Liquid-Solid Transitions with Applications to Self-Assembly

by

Aaron S. Keys

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemical Engineering)
in The University of Michigan
2010

Doctoral Committee:
Professor Sharon C. Glotzer, Chair
Professor John Kieffer
Professor Michael J. Solomon
Professor Robert M. Ziff

The most incomprehensible thing about the universe
is that it is comprehensible.

— Albert Einstein

For my parents, Ellen and Lee

Acknowledgments

I thank my advisor Sharon Glotzer for allowing me to pursue problems that I am passionate about. I thank her for looking out for my interests, even, at times, in deference to her own, and for always putting me in a position to succeed. I thank her for fostering an environment of hard work and excellence in the Glotzer group, and for her enthusiastic dedication to research that is contagious with all of her students.

I thank the other faculty members that I have had the pleasure of interacting with. I thank Dave Kofke, Daan Frenkel, Mike Solomon, Bob Ziff and John Kieffer for their interesting conversations and mentoring, and for constructive feedback regarding my work. I thank Doug Durian for allowing me to collaborate on a fascinating project. I thank David Chandler, whose remarkable creativity and intense dedication to research has been a source of inspiration to me, as well as so many other young scientists.

I thank my mentors Mark Horsch, Magnus Bergroth, Yeshi Gebremichael, Michael Vogel, and Naida Lacevic for helping me to lay the early foundations for what eventually became this thesis. I also thank my contemporaries Chris Iacovella, Josh Andersen, Michael Engel, Amir Haji-Akbari, Lester Hedges, and Trung Dac Nguyen, for all that they have taught me throughout our time working together. I thank Chris Iacovella, Eric Jankowski, Stephanie Teich-McGoldrick, *et al.* for making work fun all of those years. I thank all of the Glotzer group members, past and present, for their friendships – many of which will last a lifetime.

I thank my family for encouraging me to follow my dream of becoming a scientist. I thank my younger siblings Bob and Jessy for their admiration ;). Truth be told, the feeling is mutual. I thank my grandmother Elizabeth for her love and support, and for filling my childhood with happiness during all of the time we spent together. I thank my grandmother Ethel for her love and inspiration– knowing that the daughter of immigrants with almost no formal schooling could become a Harvard-educated professor will always remind me of the great things that people can accomplish through education. Although he is no longer with us, I thank my grandfather Saul for teaching me to stay positive and appreciate each day.

Most of all, I thank my mother Ellen and my father Lee. I thank my dad for teaching me

the value of hard work and family, values which have served me well throughout my time at UofM. I thank my mom for always emphasizing the importance of education, all the little things that she has done (and continues to do) for me to make my life easier and happier. I thank both of my parents for their unwavering love and support. It is easy to reach for your goals when you know you will always have someone in your corner.

Table of Contents

List of Tables

Table

List of Figures

Figure

List of Appendices

Appendix

List of Abbreviations

API Application Programmer Interface

BCC Body Centered Cubic

BLJ Binary Lennard-Jones

BOD Bond Order Diagram

BOP Bond Order Parameter

BUBBA Bottom Up Building Block Assembly

CNA Common Neighbor Analysis

CNT Classical Nucleation Theory

CPU Central Processing Unit

CRR Cooperatively Rearranging Region

DFT Discrete Fourier Transform

DH Dynamical Heterogeneity

DSHT Discrete Spherical Harmonics Transform

DZ Dzugutov

DQC Dodecagonal Quasicrystal

FCC Face Centered Cubic

FEP Free Energy Perturbation

FK Frank-Kasper

FIRE Fast Inertial Relaxation Engine

FENE Finitely Extensible Non-linear Elastic

FLD Frustration Limited Domain

GPU Graphics Processing Unit

HCP Hexagonal Close Packed

HS Hard Sphere

ICP Iterative Closest Point

KA Kob-Andersen

KCM Kinetically Constrained Model

LJ Lennard-Jones

LJG Lennard-Jones Gauss

MD Molecular Dynamics

MC Monte-Carlo

MCT Mode Coupling Theory

MMC Metroplis Monte-Carlo

MSD Mean Squared Displacement

MSM Minimal Spherical Micelle

OO Orientational Ordering

OTF On the Fly

PCA Principle Components Analysis

PD Pentagonal Dipyramid

PEL Potential Energy Landscape

PPTIS Partial Path Transition Interface Sampling

QC Quasicrystal

RDF Radial Distribution Function

RFOT Random First Order Theory

RMS Root Mean Square

RNS Random Number Sequence

SA Self-Assembly

SHD Spatially Heterogeneous Dynamics

TEM Transmission Electron Microscopy

TI Thermodynamic Integration

TIS Transition Interface Sampling

TNS Tethered Nano Sphere

TPE Transition Path Ensemble

TPS Transition Path Sampling

TSE Transition State Ensemble

VTF Vogel-Tammann-Fulcher

WCA Weeks-Chandler-Andersen

Abstract

We study the thermodynamic and kinetic pathways by which liquids transform into solids, and their relation to the metastable states that commonly arise in self-assembly applications. As a case study in the formation of ordered metastable solids, we investigate the atomistic mechanism by which quasicrystals form. We show that the aperiodic growth of quasicrystals is controlled by the ability of the growing quasicrystal "nucleus" to incorporate kinetically trapped atoms into the solid phase with minimal rearrangement. In a related study, we propose a two-part mechanism for forming 3d dodecagonal quasicrystals by self-assembly. Our mechanism involves (1) attaching small mobile particles to the surface of spherical particles to encourage icosahedral packing and (2) allowing a subset of particles to deviate from the ideal spherical shape, to discourage close-packing. In addition to studying metastable ordered solids, we investigate the phenomenology and mechanism of the glass transition. We report measurements of spatially heterogeneous dynamics in a system of air-driven granular beads approaching a jamming transition, and show that the dynamics in our granular system are quantitatively indistinguishable from those for a supercooled liquid approaching a glass transition. In a second study of the glass transition, we use transition path sampling to study the structure, statistics and dynamics of localized excitations for several model glass formers. We show that the excitations are sparse and localized, and their size is temperature-independent. We show that their equilibrium concentration is proportional to $\exp[-J_a(1/T - 1/T_\mathrm{o})]$, where J_a is the energy scale for irreversible particle displacements of length a, and T_o is an onset temperature. We show that excitation dynamics is facilitated by the presence of other excitations, causing dynamics to slow in a hierarchical way as temperature is lowered. To supplement our studies of liquid-solid transitions, we introduce a shape matching framework for characterizing structural transitions in systems with complex particle shapes or morphologies. We provide an overview of shape matching methods, explore a particular class of metrics known as "harmonic descriptors," and show that shape matching methods can be applied to a wide range of nanoscale and microscale assembly applications.

Chapter 1
Introduction

Throughout human history, the advent of new materials and manufacturing techniques has driven technology and innovation. It is no coincidence that different epochs of human civilization are often classified by the materials used to construct tools and weaponry during those time periods, such as the Stone Age, Bronze Age, and Iron Age. Historically, inventing new materials and manufacturing techniques was as much an art as it was a science; innovations occurred either by accident or through trial and error, and trades were passed down through generations. For example, Iron-Age blacksmiths learned that by rapidly quenching the hot surface of steel blades they could greatly increase their hardness, which could then be adjusted by tempering [1]. Glassworkers learned that by mixing sand with "natron," they could obtain glasses that were rigid even when partially molten, which gave rise to a new technique known as "glass blowing" for molding glass into unique shapes [2]. In recent centuries, through the scientific and industrial revolutions these types of trades have evolved into sciences; by harnessing a fundamental understanding of physics and materials, scientific breakthroughs have engendered a huge influx of new materials and devices, as well as new technologies for manufacturing them.

The newest frontiers in materials and manufacturing technology, known as "microtechnology" and "nanotechnology," involve the construction of tiny devices on the microscale or nanoscale ($\sim 10^{-6}$m and $\sim 10^{-9}$m, respectively). To put this into perspective, the atomic radius of a single atom is on the order of 10^{-10}m. Micro and nanoscale devices have the potential to impact a variety of disciplines. They might, serve as medical devices for drug delivery or administering chemotherapy [3, 4], tiny transistors that can be used to construct computers [5,6], or photonic circuits that control the flow of light rather than electricity [7,8]. Creating such tiny devices poses a unique problem. In contrast to macroscopic devices that can be assembled by traditional top-down approaches, such tiny devices are impractical to manipulate. One potentially powerful solution to this problem is to use the method of "self-assembly" to manufacture materials and devices from the bottom-up.

Self-assembly involves devising systems with particular characteristics such that they

will naturally evolve towards desired target structures [9]. One way to accomplish this is to design specialized particle "building-blocks" that tend to assemble into particular favorable structures due to thermodynamic driving forces [10, 11]. Advances in synthetic chemistry have made it possible to create micro or nanoscale particles with a wide variety of shapes, functional groups, and interparticle interactions that can be used to such ends [11, 12, 13]. Understanding the rules by which these particles assemble requires a great deal of scientific intuition; now more than ever, we must rely on scientific knowledge rather than trial and error to invent new materials and manufacture new devices.

Predicting which structures will spontaneously assemble from a given system of nanoparticles requires a fundamental understanding of thermodynamics [14], which drives the assembly process. One of the most fundamental principles of thermodynamics is that systems evolve toward equilibrium; that is, any finite-temperature system, given sufficient time, will evolve to a state that minimizes free energy. Perhaps counterintuitively, however, direct knowledge of the minimum free energy state is insufficient for predicting the structure a given system will assemble. For applications involving condensed matter (i.e., liquids and solids), one must account for the possibility that systems can fall into "metastable states," physically-stable intermediates that form on the way to the stable state. Because liquids and solids evolve slowly, condensed systems often do not reach the minimum free energy state on a timescale that is physically relevant to the intended application. For example, the tempered steel mentioned above contains a particularly hard, metastable crystal structure known as "martensite" that forms when the steel is rapidly quenched. Over sufficiently long timescales, the atoms in the crystal rearrange into the equilibrium structure, and the hardness of the steel is lost; however, this timescale often greatly exceeds the intended lifespan of, say, Iron-Age swords or contemporary cutlery or tools for which the material is intended. Glasses, also mentioned above, are also metastable materials. On timescales on the order of 10^{32} years (much longer than the age of the universe, $\sim 10^{10}$ years), ordinary silica-based window glass will flow like liquid [15]. This, however, has little bearing on its usefulness in our everyday lives.

Because systems might form metastable states, when designing systems for self-assembly, we must consider the full thermodynamic pathway by which systems transform into their assembled states. In this work, we focus on one common pathway by which structures can self-assemble, which involves starting in a disordered liquid state and transforming into a solid state, which may be either ordered or amorphous. Typically, this type of solidification pathway might involve cooling or compressing a system of particles, or growing the particles from solution. There are two common mechanisms by which solid phases form: nucleation and growth, and vitrification. The former leads to an ordered solid (crystal-

lization) whereas the latter leads to an amorphous solid (glass formation). Although these transitions have been studied extensively in the past, many important questions regarding both remain unanswered. In the case of nucleation and growth, the fundamental thermodynamic mechanism underlying the transition has been known for some time [16, 17, 18]; however the detailed microscopic picture of how nucleation and growth occurs is elusive. This is exemplified by recent simulation studies [19, 20, 21, 22] that demonstrate that the standard classical nucleation theory (CNT [17, 18]) is insufficient for describing nucleation on a detailed particle-based level. In the case of vitrification, the fundamental nature of the transition itself is unknown. Thus, gaining insight into this mechanism is of utmost importance, and remains the cornerstone of supercooled liquids research [23, 24, 25, 26]. Here, we investigate both transitions to elucidate an microscopic picture of phase transitions that will guide researchers in solving complex problems in self-assembly.

This thesis is organized as follows: In chapter 2, we provide a theoretical overview of both the thermodynamics and dynamics of liquid-solid transitions. In chapters 3 and 4 we review the model systems and computational techniques that we use to study these transitions. In chapter 5, we study the formation of metastable ordered solids in the context of dodecagonal quasicrystals. Quasicrystals have useful photonic properties on the micro/nanoscale [27,28], and serve as a useful case study for our investigations. In chapter 6, we investigate the mechanism underlying the glass transition. To do so, we characterize the dynamics of several model glass formers, as well as an experimental granular system of macroscopic ball bearings that exhibits many of the same dynamical characteristics. In chapter 7, we introduce a new class of computer algorithms for characterizing assembled structures, which can be applied to studying thermodynamic transitions in assembled systems. We conclude by providing possible avenues for future study in chapter 8, and some closing remarks in chapter 9. While the work presented here is largely focused on the fundamental aspects of liquid-solid transitions, we have applied our knowledge of these topics to several novel self-assembly applications, which we briefly review in Appendix A. The simulation and analysis codes that we have developed to study liquid-solid transitions are reviewed in Appendix B.

The research presented in this dissertation has been published or is currently being prepared for publication. Published articles include [Z. Zhang, A.S. Keys, T. Chen, and S.C. Glotzer, "Self-Assembly of Patchy Particles into Diamond Structures Through Molecular Mimicry," *Langmuir* (2005)] [29], [A.S. Keys and S.C. Glotzer, "How Do Quasicrystals Grow?" *Physical Review Letters* (2007)] [30], [A.S. Keys, A.R. Abate, S.C. Glotzer, and D.J. Durian, "Measurement of Growing Dynamical Length Scales and Prediction of the Jamming Transition in a Granular Material," *Nature Physics* (2007)] [31], [C.R. Iacovella, A.S. Keys,

M.A. Horsch, and S.C. Glotzer, "Icosahedral Packing of Polymer-Tethered Nanospheres and Stabilization of the Gyroid Phase," *Physical Review E* (2007)], [S.C. Glotzer and A.S. Keys, "Materials Science: A Tale of Two Tilings," *Nature* (2008)] [32], [A. Haji-Akbari *et al.*, "Disordered, Quasicrystalline and Crystalline Phases of Densely Packed Tetrahedra," *Nature* (2009)] [33], and [A.S. Keys, C.R. Iacovella, and S.C. Glotzer, "Characterizing Structure Through Shape Matching and Applications to Self-Assembly," *Annual Reviews of Condensed Matter Physics* (2011)] [34]. Articles currently awaiting publication include [C.R. Iacovella, A.S. Keys, and S.C. Glotzer, "Self-Assembly of Soft Matter Quasicrystals," (2010)] [35], [A.S. Keys, L.O. Hedges, J.P. Garrahan, S.C. Glotzer, and D. Chandler, "Structure of Localized Excitations and Relaxation in Supercooled Glass-Forming Liquids," (2011)] [36], [A.S. Keys, C.R. Iacovella, and S.C. Glotzer, "Characterizing Complex Particle Morphologies Through Shape Matching: Descriptors, Applications, and Algorithms," (2011)] [37], and [A.S. Keys, C.R. Iacovella, and S.C. Glotzer, "Harmonic Order Parameters for Characterizing Complex Particle Morphologies," (2011)] [38].

