth rates of ibuprofen crystals grown from ethanol and aqueous ethanol. *Chemical Engineering Research and Design*, 90:158–161, 2012.

[109] M. Rathinam, L. R. Petzold, Y. Cao, and D. T. Gillespie. Stiffness in stochastic chemically reacting systems: The implicit tau-leaping method. *J. Chem. Phys.*, 119:12784–12794, 2003.

[110] J. S. Reese, S. Raimondeau, and D. G. Vlachos. Monte Carlo Algorithms for Complex Surface Reaction Mechanisms: Efficiency and Accuracy. *J. Comput. Phys.*, 173:302–321, 2001.

[111] A. L. Rohl. Computer prediction of crystal morphology. *Curr. Opin. Solid State Mater. Sci.*, 7:21–26, 2003.

[112] M. Rovere, D. W. Heermann, and K. Binder. The gas-liquid transition of the two-dimensional Lennard-Jones fluid. *J. Phys.: Condens. Matter*, 2:7009, 1990.

[113] I. T. Rusli, M. A. Larson, and J. Garside. Initial growth of secondary nuclei produced by contact nucleation. *AIChE Symposium Series*, 76:52–58, 1980.

[114] N. Shankland, A. J. Florence, P. J. Cox, D. B. Sheen, S. W. Love., N. S. Stewart, and C. C. Wilson. Crystal morphology of ibuprofen predicted from single-crystal pulsed neutron diffraction data. *Chem. Commun.*, pages 855–856, 1996.

[115] N. Shankland, C. Wilson, A. J. Florence, and P. Cox. Refinement of ibuprofen at 100K by single-crystal pulsed neutron diffraction. *Acta Crystallogr.*, 53:951–954, 1997.

[116] D. Shi, N. H. El-Farra, M. Li, P. Mhaskar, and P. D. Christofides. Predictive control of particle size distribution in particulate processes. *Chem. Eng. Sci.*, 61:268–281, 2006.

[117] D. Shi, P. Mhaskar, N. H. El-Farra, and P. D. Christofides. Predictive control of crystal size distribution in protein crystallization. *Nanotechnology*, 16:S562–S574, 2005.

[118] B. Smit. Phase diagrams of Lennard-Jones fluids. *J. Chem. Phys.*, 96:8639, 1992.

[119] M. A. Snyder, A. Chatterjee, and D. G. Vlachos. Net-event kinetic Monte Carlo for overcoming stiffness in spatially homogeneous and distributed systems. *Comput. Chem. Eng.*, 29:701–712, 2005.

[120] W. B. Streett, H. J. Raveché, and R. D. Mountain. Monte Carlo studies of the fluid-solid phase transition in the Lennard-Jones system. *J. Chem. Phys.*, 61:1960, 1974.

[121] Y. Suzuki, S. Miyashita, H. Komatsu, K. Sato, and T. Yagi. Crystal growth of hen egg white lysozyme under high pressure. *Japanese Journal of Applied Physics*, 33:1568–1570, 1994.

[122] P. G. Vekilov. Elementary processes of proetin crystal growth. *Prog. Crystal Growth and Charact.*, 26:25–49, 1993.

[123] P. G. Vekilov, M. Ataka, and T. Katsura. Laser Michelson interferometry investigation of protein crystal growth. *J. Cryst. Growth*, 130:317–320, 1993.

[124] P. G. Vekilov and A. A. Chernov. The physics of protein crystallization. *Solid State Physics*, 57:1–147, 2003.

[125] G. A. Vliegenthart, J. F. M. Lodge, and H. N. W. Lekkerkerker. Strong weak and metastable liquids structural and dynamical aspects of the liquid state. *Physica A*, 263:378–388, 1999.

[126] N. B. Wilding. Critical-point and coexistence-curve properties of the Lennard-Jones fluid: A finite-size scaling study. *Phys. Rev. E*, 52:602, 1995.

[127] N. B. Wilding and A. D. Bruce. Freezing by Monte Carlo phase switch. *Phys. Rev. Lett.*, 85:5138–5141, 2000.

[128] L. V. Woodcock. Computation of the free energy for alternative crystal structures of hard spheres. *Faraday Discuss.*, 106:325–338, 1997.

[129] U. Zacher and A. Mersmann. The influence of internal crystal perfection on growth rate dispersion in a continuous suspension crystallizer. *Journal of Crystal Growth*, 147:172–180, 1995.

[130] Y. Zhang and E. J. Maginn. A comparison of methods for melting point calculation using molecular dynamics simulations. *J. Chem. Phys.*, 136:144116, 2012.