Chapter 2

Background

Our work on liquid-solid transitions builds on a wealth of research in the areas of both nucleation and growth and the glass transition. In this chapter, we provide a theoretical background, as well as a current research overview, of both transitions. We begin by discussing the thermodynamics of metastable states and how this relates to both nucleation and growth and the glass transition. We then describe how these thermodynamics relate to local particle packing in assembled systems, which we exemplify for the simple case of spherical particles. We close with a relevant literature review that we will build on for our studies of liquid-solid transitions and self-assembly.

2.1 Thermodynamics, Kinetics and Metastability

In the limit of infinite system sizes and timescales, thermodynamic systems assume a state of minimum free energy [14]. Physically, "free energy" refers to the capacity for a system to do work. The free energy is given by $F = U - TS$ in the canonical (NVT) ensemble (Helmholtz free energy), or $G = U - TS + PV$ in the isothermal-isobaric (NPT) ensemble (Gibbs free energy). Intuitively, this means that systems with higher energy U or enthalpy $U + PV$ have the ability to do more work. It also means that systems with lower entropy S can do more work than those with higher entropy. (Or, to put it another way, work must be put into the system to remove entropy).

The minimum free energy state is path-independent; therefore, predicting the structure that will form in the thermodynamic limit reduces to the problem of finding the minimum free energy state. Although this is a valid approach in some cases to predicting the outcome of self-assembly experiments, the reality is, experimental systems are finite in space and time; thus the thermodynamic limit is never strictly achieved. This is particularly important when the system is driven from equilibrium, as is typical for systems that undergo liquid-solid transitions, where systems are often cooled or compressed, or the chemical

composition is changed, etc., causing the equilibrium state to change from a disordered liquid state to a solid. If a thermodynamic or kinetic barrier to equilibrium exists, the system may never reach the stable state on timescales that concern us. Such long-lived non-equilibrium states are termed "metastable."

In our studies of liquid-solid transitions, we encounter many metastable states. One frequently-encountered metastable state is the supercooled liquid state, which occurs when a liquid is rapidly cooled below its melting temperature T_m or compressed above its melting density ϕ_m. Below T_m or above ϕ_m, the liquid is metastable with respect to the ordered solid [25, 24]. When subjected to increased supercooling, the liquid must eventually solidify, either into an ordered solid or into an amorphous state known as a "glass." As explained in section 2.4, the glassy state is a liquid that has lost ergodicity (i.e., the ability for the system to sample configuration space). Thus, glasses are inherently non-equilibrium structures that are thermodynamically metastable with respect to other glasses that result from longer sampling. Often, glasses are associated with an an underlying crystalline stable state; for example most of the water in the universe exists as glassy water rather than ice [39]. As detailed in section 2.3, ordered solids may also exhibit metastability when particular thermodynamic pathways cause systems to become trapped in crystal structures that are not the stable state. For example, diamonds, which form at high pressures in the Earth's crust, are thermodynamically metastable at atmospheric conditions relative to graphite [40].

2.1.1 Ising Model Analogy: Definition and Behavior

The basic thermodynamics and kinetics underlying liquid-solid transitions can by idealized by a straightforward model known as the "Ising Model [41, 14]," which (roughly) represents a simple model magnet. The 2d square lattice Ising model consists of cells occupied by up or down spins. The energy function for the Ising model is given by [14]:

$$E = -\mu B \sum_i s_i - \sum_{i \neq j} J_{ij} s_i s_j. \tag{2.1}$$

Here, μ is the dipole moment (either -1 or 1, but taken here to be 1), B is the field strength, S_i is the value of the spin (either -1 or 1) at each lattice site i, and J is a coupling parameter between spins. If J is positive (i.e, the system is ferromagnetic), the pairwise energy is lower if adjacent cells contain the same spins. In the absence of an external field both phases (up or down spins) are equally likely. However, if an external field is applied one phase is favored over the other. From the laws of statistical mechanics, the probability of observing a given microstate ν with energy E_ν in equilibrium is $P_\nu \propto \exp(-\beta E_\nu)$, where β is the inverse

temperature. This distribution can be sampled using a Metropolis Monte Carlo simulation (see section 4.1.2).

This simple model can be used to model the basic thermodynamics and kinetics of liquids-solid transitions for systems with local interactions. In the context of liquid-solid transitions in systems of particles, the spins (-1 and 1) are analogous to the two phases: liquid (L) and solid (S). The external field B is the thermodynamic driving force, which, in physical systems, is set by adjusting the temperature, pressure, etc. such that a given phase, either L or S, has a lower free energy. The coupling parameter J is an interfacial tension between the phases. As explained in detail in section 2.2, due to a packing mismatch between particles in the liquid and solid phases, the interfacial tension is positive. Thus, to model such systems we take $J > 0$.

Consider now a system with negative value of the external field B that favors down spins (i.e., the liquid phase L). Fig. 2.1a depicts a schematic of the free energy $F(M)$ (i.e., the reversible work of magnetization) for this system as a function of the total magnetization $M = \sum s_i$. The system tends to sample configurations in the thermodynamically stable L state. Now consider what happens when we quickly change B to a positive value such that the system favors up spins (the solid phase S). In order to change from the current L state to the more stable S state, the system must form many adjacent up/down spin pairs, which increases the value of E. The ability for the system to overcome the barrier depends on the value of the interfacial tension J between the phases; for low values of J, the transition state M^* forms readily due to minor fluctuations, and the barrier E^* is easily surmounted (Fig. 2.1b). In the context of liquid-solid transitions, this scenario results in a process is known as "spinodal decomposition [42]," which we do not consider in much detail here. For larger values of J, there is an energy barrier to transforming the system to the stable S state, and the system remains metastable in L until this barrier is overcome (Fig. 2.1b). The mechanism by which the system surmounts the barrier is known as "nucleation and growth." For very high values of J, forming the transition state requires a particularly improbably configuration of spins and the system will remain metastable in L indefinitely (Fig. 2.1d).

2.1.2 Ising Model Analogy: Nucleation and Growth

Now, let us consider how the transition state M^* in Fig. 2.1c comes about in more detail. The process involves two competing effects: on one hand, there is the external field B, which favors up spins (S); on the other there is the interfacial tension, set by the coupling parameter J, which disfavors interactions between unlike spins (L-S neighbors). To minimize the energy barrier E^*, the system should maximize the number of up spins, while at the same

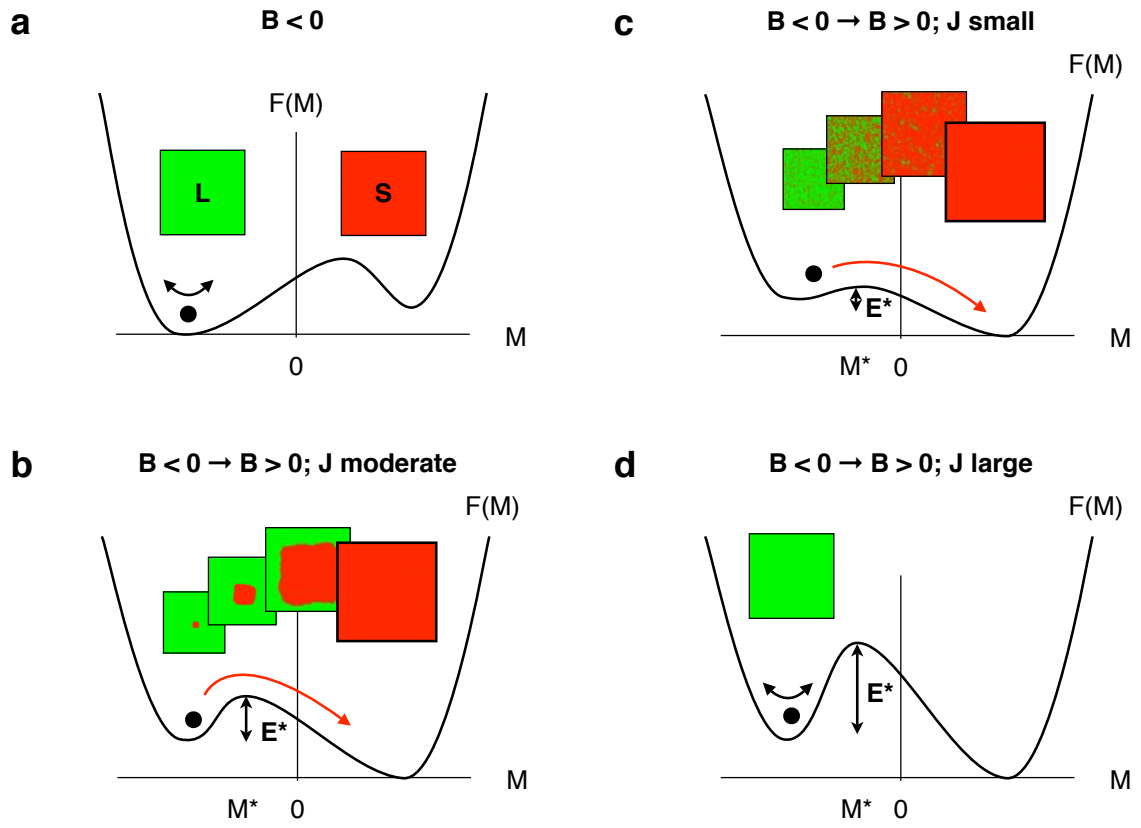


Figure 2.1 Thermodynamics and kinetics of transitions in the Ising Model. All free energy curves are very rough schematics and are not meant to be quantitative. (a) Free energy for a system as a function of the overall magnetization M, with field $B < 0$, such that the L phase, depicted in green, is favored. The system configuration, depicted by a black sphere, tends to sample the stable state, L. (b) Depiction of the transition that occurs when the field B is changed from negative to positive, such that the S phase, depicted in red, is favored. For low values of the coupling energy J, which is related to the interfacial energy, the transition resembles spinodal decomposition. (c) Depiction of the transition that occurs when the field B is changed from negative to positive, with a moderate value of J. For this moderate value of the coupling parameter, the transition occurs via a nucleation and growth mechanism. (d) Depiction of a system for which the field B is changed from negative to positive, with a very high value of J. In this case, the high interfacial tension poses an insurmountable energy barrier for the transition, and the system remains metastable in L indefinitely.

time minimizing the up/down spin neighbors. This is best accomplished by configurations that, through random fluctuations, containing a single disk-like droplet of S spins in an otherwise L system. In this case, the perimeter-to-area ratio of the droplet, and hence the interfacial tension, is minimized. As the droplet grows, its internal energy, which in 2d scales with the droplet area πr^2, grows faster than the interfacial tension, which scales with the droplet perimeter $2\pi r$. At a given critical size r^*, the two terms balance and the droplet has an equal probability to either grow or shrink. This is the transition state M^*, and the droplet is known as the "critical nucleus." Droplets smaller than the critical nucleus tend to shrink, while larger droplets tend to grow. Eventually, the droplet will grow large enough to convert the entire system to the stable S phase. This nucleation and growth mechanism is the most common pathway by which liquids transform into ordered solids. We note that, in our simplified discussion here, we assume that nucleus size r is the only important reaction coordinate for the system; however, even for idealized systems such as the Ising model, this is not strictly true [20, 43]. In section 2.3, we discuss nucleation and growth in the context of less idealized systems with explicit particles.

2.1.3 Ising Model Analogy: Ergodicity Breaking

A metastable system in L subjected to a field $B > 0$ will remain metastable until a critical nucleus forms. Let us now consider a scenario in which a system in L with a high barrier E^* (for example, the system depicted in Fig 2.1d) is subjected to a rapid quench by quickly increasing β. In this case, the critical nucleus does not have time to form, and because the temperature is low the system no longer undergoes thermal fluctuations, and becomes effectively frozen in a local energy minimizing spin configuration. In other words, the system becomes non-ergodic, as it no longer samples configuration space. The configuration of spins in which the system becomes trapped inevitably retains character of the higher temperature equilibrium L state from which the system was quenched; thus, the non-ergodic system is thermodynamically metastable with respect to both the equilibrium low-temperature L phase, which would occur over infinite timescales, as well as the ground state S phase. In particle systems, this non-ergodic phase is known as the glass, and the process by which dynamical arrest occurs is known as the glass transition. In this respect, the distinction between a metastable liquid and a glass is simply a matter of perception; the glass transition occurs when the relaxation time of the system exceeds experimental timescales. However, in practice, the process by which dynamical arrest occurs in fully atomistic systems is much more complicated than our simple discussion would imply, and is reviewed in detail in section 2.4.

2.2 Particle Packing and Thermodynamics

In the section 2.1, we discussed the thermodynamics of liquid-solid phase transitions in the context of the coarse-grained Ising model. In this section, we provide a more detailed, particle-based overview of the same topics. We begin by asserting that the free energy of condensed systems is dominated by the packing of their constituent particles [44, 45, 46, 47]. Systems with interparticle potential interactions tend to assumes packings with interatomic spacings that minimize the potential energy. Systems with excluded volume interactions tend to assume dense arrangements that maximize the accessible space for particles to jostle, thus maximizing the entropy, as well as the PV contribution to the Gibbs free energy in the NPT ensemble. These basic principles apply to condensed systems in general, regardless of the complexity of the particle shape or the interparticle interactions.

2.2.1 Case Study: Disks and Spheres

We can obtain fundamental insight into relationship between packing and thermodynamics by studying two simple models: systems of circular and spherical particles in 2d and 3d, respectively. Despite their apparent similarity, as we will describe below, these two systems represent opposite ends of the spectrum regarding how their constituent particles pack. The intuitive picture that we gain from these simple models will aide us in understanding systems involving more complex particle shapes and interparticle interactions.

Spheres and Disks: Solid Structures

We first consider the trivial case of packing disks in flat, 2d space. Assume, for the time being, that the particles are hard (i.e. they do not overlap) and monodisperse (i.e., they have the same diameter σ). Also, assume that the system is dense such that particle diffusion cannot occur. At such high system densities, the disks tend to assume dense local packings, since this maximizes the volume available for vibrational motion. This vibrational motion dominates the entropy S in the absence of particle diffusion. The densest packing locally is the hexagon [48, 47], where each disk is surrounded by 6 equidistant neighbors (see Figure 2.2). Hexagonal packing is also the densest structure globally, with a packing fraction $\phi_{max} \approx 0.91$. By connecting the centers of each disk in the hexagonal lattice, we divide space evenly into equilateral triangles. It is perhaps counterintuitive that this crystal structure maximizes entropy, since entropy is often associated with disorder. However, the formal definition of entropy is $S = k_b T \ln \Omega$, where Ω is the total number of states accessible to the

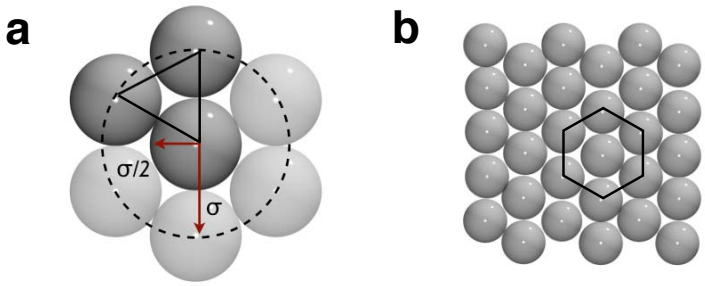


Figure 2.2 Packing disks in 2d. (a) The 7 particle hexagon, which divides space into equilateral triangles. The hexagon maximizes the packing density locally. (b) The hexagonal lattice, composed of particles in hexagons locally. The hexagonal crystal maximizes packing density globally.

system [14]. Thus, since the crystal allows for the most vibrational states, it minimizes the free energy through entropic stabilization. For systems with interparticle interactions, this entropic effect is sometimes augmented by a potential energy effect. For example, since all interparticle distances in a hexagonal crystal are equivalent, particles can lie at the minimum of spherically symmetric attractive pair potentials, such as the Lennard-Jones potential (see section 3.1), thus minimizing the potential energy U as well. In other cases, potentials can be contrived to override the entropic effect, a topic that we explore in more detail in Chapter 3.

Now consider the same problem as above, except in three dimensions. Whereas packing disks in 2d is relatively straightforward, packing spheres in 3d is much more complex [44,47]. The closest analogue to the 2d hexagon in 3d is the icosahedron, which consists of a central sphere is surrounded by 12 roughly equidistant neighbors (see Fig. 2.3a). Like the hexagon, the icosahedron is a dense local packing that allows particles to lie near the minimum of the potential well of spherically symmetric attractive potentials [49]. Just as hexagons divide space into equilateral triangles, icosahedra divide space into tetrahedra, the 3d analogue of the equilateral triangle. However, whereas the equilateral triangles formed by hexagons are ideal, the tetrahedra formed by icosahedra are slightly distorted. As depicted in Fig. 2.3b, the tetrahedra must stretch along their outer edge to fill the 7° gap left between undistorted tetrahedra. As a result, icosahedra do not tile 3d space, resulting in packing frustration for long-range icosahedral structures.

Due to the inherent packing defects associated with long-ranged icosahedral structures, crystalline structures in 3d take on an altogether different packing paradigm. The densest global packing for 3d spheres is achieved by stacking 2d hexagonal layers. The layers may be stacked in an infinite number of ways to give rise to “close-packed” arrangements, all

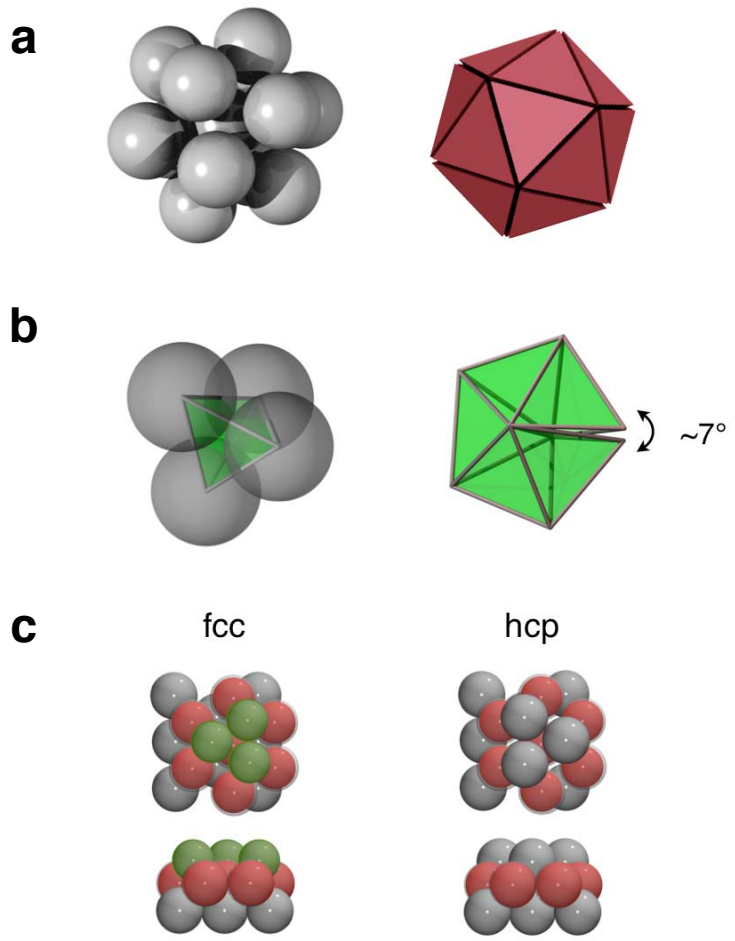


Figure 2.3 Packing spheres in 3d. (a) The image depicts the 13-particle icosahedron, which divides space into tetrahedra. Like its 2d analogue the hexagon, the icosahedron maximizes the packing density locally. (b) Packing tetrahedra. Whereas equilateral triangles tile 2d space, tetrahedra must distort to fill 3d space, as exemplified by the $\sim 7^\circ$ gap that appears when attempting to pack five ideal tetrahedra into a pentagonal dipyramid. (c) Close-packed crystals. Due to the inherent packing frustration for icosahedral structures, the densest packings in 3d are layered hexagonal structures, such as the fcc and hcp crystals. The structures differ in the stacking sequence; whereas the fcc structure has an ABC layering pattern, the hcp structure has an alternating AB pattern.

of which have a packing fraction of $\phi = \frac{pi}{\sqrt{18}} \approx 0.74048$ [50]. Two simple close-packed arrangements are the fcc structure (also known as cubic-close packed) and the hcp structure, which involve AB stacking, and ABC stacking, respectively. Although these structures do not divide space entirely into tetrahedra, they minimize free energy for many systems because of their high density. Another common crystal is the bcc structure with a packing fraction of $\phi \approx 0.681$. Although the packing fraction is lower for bcc than for fcc/hcp, the bcc structure is sometimes more energetically favorable for some specific interparticle interaction potentials [51, 52].

Spheres and Disks: Liquid Structures

In general, liquid structure can be understood in terms of random deviations from the ideal local packing [47]. Due to thermal fluctuations, liquid systems of 2d disks will contain some particles with more than six neighbors, and some with fewer, in addition to ideal hexagons. Local configurations with packing defects tend to be less stable than defect-free hexagonal configurations. Thus, as the temperature is lowered or the density is increased, the number of defects decreases as well.

Liquid structure in 3d systems of spheres can be understood in terms of local deviations from the ideal icosahedral packing [47]. Icosahedra can be decomposed entirely into pentagonal dipyramids (Fig. 2.3b), where five tetrahedra share a common fivefold edge. Thermal fluctuations give rise to defects, where more than or fewer than five tetrahedra share a common edge. In contrast to the case of disks, where defect pairs of opposite "sign" are equally stable, in the case of spheres, higher-order (e.g., sixfold, sevenfold) defects are more stable than lower-order (e.g, fourfold, threefold) defects. This is due to the fact that pentagonal dipyramids consist of tetrahedra that are slightly stretched rather than compressed (see Fig. 2.3b). As the temperature decreases or the density increases, defects annihilate to form more icosahedra locally. However, since icosahedra do not tile 3d space, some defects inevitably remain, even as the system approaches very low temperatures and/or high pressures.

Spheres and Disks: Liquid-Solid Transitions

Now let us consider how particle packing changes when the thermodynamic driving force (e.g., the temperature, pressure, etc.) is changed such that the system favors the solid phase over the liquid phase. Recall, from section 2.1.1 that the thermodynamic behavior of a changing system is strongly tied to the interfacial tension between the phases, which, in

the case of the Ising model, is dictated by the coupling parameter J. For systems of disks, hexagonal packing is favored both locally and globally; thus the interfacial tension between the liquid and solid phases is negligible. When the solid phase becomes more stable, all of the local defects disappear, resulting in a hexagonal crystal. Since the interfacial tension between the phases is small, the transition from a liquid to a hexagonal crystal is reminiscent of critical behavior, similar to Fig. 2.1b. Crystallization is characterized by a growing "hexatic" (i.e., hexagonal) lengthscale, that eventually diverges as the system forms a hexagonal crystal [53, 48, 54].

In systems of spheres, the inherent structural mismatch between locally-favored icosahedral packing and globally-favored crystal structures, such as fcc, hcp, and bcc packings results in a high interfacial tension. Thus when the thermodynamic driving force is changed such that the system favors the solid phase over the liquid phase, there is a free energy barrier to transitioning from a liquid to a solid, similar to the scenario depicted in Fig. 2.1b. In this case, the transition for spheres resembles typical nucleation and growth behavior. Notice that despite their outward similarity, our two example cases, disks and spheres, present two limits of packing and thermodynamic behavior; we can gain general insight from these systems that can be applied to more complex systems with analogous local/global packing scenarios, regardless of the shape of the underlying particles.

2.3 Nucleation and Growth

With a solid understanding of sphere packing, we can now revisit the concept of "nucleation and growth" in particle-level detail. Analogously to the Ising model, a system of liquid particles must form a large crystalline droplet known as the "critical nucleus" before the barrier posed by the interfacial tension between the phases can be overcome, and the system can transform from liquid to solid. This process is known as "homogeneous nucleation [16, 55, 56, 57]," and occurs due to random fluctuations within the local liquid structure. Another mechanism by which nucleation and growth can occur is through the introduction of a disturbance, such as a speck of dust, an artificial boundary, or a seed crystal into the system. This process, known as "heterogeneous nucleation," allows the nucleation barrier to be surmounted more easily because disturbances to the liquid structure typically increase the probability of forming a critical droplet [58]. Although heterogenous nucleation is more likely in experimental systems, we focus primarily on homogeneous nucleation since the unpredictable nature of the disturbance makes heterogenous nucleation difficult to characterize.

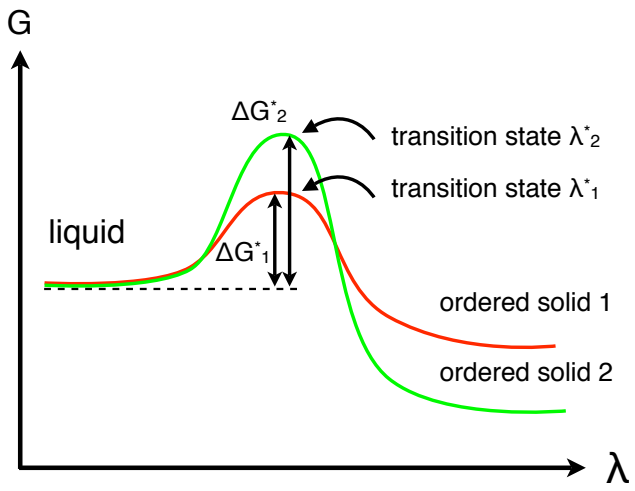


Figure 2.4 The schematic depicts the free energy G as a function of the reaction coordinate λ for two nucleation pathways. Although pathway 2 (green line) results in more stable solid than pathway 1 (red line), the system tends to follow pathway 1 since the free energy barrier to nucleation is lower (i.e., $\Delta G_1^* < \Delta G_2^*$).

Nucleation is an activated process since most nuclei are unstable and tend to revert back to the liquid. The formation of a stable nucleus requires that the free energy lost by converting particles from the liquid to the solid outweigh the free energy gained due to the interfacial tension at the nucleus surface. Thus, like a chemical reaction, nucleation is a fluctuation-driven process that requires activation energy to overcome a free energy barrier (see Figure 2.4).

A particularly important result of the free energy barrier to nucleation is that the system evolves along the reaction coordinate that poses the lowest free energy barrier [16, 55, 59]. Therefore, the ordered solid that forms need not be the stable state (i.e., the structure with the lowest free energy), only the structure with the lowest free energy barrier to formation. As a result, many systems form metastable ordered solids that minimize the free energy barrier to nucleation.

2.3.1 The Critical Nucleus

The nucleus configuration that sits atop the free energy barrier is known as the "critical nucleus." Once the critical nucleus forms, there is no free energy barrier to further growth and the system can freely solidify. The critical nucleus represents a transition state; therefore, it has an equal probability of evolving either backward or forward along the reaction

coordinate (i.e., of "committing [60]" to either the liquid or the solid phase). As the critical nucleus is essential for understanding how crystallization occurs for a given system, many studies focus on identifying critical nuclei and understanding the reasons for their stability [57, 61, 62, 21, 22].

2.3.2 Classical Nucleation Theory

A mathematical framework for modeling the formation of the critical nucleus is given by classical nucleation theory (CNT) [17, 18]. CNT models the excess free energy of nucleation ΔG as a function of nucleus radius r for a spherical, incompressible nucleus:

$$\Delta G(r) = \frac{4}{3}\pi r^3 \rho_s \Delta\mu + 4\pi r^2 \gamma \tag{2.2}$$

Here, we have assumed that we have a 3d system, where $\frac{4}{3}\pi r^3$ is the nucleus volume, ρ_s is the solid density, $\Delta\mu$ is the chemical potential difference between liquid and solid, $4\pi r^2$ is the surface area, and γ is the interfacial tension. The nucleus is approximated as being perfectly spherical, since this maximizes the volume-to-surface area ratio and thus minimizes ΔG. At the critical size r^*, ΔG goes through a maximum, indicating the transition state. Taking the derivative of $\Delta G(r)$ an equating to zero gives

$$r^* = \frac{-2\gamma}{\rho_s \Delta\mu}. \tag{2.3}$$

Inserting this value for r into our expression for $\Delta G(r)$ gives

$$\Delta G^* = \frac{16\pi}{3}\frac{\gamma}{(\rho_s \Delta\mu)^2}. \tag{2.4}$$

Although CNT provides a rough picture of the free energy barrier to nucleation, in reality, it is highly idealized, since nuclei are not, in general, perfectly spherical and the factors ρ_s, $\Delta\mu$, γ are not constants, but rather complex functions of nucleus structure and the local topology of the surrounding liquid [20, 19]. In general, the nucleus may change shape, structure, etc. as they grow in order to minimize free energy along the reaction coordinate [20, 59]. Thus, for many systems, the free energy curve depicted in Figure 2.4 may be more accurately represented by a multidimensional free-energy surface. Recently methodology has been developed to accurately measure the true reaction coordinate for

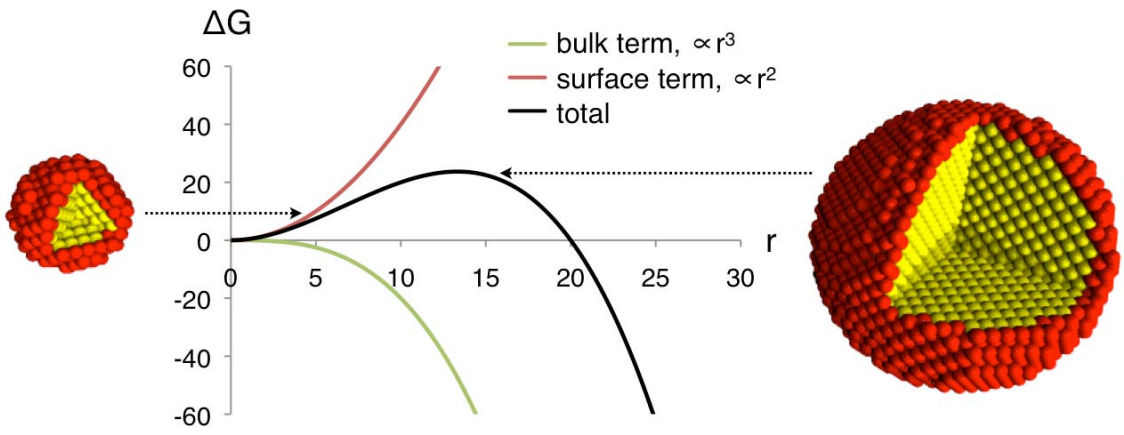


Figure 2.5 Free energy ΔG plotted as a function of nucleus radius r for a perfectly spherical nucleus. The number of particles in the nucleus core (yellow) scales with r^3 while the number of particles on the nucleus surface (red) scales with r^2. The nucleus with a radius that maximizes ΔG is known as the "critical nucleus."

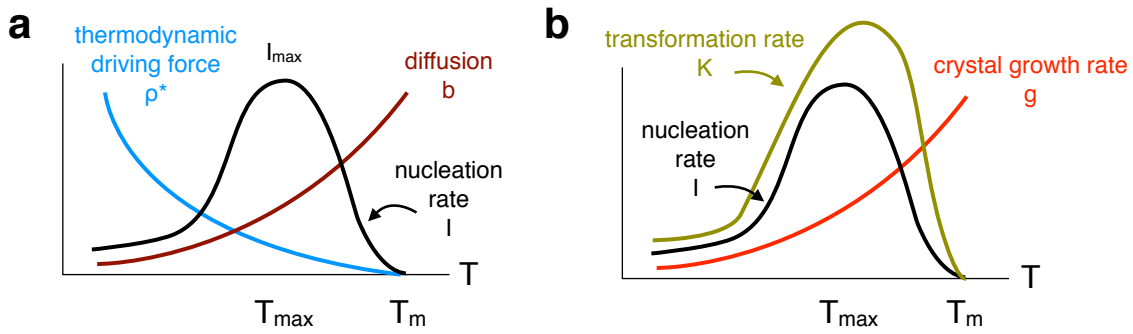


Figure 2.6 Nucleation Kinetics. (a) Schematic depiction of the nucleation rate as a function of T. At high temperatures, nucleation is thermodynamically-limited. At low temperatures, nucleation is diffusion-limited. At an intermediate temperature T_{max}, the nucleation rate is at a maximum I_{max}. (b) Schematic depiction of overall liquid-solid transformation rate as a function of temperature. The overall rate is a product of the nucleation rate and the crystal growth rate.

nucleation (and other rare events) via computer simulations (see section 4.4).

2.3.3 Nucleation Rate and Growth Rate

Now that we have established the thermodynamics of nucleation and growth, we can consider the kinetics of crystal formation. In analogy with a standard chemical reaction, we can

associate a "rate" with nucleation, which is modeled by [25]:

$$I = \rho^* b. \tag{2.5}$$

Here, ρ^* is the population density of critical nuclei, and b is the rate of particle attachment to the nucleus surface. From Boltzmann statistics, we can write

$$\rho^* = \rho_0 \exp\left(\frac{-\Delta G^*}{k_b T}\right). \tag{2.6}$$

The term ρ_0 is the density of particles in the system. Substituting our approximate expression for ΔG^* from equation 2.4 above, we obtain:

$$\rho^* = \rho_0 \exp\left[\frac{16\pi}{3}\frac{\gamma}{k_b T(\rho_s \Delta\mu)^2}\right]. \tag{2.7}$$

Similarly, we can write an expression for b according to:

$$b \propto \exp\left(\frac{-\Delta G_d}{k_b T}\right). \tag{2.8}$$

Here, ΔG_d is the activation energy for bulk diffusion. While b, decreases with temperature, we can see from equation 2.7 that ρ^* increases with T. This is depicted in Fig. 2.6a, which shows that the nucleation rate goes through a maximum I_{max} at particular T_{max}. For $T < T_{max}$, critical nuclei form easily, but I is limited by slow diffusion. For $T_{max} < T < T_m$, I is limited by the thermodynamic driving force.

The overall transformation rate for particles from the liquid phase to solids phase K is given by the product of the nucleation rate I and the growth rate g:

$$K = Ig \tag{2.9}$$

The growth rate to a critical nucleus is again driven by diffusion, modeled by an Arrhenius expression similar to equation 2.8. The overall transformation rate is depicted in Fig. 2.6b.

2.4 The Glass Transition

It is clear from Fig. 2.6b that if a liquid is supercooled sufficiently far below T_m, the rate of transformation from the liquid to solid becomes negligible and the system stays trapped in the mestastable liquid state indefinitely. At some point, the temperature becomes sufficiently

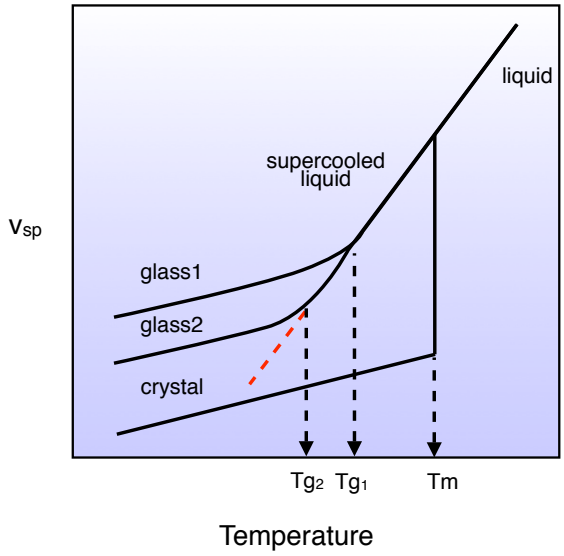


Figure 2.7 Schematic depicting the specific volume as a function of temperature for a typical supercooled liquid. Here, "glass 2" is cooled more slowly than "glass 1" and falls out of equilibrium at a lower lower glass transition temperature T_g.

low that the particles in the metastable liquid can no longer relax over the timescale of a typical experiment, and the system becomes kinetically trapped and effectively solid. The temperature at which the relaxation time crosses the experimental timescale is known as T_g, the glass transition temperature [26]. In this sense, vitrification is a purely dynamical phenomenon, since the glass and the liquid and their corresponding thermodynamics are interchangeable based on an arbitrary definition of the experimental timescale. For example, the familiar silica-based window glass, with relaxation time $\tau = 10^{32}$ years, would be defined as a liquid for experiments spanning, say, 10^{33} years [15]. However, for most supercooled liquids, the temperature dependence of the relaxation time is sufficiently strong that T_g is a well-defined material property (to within a few degrees Centigrade) and is essentially independent of any reasonable definition of the experimental timescale. Aside from the experimental timescale, T_g also has a weak dependence on the cooling rate. A liquid that is cooled more slowly has more opportunity to rearrange structurally and will fall out of (metastable) equilibrium and form a glass at a slightly lower T_g than a rapidly quenched melt with a structure more reminiscent of the high temperature equilibrium liquid (see Figure 2.7).

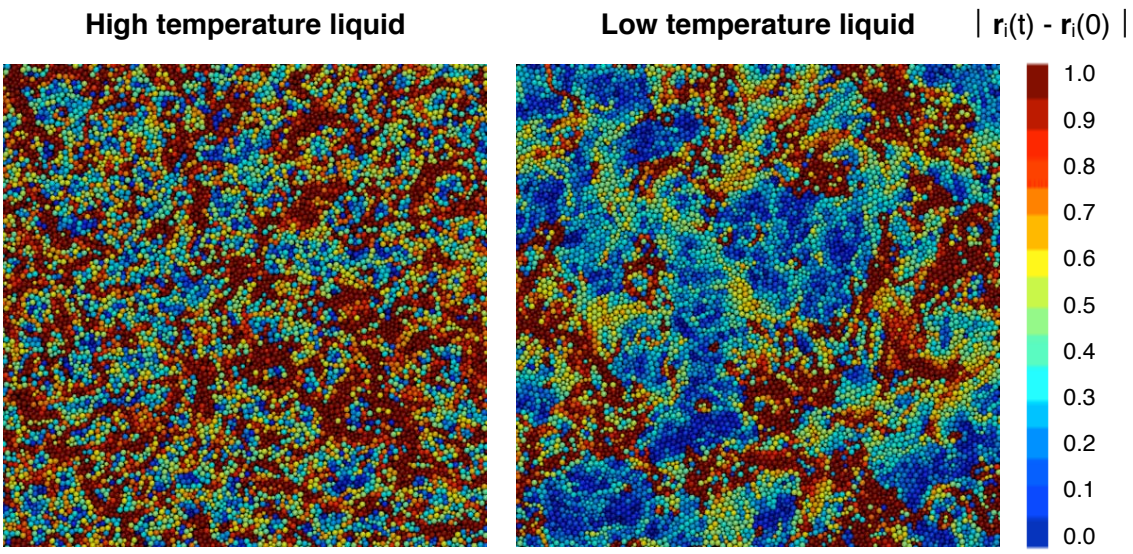


Figure 2.8 Spatially heterogenous dynamics in a 2d supercooled liquid. Particles are colored based on their displacement $|\mathbf{r}_i(t) - \mathbf{r}_i(0)|$ over a time window $t \sim \tau_\alpha$. Notice that as temperature decreases, particle mobility becomes increasingly spatially correlated.

2.4.1 Supercooled Liquid Phenomenology

The argument outlined above implies that the glassy state is simply a supercooled liquid that has lost ergodicity. However, this overly-simplified picture fails to tell the whole story; in practice, supercooled liquids exhibit several highly unusual and mostly unexplained dynamical and thermodynamic properties. In this section, we provide a cursory review these unique properties of supercooled liquids.

Spatially Heterogeneous Dynamics

Based on the arguments above, a glass can be considered a very slow supercooled liquid. That is, on long enough timescales, glasses will relax structurally and flow like liquids. Thus, we might expect that if we watched the particles in a glass in "fast-forward," they would move just like those in a liquid. Interestingly, molecular simulations [63] and, more recently, experiments [64, 65, 66] have shown that this is not the case. As liquids become increasingly supercooled, particle dynamics become increasingly correlated, or "spatially heterogeneous [67, 68]." More specifically, systems exhibit spatially heterogenous dynamics (SHD) if (1) it is possible to select a dynamically distinguishable subset of particles and (2) these dynamics are spatially correlated. Fig. 2.8 depicts SHD for a 2d supercooled liquid over a timescale of about one α relaxation time.

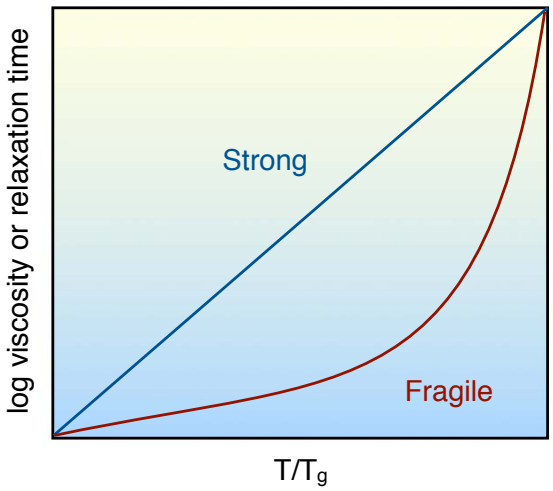


Figure 2.9 Depiction of the Angell plot [78], showing strong versus fragile relaxation behavior.

SHD has been extensively characterized in a variety of systems, including simulated atomistic and molecular glass formers [69, 70, 71, 72], and experimental systems such as colloidal suspensions [64], and, as we will show in Chapter 6, granular matter [65, 66, 31]. Hallmark features of SHD include the cooperative motion of highly mobile particles in strings [73], which aggregate into clusters [74], and the cooperative motion of relatively immobile particles as they break free from their cages [75, 76]. These features are defined in more quantitative detail in Chapter 4.

Decoupling

Another interesting property of supercooled liquids is a decoupling that occurs between the translational and the rotational diffusion coefficients and the viscosity. For relatively high temperatures, both the translational and the rotational diffusion coefficients in supercooled liquids are inversely proportional to the viscosity. For temperatures below approximately $1.2T_g$, a decoupling occurs between translational diffusion and viscosity, and between the rotational and translational diffusion [77, 25]. However, the relationship between rotational diffusion and viscosity remains intact. As a result, near the glass transition temperature particles translate much faster than expected based on their viscosity.

Super-Arrhenius Relxation

Another important phenomenon observed for some supercooled liquids is the breakdown of the Arrhenius temperature dependence of viscosity at low temperatures [26, 25]. While the viscosity of certain liquids (SiO_2, GeO_2, etc.) exhibits Arrhenius behavior for a full range of T, the viscosity of other liquids (ortho-Terphenyl, Chlorobenzene, etc.) undergoes a crossover at low T where the viscosity begins to diverge much more rapidly than an exponential function (see Fig. 2.9). Such liquids have been characterized as "strong" and "fragile," respectively in reference to how closely they adhere to an Arrhenius temperature dependence [78]. Mathematically, the relaxation behavior of strong liquids can be characterized by $\tau = A\exp(E/k_bT)$, where A and E are constants and τ is the relaxation time. One relationship that has been shown empirically to fit transport properties reasonably over 2-4 orders of magnitude for a range of fragile liquids is the Vogel-Tamman-Fulcher (VTF) equation [79, 80, 81]:

$$\tau = A\exp[B/(T-T_0)] \tag{2.10}$$

Here A and B are constants and T_0 is the glass transition temperature. In addition to the VTF equation, several alternate fits are possible for fragile glass formers [82, 83, 84, 85, 86].

Thermodynamic Singularity

It follows from our discussion at the beginning of this section that, theoretically, if a liquid is quenched slowly enough over long enough times, it will never fall out of equilibrium. That is, the system will follow the thermodynamic pathway traced out by the red dashed line in Fig. 2.7. As pointed out by Kauzmann [87], the decrease in entropy with temperature occurs at a much higher rate for a supercooled liquid following such a path than for either the glass or for the ground state crystal (see Fig. 2.10). (Notice that this can be qualitatively inferred from Fig. 2.7 as well; as explained in section 2.2, the entropy is directly related to the accessible volume per particle, which is inversely related to $v_{sp} = 1/\rho$). When extrapolated to a temperature T_K, the entropy of the supercooled liquid equals that of the crystal. This poses a paradox, since, as Kauzmann states: "*certainly the entropy of the liquid can never be very much less than that of the solid* [87]." When extrapolated further, the liquid entropy becomes negative at some $T > 0$, which is physically impossible since entropy is inherently non-negative. Kauzmann theorized that at some point, the derivative $\delta S/\delta T$ (i.e., the temperature times the heat capacity T_C must change sharply in order to avoid such an "entropy crisis." Indeed this is precisely what occurs at the glass transition (see Fig. 2.10). Kauzmann hypothesized that such a transition must occur at or above T_K, since the glass

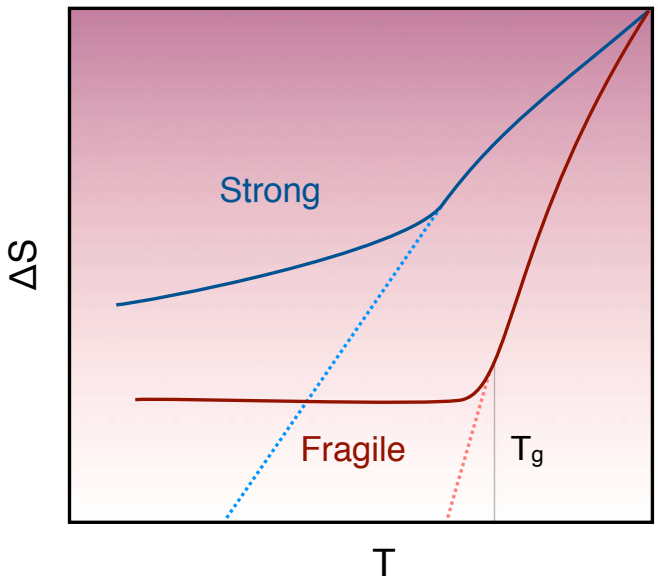


Figure 2.10 Depiction of the entropy S as a function of temperature T, similar to the plot shown in reference [24]. When the supercooled liquid falls out of equilibrium at T_g, the heat capacity, given by the derivative $1/T(\delta S/\delta T)$ changes rapidly. If the liquid could be cooled infinitely slowly such that T_g was avoided, the entropy would become negative at some finite T (dashed lines).

and crystal have approximately the same heat capacity.

Thus, it would seem that the glass transition is a thermodynamic inevitability. However, in practice, this scenario of a thermodynamic glass transition is always avoided since, in practice, supercooled liquids always either crystallize or vitrify before the Kauzmann temperature is reached. The true nature of the glass transition and the importance of the Kauzmann paradox remains a matter of intense debate within the scientific community [78, 24].

2.4.2 Theories of the Glass Transition

The unique properties of supercooled liquids outlined above have inspired many theories to explain their behavior. A full review of the theories of the glass transition, which span over 50 years of research, is beyond the scope of the current discussion. Instead, we very briefly review the theories that are most relevant to the work that we will present in the following chapters.

Potential Energy Landscape

The potential energy landscape (PEL) abstraction, originally introduced by Goldstein [88], gives an intuitive topographic picture of supercooled liquids. The PEL is not a theory of the glass transition *per se*, but rather a convenient description of liquids from which theories can be constructed. In the PEL picture, deeply supercooled liquids become temporarily trapped in local potential energy minima known as "inherent structures [89, 90]." At high temperatures, the liquid can escape energy minima easily and structural relaxation is rapid. At lower temperatures, the system becomes trapped in energy minima for longer periods of time and structural relaxation is more sluggish. The point at which the system becomes completely trapped in the same inherent structure corresponds to the "dynamical" glass transition [24].

Because inherent structures do not account for thermal vibrations, the PEL itself is independent of temperature. However, the space of minima sampled becomes constricted as temperature decreases. This idea has lead to various thermodynamic and statistic descriptions of the PEL, which have informed speculation on the nature of the mode-coupling crossover (see next section) and the Kauzmann paradox [90, 24]. The PEL picture is successful in describing the cooling rate dependence on T_g in that rapidly-cooled liquids visit regions of the PEL that are not accessible in equilibrium at lower temperatures and become kinetically trapped. The PEL picture can also explain the strong-fragile distinction, in terms of the topography of the potential energy landscape. Fragile liquids have particularly heterogeneous landscapes with large "metabasins" containing many inherent structures, whereas strong liquids do not [89, 91].

Mode Coupling Theory

A well known theory of the glass transition is given by mode-coupling theory (MCT) [92]. MCT is a first-principles approach starting only with the structure of the liquid, measured by the structure factor $S(q)$ (see section 4.5.2). The main physical effect that MCT is thought to describe properly is "caging," whereby particles in a dense system must move cooperatively to escape from their cage created by neighboring particles. At some critical temperature T_0 (or density ϕ_0), particles can no longer escape from their cage, and the system undergoes a glass transition. This is modeled by

$$\tau \propto \exp\left[\frac{E}{T - T_0}\right]. \tag{2.11}$$

The variable τ is the relaxation time. Clearly, equation 2.11 exhibits a singularity at T_0, and the relaxation time diverges. In this respect MCT is successful in that it predicts a glass transition at a nonzero temperature. Moreover, the functional form for τ vs T predicted by equation 2.11 gives reasonable agreement for many simulated liquids and colloidal suspensions, provided they are not significantly supercooled. However, in practice the predicted value of T_c always falls above the true glass transition temperature $T_c > T_g$ [93]. This is thought to result from a crossover in the relaxation mechanism of in deeply supercooled liquids, that is not accounted for in the theory. Debennedetti, Stillinger and coworkers have hypothesized that this crossover corresponds to a crossover from diffusion-dominated dynamics to energy-landscape-dominated dynamics [89, 24]. An additional drawback of MCT, in its original form, is that it fails to clearly predict SHD, or decoupling [93]. However, more recent studies have shown that a diverging dynamical lengthscale, analogous to a dynamical critical point, can be obtained within the mode coupling formalism, which is consistent with the observation of a growing lengthscale for SHD as the glass transition is approached [94, 95].

Cooperatively Rearranging Regions

More than 40 years ago, Adam and Gibbs proposed a theory of the glass transition based on the premise that supercooled liquids relax structurally through cooperatively rearranging regions (CRRs), which grow larger in size as temperature decreases [96]. The subsequent discovery of SHD is seemingly a verification of this idea [97]. The idea of CRRs has also been related to metabasin transitions within the PEL [98, 99]. The main result of the Adam-Gibbs theory is that the structural relaxation time is related to the configurational entropy, defined as the contribution to the entropy from changing configurations, but not vibrations; $S_c = S - S_{vib}$:

$$\tau \propto \exp(A/TS_c). \tag{2.12}$$

Here, A is a constant given by $A = \Delta\mu S_c^*/k_b$, where μ is the barrier hindering rearrangement, S_c^* is the configuration entropy of the minimal CRR, which can be related to the integrated excess heat capacity. When plugging in this expression for S_c^*, the Vogel-Fulcher temperature dependence is obtained.

Several theories in the spirit of Adam and Gibbs have since been proposed. For example, the random first order theory of the glass transition (RFOT) of Wolynes and coworkers improves on Adam and Gibbs by providing a microscopic definition of the CRRs activation energy, and predicts structural relaxation fits that correspond closely to physical relaxation data [100, 101]. An alternate viewpoint is given by Langer, where the strings of mobile

particles observed in many supercooled liquids may serve as CRRs [102].

Kinetically Constrained Models

In recent years, a great deal of insight into glassy systems has been attained from kinetically constrained models (KCMs). These types of models were first introduced by Fredrickson and Andersen in the 1980s [103, 104, 105]. KCMs are simple coarse-grained lattice models of particle dynamics in glassy liquids that attempt to capture the fundamental elements of glass transition phenomenology. The models typically begin with the assumptions that particle dynamics are sparse, and that dynamics are facilitated [106]. The latter assumption follows directly from spatial constraints present in dense glassy systems; particle motion creates structural defects or empty space locally, and thus "facilitates" the subsequent motion of nearby particles. Particular successes of KCMs include the prediction of a universal fit for relaxation behavior to a parabolic law [106, 107]:

$$\frac{\tau}{\tau_{\mathrm{o}}} \propto \exp\left[J^2\left(\frac{1}{T}-\frac{1}{T_{\mathrm{o}}}\right)^2\right], \quad T < T_{\mathrm{o}} \tag{2.13}$$

Here, J is an energy scale for a fundamental unit of particle motion within the system, and T_{o} is an onset temperature below which the assumptions of sparse, facilitated dynamics hold. This theoretical fit has been shown to hold for a vast range of experimental and simulated glass formers [107]. Another recent success of KCMs is the prediction of a unique first-order transition in "space-time," which is suspected to provide the thermodynamic driving force behind SHD [108, 109]. In this picture, the growing size of heterogeneous regions near the glass transition is a manifestation of the underlying first order transition. This has since been verified for a fully atomistic system [109]. Another notable success of KCMs is their ability to explain decoupling [110, 111]. In Chapter 6, we further explore that relationship between fully atomistic glass formers to KCMs.

Other Theories

Several other theories of the glass transition have been proposed that we will not explore in detail. One famous theory is the "free volume approach" formulated by Turnbull and Cohen in the 1960s [112, 113, 86]. Here, it is theorized that the glass transition results from a disappearance of excess volume, or volume that can be redistributed within the system without a significant energy change. Another interesting theory is given by the frustration

limited domains (FLD) picture of Kivelson, Tarjus and coworkers [114, 115, 116]. This theory stipulates that super-Arrhenius relaxation behavior can be related to the tendency for deeply supercooled liquids to form locally stable domains that are frustrated due to structural constrains that prevent them from spanning space. The crossover to super-Arrhenius behavior is thought to occur at the point that the liquid composed of individual particles begins to resemble a hierarchy of interacting FLDs. In addition to these theories, several other theories have been developed over the 50+ years of glass transition research that we do not have time to mention here, in the interest of brevity. However, it is important to bear these ideas in mind when studying supercooled liquids so that we can to draw connections wherever possible.

Chapter 3
Model Systems

In our study of liquid-solid transition, we employ several different model systems. In this chapter, we introduce our model systems and the basic physical characteristics that they are meant to capture. We begin by introducing some standard condensed systems, consisting of spherical particles with hard or van der Waals interactions. We then consider slightly more complex models that form polytetrahedral ordered structures, which serve as a convenient prototype for our study of metastable ordered phases in Chapter 5. Finally, we introduce some model supercooled liquids, which are used for our study in Chapter 6. Although we use several other more complex models in our studies of structural metrics in Chapter 7, we forgo a discussion of those models here.

3.1 Simple Models of Condensed Matter

Since we are interested in studying general phenomena rather than modeling specific systems, we require models that capture the basic physical behavior of interest, but minimize computational effort and complexity. For our study of liquid-solid transitions, we are particularly interested in systems that solidify via nucleation and growth, since they are more likely to form metastable crystals and deeply supercooled liquids than those that solidify by spinodal-decomposition (see section 2.1.1). Per our discussion of section 2.2.1, it is clear that 3d systems of spherical particles can be supercooled and form ordered solids via nucleation and growth. We briefly outline some standard models of 3d spherical particles below.

Hard Sphere Potential

Perhaps the simplest model of a condensed system is given by the hard-sphere model. The hard potential accounts for excluded-volume interactions between particles [117]. For

spherical particles with diameter σ, the potential is defined as

$$U_{HS}(r) = \begin{cases} \infty & r < \sigma \\ 0 & r \geq \sigma \end{cases}. \tag{3.1}$$

If the distance between particles r is less than or equal to the particle diameter σ, the potential is infinite, and zero otherwise. The hard-sphere potential is depicted in Fig. 3.1a. Although many condensed systems involve particles with some interactions, such as van der Waals, bonded, electrostatic, depletion interactions, etc., hard spheres are often used as a minimal model for condensed systems, and, in many cases, capture a large amount of the relevant behavior [118] (see section 2.2). Because the hard-sphere potential is discontinuous, it is typically implemented in Monte-Carlo simulations rather than molecular dynamics simulations (see section 4.1.2).

Lennard-Jones Potential

The Lennard-Jones (LJ) potential accounts for both excluded-volume interactions and van der Waals attractions between particles [117]. The "12-6" form of the potential (see Figure 3.1) is most common:

$$U_{LJ}(r) = 4\varepsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right]. \tag{3.2}$$

Here, r is the separation between two particles, σ is the particle diameter, ε is the well depth parameter, which is proportional to the inverse of the temperature $1/k_BT$. The first term, proportional to $1/r^{12}$, is chosen to give the particles a strong, yet continuous excluded volume interaction, which is ideal for MD simulations. The second term accounts for the van der Waals attractive interaction, which is determined theoretically to be proportional to $1/r^6$. The LJ potential is depicted in Fig. 3.1b.

Weeks-Chandler-Andersen Potential

In many cases, we require a soft-sphere potential similar to the LJ potential, but with no attractive interaction. One way to accomplish this is to simply remove the $1/r^6$ term from the LJ potential. Another common soft-sphere potential is given by the Weeks-Chandler-Andersen

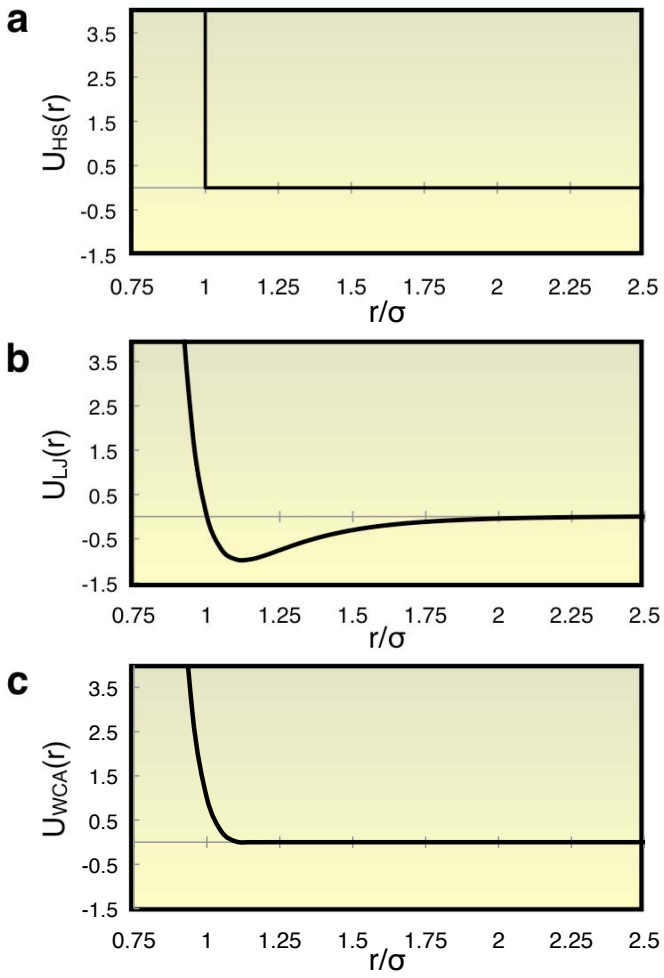


Figure 3.1 Condensed matter pair potentials. (a) The hard sphere (HS) potential. (b) The Lennard-Jones (LJ) potential. (c) The Weeks-Chandler-Andersen (WCA) potential.

(WCA) potential [118]:

$$U_{WCA}(r) = \begin{cases} U_{LJ}(r) - U_{LJ}(r = 2^{1/6}\sigma) & r \leq 2^{1/6}\sigma \\ 0 & \text{otherwise.} \end{cases} \tag{3.3}$$

The WCA potential is a truncated and shifted instance of the LJ potential, where the potential is cutoff at the minimum of the LJ potential, $2^{1/6}\sigma$. The WCA potential is depicted in Fig. 3.1c.

3.2 Metastable Ordered Solids

As mentioned previously, one of our primary goals in this body of work is to investigate the formation of metastable ordered solids. The models mentioned above are sufficient for observing crystallization in supercooled liquids, and have been used in many studies of nucleation and growth [57, 62]. However, these models tend to form thermodynamically stable crystals from the melt. To begin to formulate a model for forming metastable crystal structures, recall that from section 2.3 that metastable ordered solids arise when a particular crystal structure has a lower free energy barrier to nucleation than the stable crystal. The relative size of the nucleation barrier for different structures is dictated by the interfacial tension between the liquid and solid phases, which arises due to a structural mismatch between the phases. As outlined in section 2.2, the liquid structure in the 3d systems of model spheres reviewed above is typically icosahedral, whereas the solid-phase structure is close-packed, giving rise to a high interfacial tension. In some cases, specific interaction potentials can be designed to stabilize solid phases that retain some icosahedral character in the solid phase [119, 120]. Since these solid phases are structurally similar to the icosahedral liquid, they are good targets for forming metastable states. In this section, we review a particular class of solid structures with icosahedral local ordering known as Frank-Kasper (FK) structures [121, 47], which will serve as the basis for our studies of metastable solid phases.

3.2.1 Frank-Kasper Phases: Formation

To understand FK structures, recall from section 2.2.1 the process that occurs when we cool a liquid composed of 3d spherical particles. The liquid structure can be understood in terms of local deviations from ideal icosahedral ordering [44] (Fig. 3.2a). As the liquid is cooled, these defects disappear, but, due to packing constraints some defects inevitably remain, even at low temperatures. As a result, the liquid consists of a mixture of icosahedra and other defective coordination shells known as FK polyhedra (Fig. 3.2b). The stability of the different FK polyhedra can be rationalized in terms of minimizing defects [44, 47]. The icosahedron is the optimal coordination shell, since icosahedra contain no defects. The second most stable coordination shell is the Z14 configuration, which contains two relatively favorable six-fold defects. The third most stable packing is the Z15 configuration, which contains three six-fold defects. The fourth most stable packing is the Z13 configuration, which contains two six-fold defects and one unfavorable four-fold defects. Other configurations containing more defects, and higher-order three-fold or seven-fold defects, are also possible at high

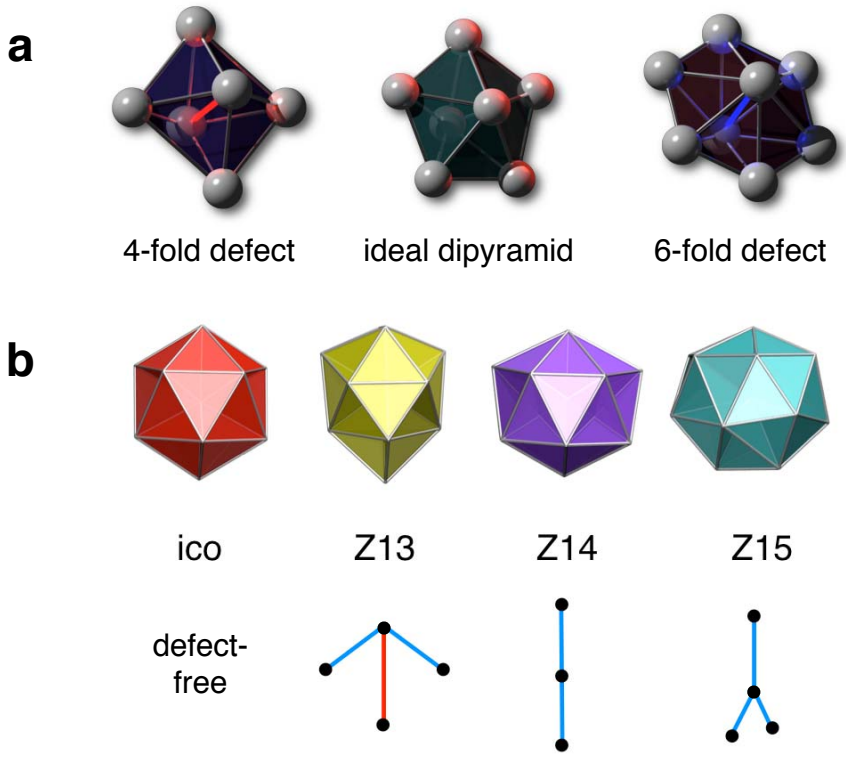


Figure 3.2 Frank-Kasper polyhedra [44, 47]. (a) Depiction of "disclination defects [47]." The pentagonal dipyramid is the ideal structure. The 6-fold defect is favored over the 4-fold defect. (b) Depiction of FK polyhedra with coordination number Z, which form the coordination shell of a particle at the center of the polyhedron. Neighboring particles sit at the vertexes. The defect network corresponding to each shell is depicted below.

temperatures (see Fig. 3.3a), but are very improbable at lower temperatures [47].

Even at zero temperature, 3d polytetrahedral systems inherently contain a positive number of defects since icosahedra can not tile space. A given system can minimize these defects by arranging FK coordination shells into an ordered crystal [121]. These systems minimize energy by (1) ensuring that all defects are of the relatively favorable sixfold variety and (2) arranging defects into an ordered network [47]. In order for the system to arrange defects into an ordered network, the particles themselves must arrange into an ordered structure. The defect network may exhibit either periodic or random connectivity. In the case that the defect network is periodic, the resulting FK crystal is periodic as well. In the case that the defects form a network with random connectivity, the system forms a quasicrystal [122, 123, 119], an ordered solid with long-range orientational ordering, but no translational symmetry. The formation of a defect network for a quasicrystalline FK structure is depicted in Fig. 3.3b. The network is layered such that when viewed from the depicted angle, the defects form stacks into the page. The nodes at which 3 defects meet are Z15 coordination shells. The

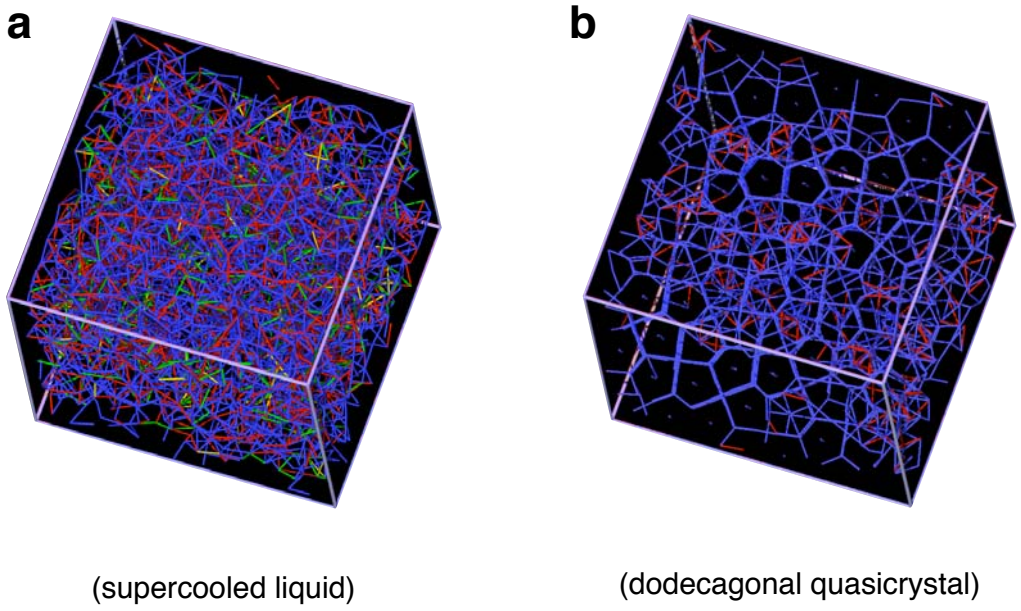


Figure 3.3 Disclination defects [44, 47] in a system forming an FK quasicrystal. In the liquid, many defects are present. Defects are colored based on their type (3-fold = yellow, 4-fold = red 6-fold = blue, 7-fold = green). After the system transforms to an ordered solid, the mostly 6-fold defects remain. These defects form an ordered network with random connectivity, which corresponds to a 12-fold quasicrystalline real-space structure.

nodes that appear as dots (pointing out of the page) are Z14 configurations. Notice that almost all defects are the 6-fold variety. Other 4-fold defects occur, but only due to thermal fluctuations – many ideal FK crystals do not contain any 4-fold defects.

3.2.2 Frank-Kasper Phases: Structure

Many different types of crystals can result from these ordered defect networks. The structures can be characterized by connecting columns of Z14 particles, which appear as dots in Fig. 3.3b, since they are lines perpendicular to the plane. In the real-space structure, these dots correspond to the centers of the Z14 atoms, which look like 12-fold rings when viewed into the plane. Connecting the rings gives a tiling pattern that can be used to interpret the long range structure of the system. Fig. 3.4a shows the tiling pattern for a dodecagonal quasicrystal. There are several different types of tiles that can form in the system [52], as depicted in Fig. 3.4b. FK structures with periodic arrangements of these tiles are known as "periodic approximants" of the quasicrystal. Some simple approximants are shown in

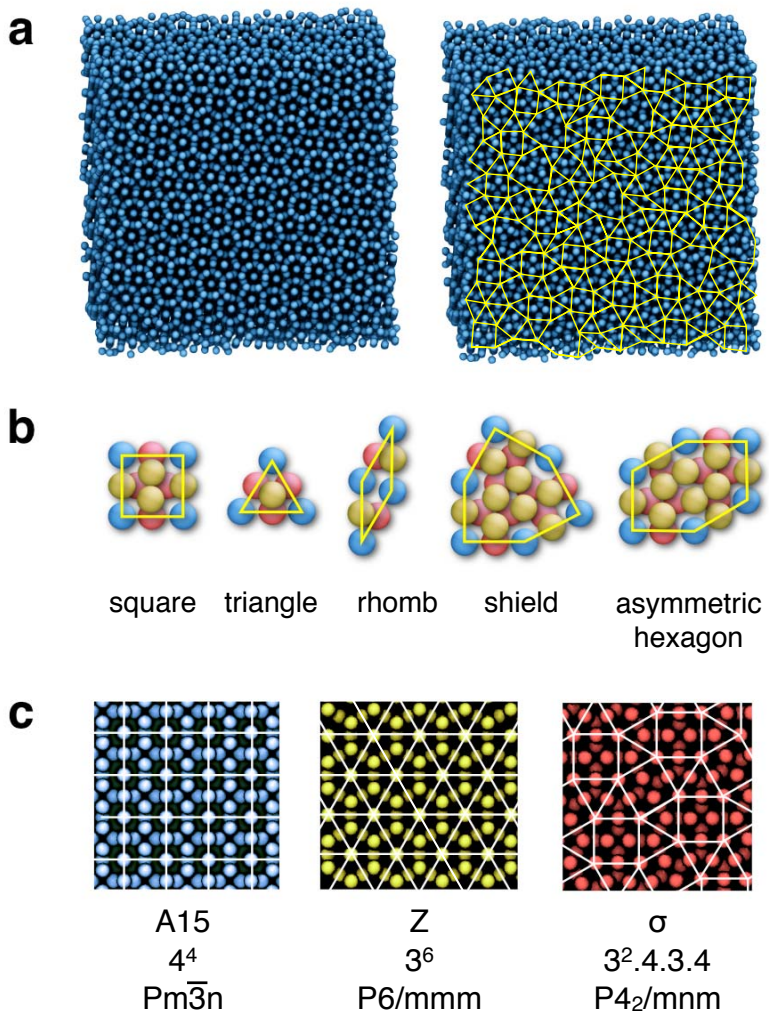


Figure 3.4 FK phases. (a) A dodecagonal quasicrystal [119] FK structure shown with (right) and without (left) tiles. The tiles are formed by connecting the centers of the 12-member "rings" of particles, which are in the Z14 configuration. (b) Valid tiles for FK structures [52]. The particles at the centers of the 12-member rings are colored blue. (c) Quasicrystal approximants. Approximants are simple FK crystals with periodic arrangements of the tiles. The three most simple approximants, the A15, Z, and σ structures, are shown along with their vertex configuration, and space group.

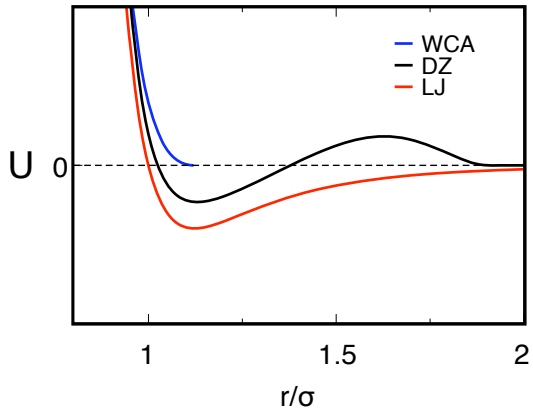


Figure 3.5 Dzutugutov (DZ) potential [127]. The potential is plotted along with the LJ and WCA potentials for comparison.

Fig. 3.4c. The tilings for these structures correspond to Archimedean tilings [124, 32], simple tilings with a single vertex configuration. Archimedean tilings have been observed in many other systems [124, 125], and are the focus of perspective article that we review in Appendix A.

Since quasicrystals and approximants are identical locally, they often exhibit many similar physical, chemical, and optical properties [126]. In practice, quasicrystals and approximants often form under the same conditions and whether one obtains a quasicrystal or approximant may depend on minor differences in the preparation method, such as the cooling rate [126, 123]. Quasicrystals and approximants differ only in their long range rotational symmetry, where quasicrystals takes on forbidden crystallographic symmetries such as 8-fold, 10-fold, or 12-fold (as is the case for the dodecagonal quasicrystal that we have described above). In Chapter 5, we investigate the difference in the growth mechanism that causes systems to form quasicrystals versus approximants.

3.2.3 Dzugutov Potential

We can use simple model pair potentials to study FK crystals and quasicrystals in molecular simulations. One simple pair potential that is known to produce FK crystals is the Dzugutov potential [127, 52]. The Dzugutov potential is similar to the standard LJ potential, except that it includes an additional repulsive interaction centered at $\sim 1.628\sigma$ (see Fig. 3.5). By design, the repulsion coincides with the interatomic spacing for fcc, hcp, and bcc crystals, thus increasing the energy for these structures and hindering their formation [52]. As a result the system does not form these crystals, but rather forms FK crystals and quasicrystals at

many statepoints [119, 52]. In addition to modeling FK phases, the Dzugutov potential has been successfully applied to modeling metallic liquids [127, 119, 128, 129, 70]. Thus, this simple potential provides a relatively realistic model system in which to study the formation of a metastable solid phase.

The Dzugutov potential is given by [127]:

$$U_{DZ}(r) = A\left(r^{-m} - B\right) exp\left(\frac{c}{r-a}\right)\theta\left(a-r\right) + Bexp\left(\frac{d}{r-b}\right)\theta\left(b-r\right). \tag{3.4}$$

Here, Θ is the Heaviside step function, given by:

$$\theta\left(x\right) = \begin{cases} 1 & x \geq 0 \\ 0 & otherwise \end{cases}. \tag{3.5}$$

The standard form of the Dzugutov potential is implemented using $A = 5.82$, $B = 1.280$, $a = 1.87$, $b = 1.94$, $c = 1.10$, $d = 0.27$, and $m = 16$. Other variations of similar potentials are also known to form FK structures and quasicrystals [130, 131]. For example, reference [131] reported the formation of dodecagonal quasicrystal approximants in the Lennard-Jones Gauss system, which is similar to the DZ system, but with purely attractive particle interactions for $r > \sigma$.

3.2.4 Other Models

In addition to the Dzugutov potential, we explore other models that can be used to form FK crystals and quasicrystals. These include models with more complex particle shapes and functionality, which are specifically applicable to colloidal and nanoscale systems. A detailed study of these models is provided in Chapter 5.

3.3 Glassy Liquids

As mentioned in section 3.1 above, several simple pair potentials can be used to model supercooled liquids. Historically, many studies of supercooled liquids have been based on these models [132, 47, 57, 62]. However, systems such as the HS, LJ, and WCA systems have the drawback that they eventually crystallize, making it impossible to study the dynamics of these systems at high degrees of supercooling.

One solution to this problem is to introduce size polydispersity into the system, which

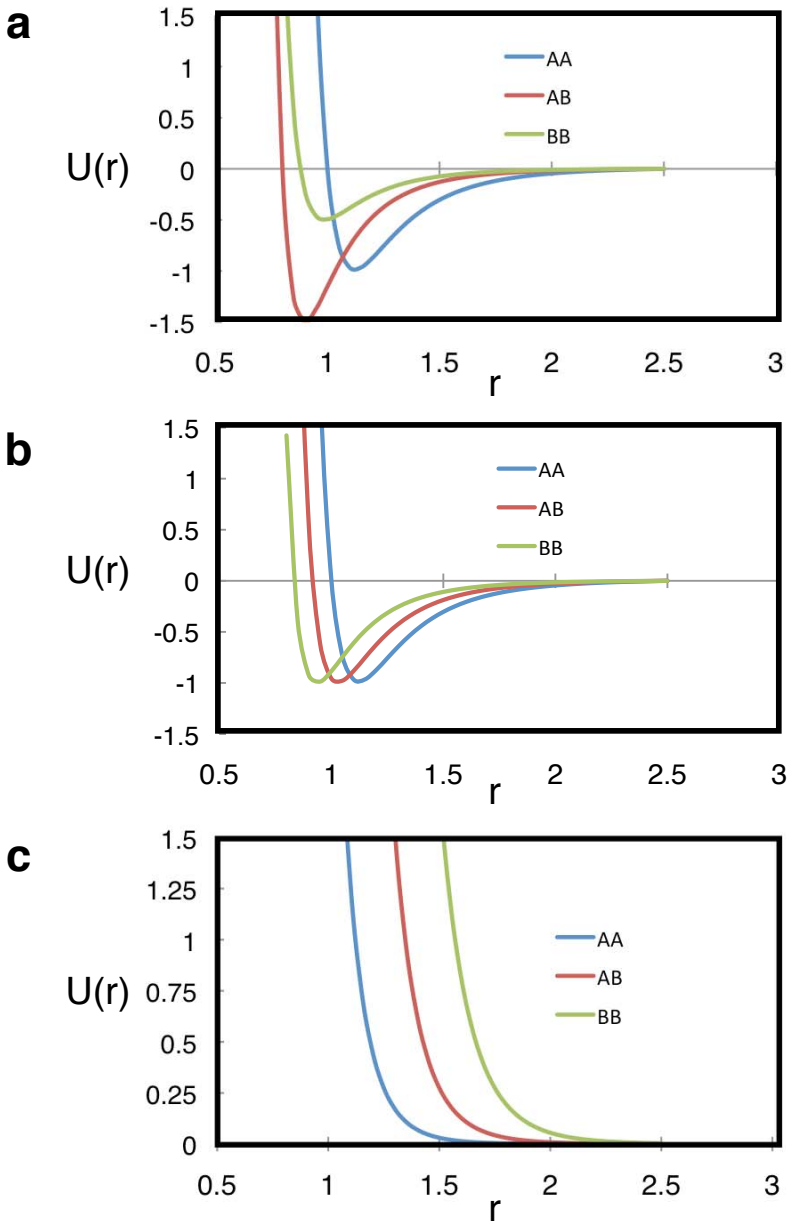


Figure 3.6 Binary mixture models of glassy liquids. (a) The KA system [133]. (b) The Wahnström system [134]. (c) The Harowell 2d soft sphere system [135].

frustrates crystalline packing and increases the nucleation barrier [61]. A simple way to introduce this type of polydispersity into a system is to consider binary mixtures consisting of two different particle sizes [133, 134]. This is computationally simpler to implement than, say, systems with continuous polydispersity [61], where particle diameters are taken from a normal distribution; in the binary case, there are only three different pair interactions to define, for AA, AB, and BB pairs. One common binary system is given by the Kob-Andersen (KA) binary LJ system [133], shown in Fig. 3.6a. The parameters for the KA system are shown in the table below. Notice that this potential includes a non-geometric mixing rule for particle diameters (that is, $\frac{1}{2}(\sigma_{AA} + \sigma_{BB}) \neq \sigma_{AB}$). A similar binary LJ model with a geometric diameter mixing rule is given by the Wahnström system [134]. Notice that this system includes two different particle masses. According to WCA theory, the attractive portion of the LJ potential has little bearing on the overall behavior of the system [118]; thus we often construct WCA versions of these LJ systems by cutting off and shifting each potential at $2^{1/6}\sigma$. This saves computational effort, since it requires fewer force calculations (see section 4.1.4). We study the KA and Wahnstrom systems, as well as their WCA counterparts in the context of drawing connections between fully atomistic liquids and kinetically constrained models in Chapter 6

	d	$\mathbf{f_A}$	$\sigma_{\mathbf{AA}}$	$\sigma_{\mathbf{AB}}$	$\sigma_{\mathbf{BB}}$	$\varepsilon_{\mathbf{AA}}$	$\varepsilon_{\mathbf{AB}}$	$\varepsilon_{\mathbf{BB}}$	$\mathbf{m_A}$	$\mathbf{m_B}$
KA [133]	3	0.8	1.0	0.8	0.88	1.0	1.5	0.5	1.0	1.0
Wahnström [134]	3	0.5	1.0	11/6	5/6	1.0	1.0	1.0	2.0	1.0

A similar method can be applied to study glassy dynamics in 2d systems, where monodisperse systems of disks easily form hexagonal crystals (see section 2.2.1). Binary mixtures are often studied with particle diameter ratios 1.4:1 [136, 135, 137, 31] (see Fig. 3.6c). This prevents the particles from forming equilateral triangles locally, and thus prevents hexagonal ordering from persisting over long ranges. In these systems, the minimum free energy state is typically a phase-separated structure consisting of two pure-component hexagonal crystals with an interfacial layer between. However, the free energy barrier for forming such phases is often sufficiently high that the system remains metastable over long timescales. The liquid is characterized by short ranged hexagonal order, where particles form as many locally favorable hexagonal configurations as possible [138, 139]. As long as the system remains well-mixed, larger clusters will become frustrated and tend to shrink. We explore this type of system in the context of a granular system of air-driven ball-bearings in Chapter 6.

Chapter 4

Methodology

Throughout this work, we use a broad range of molecular simulation and data analysis techniques. In this chapter, we review the methodology that we will employ. We first review the basic principles of molecular dynamics [117] (MD) and Monte-Carlo (MC) simulations [140]. We then provide an overview of some more specialized techniques used to study rare, activated events. Examples of rare events that we are interested in include nucleation and growth, and particle diffusion in glassy liquids. Understanding transitions often requires that we analyze the structure and dynamics of particle systems. Thus, we review some standard analysis techniques used to characterize structure and dynamics in standard condensed systems. In many cases, these techniques are too simplified for treating the complex structures encountered in studies of self-assembly; thus, we return to the problem of characterizing structural transitions for these complex systems in Chapter 7.

4.1 Molecular Simulations

For the purposes of this study, we employ both molecular dynamics (MD) and Monte-Carlo (MC) simulations [117, 140]. These methods represent the fundamental basis from which more specialized molecular simulation techniques are often derived. In this section, we review the basic principles of MD and MC simulations.

4.1.1 Molecular Dynamics

MD is an intuitive approach to simulating systems of particles. Here, particles in a computer simulation interact according to the laws of Newtonian physics, and therefore inherently adhere to the laws of statistical mechanics. A rudimentary MD simulation involves calculating the instantaneous forces on each particle in the simulation cell according to and then applying Newton's equations of motion to predict the future state of the system [117].

The force on each particle i can be computed from $\mathbf{f}_i = -\vec{\nabla}E_i$, where $\vec{\nabla}E_i$ is, for example, obtained by differentiating a pair potential defined in the previous chapter:

$$\mathbf{f}_i = \vec{\nabla}E_i = \sum_{j=1}^{N} \frac{\delta U}{\delta r}(|\mathbf{x}_i - \mathbf{x}_j|). \tag{4.1}$$

Integrating the system forward in time involves applying a numerical approximation, since it is not feasible to solve the system of differential equations that describe the mechanics of the many-body system analytically. The equations of motion are integrated by choosing a discrete timestep as an integration variable and stepping piecewise throughout time. One common numerical integration scheme is given by the Velocity-Verlet algorithm [140]:

$$\begin{aligned} \mathbf{x}(t+\Delta t) &= \mathbf{x}(t) + \mathbf{v}(t)\,\Delta t + \frac{1}{2}\mathbf{a}(t)(\Delta t)^2, \\ \mathbf{v}(t+\Delta t) &= \mathbf{v}(t) + \frac{1}{2}[\mathbf{a}(t) + \mathbf{a}(t+\Delta t)]\Delta t. \end{aligned} \tag{4.2}$$

Notice that the accelerations $\mathbf{a}$ are computed by $\mathbf{f}/m$. The Velocity-Verlet scheme uses the average of the force at two different times to update the velocities and is more stable than, say, standard Euler integration [140]. The pseudocode for a rudimentary MD algorithm using a pair potential and the Velocity-Verlet integrator is given below:

```
//algorithm for NVE molecular dynamics:
//initialize variables:
set x = init_particle_positions()
set v = init_particle_velocities()
set m = init_particle_masses()
set dt = init_integration_timestep()

//initialize forces:
for i in num_particles
   set f(i) = 0.0
   for j in num_particles
      if (i != j)
         set f(i) += compute_fpair(i, j)
      end if
   end for
end for

//do timesteps:
for t in num_timesteps
   for i in num_particles
      set x(i) = x(i) + v(i)*dt + 0.5*f(i)/m(i)*dt^2
      set vhalf(i) = v(i) + 0.5*f(i)/m(i)*dt
   end for
   for i in num_particles
      set f(i) = 0.0
      for j in num_particles
```

```
            if (i != j)
               set f(i) += compute_fpair(i, j)
            end if
         end for
      end for
      for i in num_particles
         set v(i) = vhalf(i) + 0.5*f(i)/m(i)*dt
      end for
   end for
```

The former MD algorithm samples the microcanonical (NVE) ensemble, where the number of particles (N), the volume (V) and the total energy (E), determined by the sum of the initial kinetic and potential energy of the particles, are all conserved. The algorithm requires additional steps to sample different ensembles, such as the canonical (NVT) and isothermal-isobaric (NPT) ensembles; however, we postpone this discussion until section 4.1.3.

4.1.2 Monte-Carlo Methods

An altogether different approach to simulating particle systems is given by the MC method, which is based on applying the laws of statistical mechanics directly to sample statistical ensembles [140]. This contrasts with MD, which assumes that the laws of Newtonian physics hold and satisfies the laws of statistical mechanics as a natural outcome. In contrast, MC schemes involve generating configurations at random and then accepting or rejecting them with a weight that is proportional to an underlying statistical distribution function. For example, we know from the Boltzmann distribution law [14] that the probability of observing a microstate ν in the canonical (NVT) ensemble $P_\nu = \exp(-\beta E_\nu)/Z$, where β is the inverse temperature and Z is the partition function $Z = \sum_i \exp(-\beta E_i)$. Thus, we can formulate a naive MC algorithm by randomly generating configurations ν, measuring their energy E_ν and accepting them into our simulated NVT ensemble with a weight proportional to P_ν:

$$h_\nu = \begin{cases} 1 & \text{if } P_\nu / P_{E_\nu} > \mathrm{R}[0,1] \\ 0 & \text{otherwise.} \end{cases} \tag{4.3}$$

Here, h_ν is a binary number indicating whether ν is accepted into our ensemble and P_{E_ν} is the probability of generating a configuration with energy E_ν at random. In this sense, MC is not a "simulation" method, but rather a numerical sampling scheme.

The naive MC algorithm described above is not applicable in practice, since we can not enumerate the partition function Z for any non-trivial system. (Notice that if we could enumerate Z, this would, in many cases, preclude running a simulation in the first place). An elegant solution to this problem is to reformulate our sampling criterion as a ratio of

probabilities, thus eliminating Z. The relative probability of observing two microstates ν_i and ν_j in our ensemble is

$$\frac{P(\nu_i)}{P(\nu_j)} = \frac{\exp(-\beta E_{\nu_i})}{\exp(-\beta E_{\nu_j})} = \exp[-\beta(E_{\nu_i} - E_{\nu_j})] \tag{4.4}$$

Therefore, if we start with a microstate ν_i and make a small perturbation such that we generate a new microstate ν_j, the probability of accepting this change is given by $P(\nu_i \rightarrow \nu_j) = \exp[-\beta(E_{\nu_j} - E_{\nu_i})]$. In practice, this can be implemented by drawing a random number R and accepting the move if

$$R[0,1] < \min\{1, \exp[-\beta(E_{\nu_j} - E_{\nu_i})]\}. \tag{4.5}$$

For our algorithm to obey detailed balance, we should perturb our microstates consistently, such that $P_{\nu_i} P(\nu_i \rightarrow \nu_j) = P_{\nu_j} P(\nu_j \rightarrow \nu_i)$. This scheme is known as the Metroplis MC (MMC) algorithm [140]. A short pseudocode for the MMC algorithm is given below.

```
//algorithm for Metroplis Monte Carlo (particle translations only):
set x = init_particle_positions()
set dx_max = init_max_perturbation()

for t in num_mmc_steps
   for i in num_particles
      set x_old = x(i)
      set e_old = 0.0

      //compute old energy
      for j in num_particles
         if (i != j)
            set e_old += compute_epair(i, j)
         end if
      end for

      //perturb particle position
      x(i) += rand(-1, 1)*dx_max

      //compute new energy
      set e_new = 0.0
      for j in num_particles
         if (i != j)
            set e_new += compute_epair(i, j)
         end if
      end for

      //apply MMC acceptance criterion
      if (rand(0,1) < exp(-beta*(e_new - e_old))
         set x(i) = x_old
      end if
end for
```

Notice that whereas MD involves generating a series of configurations that are sequential in time, MC involves generating random configurations that are non-sequential. Thus MC is not typically applied to studying particle dynamics, although it is possible in some cases, if proper care is taken [141], or hybrid methods are employed [142]. Aside from this difference, the two methods yield a very comparable level of information.

4.1.3 MD and MC in Various Ensembles

When studying liquids and solids, we are typically interested in simulating the NVT or NPT ensembles. The modifications that must be made to the schemes outlined above in order to to simulate these ensembles are covered elsewhere [140, 143]; therefore, we will not repeat them here. However, when using these methods, it is useful to have a basic mathematical insight into how they are implemented, either for the sake of troubleshooting, or for inventing new methods. Methods for extending MD and MMC to other ensembles, in general, fall into one of two categories: (1) methods based on fictitious Newtonian mechanics and (2) methods based on stochastic processes. These are the mathematical tools underlying conventional MD and MC, respectively.

Schemes based on fictitious Newtonian dynamics involve associating fictitious Newtonian variables (positions, velocities, masses and driving forces) with abstract objects in the system. Examples of such abstractions include thermostast or barostats that are coupled to measurable variables in the system, such as the temperature or pressure. Fictitious Newtonian variables need not represent physical objects like particles; rather they are purely mathematical abstractions with no physical analogue. In the context of a computer algorithm, it makes little difference whether the variables can be mapped to physical objects or not; in this sense Newtonian dynamics is simply a numerical sampling scheme. Since the underlying Hamiltonian for such a system contains particles, as well as other (abstract) objects, these methods are often called "extended system methods."

One classical example of this type of scheme is the Nose-Hoover thermostat [144], which rescales the particle velocities based on the thermostat position so that the system strictly samples the NVT ensemble. When the measured temperature deviates from the thermostat setting, this introduces a driving force that changes the thermostat position, which in turn determines how the velocities will be rescaled. An analogous method is given by the NPT MD scheme of Andersen-Hoover [145]. In this case, a fictitious "barostat" is coupled to the system pressure, which is adjusted by rescaling the simulation cell. Several other examples of schemes based on Newtonian dynamics are present throughout the literature. One famous example is given by the Car-Parinello MD (CPMD) scheme for simulating

quantum systems [146]. In this case, the fictitious Newtonian variables are coupled to the error in solving for the instantaneous electronic force field via density functional theory, which is then used to integrate the equations of motion in MD. With all of these schemes, the main adjustable parameter is the mass of the fictitious objects; if the objects are too "light" they will react too quickly to deviations between the ideal and measured coupled variable. If they are too "heavy," they will react too slowly and exhibit lag. In either situation, this can cause erroneous sampling of the statistical ensemble; thus care must be taken to tune the parameters such that the desired ensemble is produced [140].

An alternative method for sampling non-standard statistical ensembles is to apply stochastic processes to the system to mimic the effect of a thermodynamic variable. For example, the NVT MD scheme of Andersen involves simulating random collisions with a heat bath by drawing velocities from the expected Gaussian distribution and assigning them to particles at random [140]. Similar schemes have been used in more specialized cases; for example, reference [147] uses a stochastic collision scheme to control temperature and pressure in a non-periodic system with a finite crystal grain. In MMC simulations, alternate ensembles can be sampled by modifying the Hamiltonian used for accepting trial moves (i.e., although we used $\mathscr{H} = E_V$ for the NVT ensemble, other $\mathscr{H}$ can be used). Several other types of reversible trial moves can be attempted in analogy with trial particle moves in conventional NVT MMC to sample ensembles of microstates that correspond to these underlying Hamiltonians. For example, to sample the NPT ensemble, trial simulation cell expansion/compression moves are attempted to constrain the average system pressure to a set value [140].

4.1.4 Optimizing Simulations

The pseudo codes for MD and MC simulations given above make no assumptions about the form of the intermolecular force-field, other than that it is pairwise additive. However, for the systems that we will consider in our studies here, we can safely assume that the force-fields are short-ranged (see section 3.1). That is, particles only interact with their neighbors within a short cutoff range; all other interactions can be ignored. Constraining our potentials to short-range interactions allows us to significantly optimize the force and energy calculations for MC and MD simulations. Consider the problem of computing the energy for a particle *i*:

```
//algorithm for computing energies or forces using O(N^2) calculation:
set e_i = 0.0
for j in num_particles
   if (i != j)
```

```
      set e_i += compute_epair(i, j)
    end if
end for
```

This computation scales with the number of particles N; thus performing the same computation for all N particles is of the order $O(N^2)$. For a short ranged potential, we can grid space into a "cell list [117]." We can then constrain our potential calculation to particles in nearby cells:

```
//algorithm for computing energies using a cell list:
set e_i = 0.0
set cell_i = get_cell_containing_particle(i)
for cell_j in get_neighbors_of(cell_i)
   for j in get_particle_list(cell_j)
      set e_i += compute_epair(i, j)
   end if
end for
```

In this algorithm, we include each cell in its own list of neighboring cells such that pair interactions are computed for particles within the same cell as well as neighboring cells. In practice, the cell size can be adjusted for computational speed. A smaller cell size results in fewer pair calculations outside of the interaction range, since the range of positions that a particle can occupy is narrower in a small cell than for a larger cell. Additionally, smaller cells are advantageous in that the neighboring cells more closely represent the spherical shape of the interaction range than do larger cells. However, if the cell size is small and the number of cells is large, the list of neighboring cells within the interaction range for each cell will be also large. This results in large overhead for retrieving neighboring particles from cells. Thus, the optimal size for a given system must be tested via trial and error.

For MD simulations involving condensed systems, particles typically do not move very far over a given timestep. Therefore, the list of particles that a given particle interacts with changes little as well. In many cases, MD simulations can be greatly optimized by periodically constructing a "Verlet list" (also called a "neighbor list") so that, for a given particle, the same neighbors are looped over to compute the forces for several timesteps [117]:

```
//algorithm for computing forces using a neighbor list:
for t in num_timesteps
   ...
   for i in num_particles
      set r = distance(x(i), x0(i))
      if r > r_cut + r_buff
         call rebuildnlist(nlist)
         set x0 = x
      end if
   end for

   for i in num_particles
```

```
        set f(i) = 0.0
        for j in nlist(i)
              set f(i) += compute_fpair(i, j)
        end for
      end for
      ...
  end for
```

Building a neighbor list involves constructing a list of all particles within a range $r_{cut} + r_{buff}$, where r_{buff} is a small buffer. In many cases, the process of building the neighbor list may be optimized by using a cell list to compute the distances. The neighbor list can be reused until a particle has moved far enough to exceed the buffer; at this point the list must be rebuilt. The choice of the buffer can be optimized for computational speed; a buffer that is too small will require frequent rebuilds, whereas a buffer than is too large will result in extra force computations.

In addition to using optimized software algorithms such as cell lists and neighbor lists, another method is to optimize codes for the particular hardware that code runs on. MD simulations with pair forces are highly data-parallel; that is each MD timestep can easily be broken down into several smaller computations that can be carried out simultaneously [148, 149]. One method for parallelizing MD simulations is to split the neighbor list, force, and integration computations over several CPUs. For example, the LAMMPS MD package uses a powerful spatial decomposition algorithm to achieve excellent scaling with the number of CPUs, even for very large systems [148]. MD simulations with system sizes N on the order $\sim 1 \times 10^5$ to $\sim 1 \times 10^7$ can be greatly optimized by running simulations of graphics processing units (GPUs) [149]. In this case, the simulation code launches thousands of threads, each of which handle computations for only a few particles. Since particles are independent of one another when computing the force or integrating the equations of motion, MD lends itself to this type of algorithm and speedups on the order of 2 orders of magnitude can be achieved. In our study of glassy dynamics in Chapter 6, we use the HOOMD-blue code to run fast MD simulations on the GPU.

4.2 Potential Energy Minimization

In many instances during our study, we require a method for creating potential-energy-minimized configurations, both for the purpose of making connections with the PEL (see section 2.4.2), and also for removing thermal noise from systems in the context of data analysis schemes. The basic goal of energy minimization is to quickly move the system to the nearest basin of attraction in the PEL (i.e., the nearest local potential energy minimum).

By definition, energy-minimization can be accomplished trivially by instantaneously setting the temperature in an MC or MD simulation to $\sim T = 0$ and performing simulation steps until the energy converges. (Note that for an MD simulation, T must remain slightly above zero; otherwise the particles will never move). While this method is straightforward, in practice, it is also computationally expensive. Thus, several standard methods have been developed to obtain fast convergence to local potential energy minima, including, for example, the conjugate gradient and steepest descent methods [150]. For our studies, we use an alternate method that utilizes an existing MD integrator for quickly converging to an energy minimum [151].

The equations of motion for the FIRE energy minimization scheme are given by [151]:

$$\mathbf{a}(t) = \mathbf{f}(t)/m - \gamma(t)|\mathbf{v}(t)|[\hat{\mathbf{v}}(t) - \hat{\mathbf{f}}(t)]. \tag{4.6}$$

Here, the hat denotes a unit vector. The second term has the effect of steering the velocity vector of the system in the direction of the force vector. The function $\gamma(t)$ is scaled up along with the timestep as long as $\mathbf{v}(t) \cdot \mathbf{f}(t) \geq 0$. However, if $\mathbf{v}(t) \cdot \mathbf{f}(t) < 0$, $\gamma(t)$ and the timestep are reset and the velocity is set to zero. Physically, this means that so long as the velocity of the system is in the direction of the force (i.e., the system is moving downhill), the system is integrated along this direction with larger and larger steps. If the velocity is in the opposite direction of the force (i.e., the system is moving uphill), the system is stopped by setting the velocities to zero. The system, now at rest, is pulled along a new direction, and the process starts again. The FIRE algorithm is presented in detail in reference [151].

4.3 Free Energy Calculations

In the thermodynamic limit, physical systems assume a minimum free energy state. Thus, knowing the minimum free energy state is important, since it tells us whether a given structure is thermodynamically stable or metastable. The free energy cannot be computed directly as an ensemble average as can the internal energy or kinetic energy. Thus, specialized techniques must be employed. In this section, we review some standard methods for computing free energies. Since the realm of free energy calculations is rather expansive, we narrow our focus to the methods that we use to obtain results for the work presented in Chapters 5.

4.3.1 Thermodynamic Integration and Free Energy Perturbation

The free energy difference between two systems A and B is given by the work required to change system A to system B [152]:

$$W = \Delta F \equiv F_A - F_B. \tag{4.7}$$

The change can be a change in pressure, volume, particle interactions, etc. Equation 4.7 assumes that the change is made infinitely slowly (i.e., it is a reversible thermodyanmic process). If the system is changed at a finite rate, work is dissipated during the irreversible process and the relation becomes [152]:

$$\overline{W} \geq \Delta F. \tag{4.8}$$

The overline denotes an ensemble average over all initial states in A. Jarsynski showed that the free energy can be expressed in terms of a path ensemble average [152, 153]:

$$\Delta F = -\frac{1}{\beta} \ln \overline{e^{-\beta W}}. \tag{4.9}$$

Here, the work for each realization of the path z from A to B is given by:

$$W = \int_0^{t_s} \frac{\delta H_\lambda}{\delta \lambda} [(\mathbf{z}(t)] \dot{\lambda} dt. \tag{4.10}$$

Here, $\dot{\lambda}$ is the switching rate with units t_s^{-1} and $\mathbf{z}(t)$ is a trajectory (i.e., a series of contiguous microstates). Therefore, the free energy measurement involves computing the average work for an ensemble of paths of length t_s where we modify the switching parameter λ. When λ is 0, the system is in state A, and when λ is 1, the system is in state B. Notice that these paths can be highly non-equilibrium.

Now, consider what happens in the limit of the switching time t_s [152]. For infinitely slow switching (i.e., a quasi-static process), the equation becomes:

$$\Delta F = \int_0^1 \left\langle \frac{\delta H_\lambda}{\delta \lambda} \right\rangle_\lambda d\lambda. \tag{4.11}$$

This is the famous Kirkwood equation [154], which is the basis for the thermodynamic integration (TI) method. The equation involves computing the free energy as an ensemble average of the derivative of the Hamiltonian with respect to the coupling parameter. In practice, this involves running several simulations with different values of coupling parameter λ and numerically integrating the expression in equation 4.11. Notice that this is a

more general form of the standard thermodynamic integration used by experimentalists to measure ΔF from the EOS variables [140].

In the limit of infinitely fast switching time, the equation becomes:

$$\Delta F = -\frac{1}{\beta} \ln \langle e^{-\beta \Delta H} \rangle_0. \tag{4.12}$$

In other words, the work is proportional to the difference in the Hamiltonian for state A and state B evaluated at the initial point $t = 0$ (i.e., when the system is in state A). This method is known as free energy perturbation (FEP) [155]. In practice, the states A and B must by sufficiently close that the difference (and hence the error) is small. As a result, FEP is typically carried out in several windows. Since the windows are independent, FEP is well suited for an "embarrassingly parallel" computational scheme, where several completely independent simulations are run on different processors.

When using any of the three methods above, we must parameterize the system as a function of the switching variable λ. That is, we must modify the Hamiltonian of the system such that for $\lambda = 0$, the system is in state A, and for $\lambda = 1$, the system is in state B. Several examples of this are given in Chapter 5 in the context of computing free energies for dimer-sphere mixtures.

4.3.2 Umbrella Sampling

As mentioned above in the context of the FEP method, we can obtain the free energy difference between two states 1 and 0 by computing ΔU based on configurations sampled in state 0:

$$\exp(-\beta \Delta F) = \langle \exp(-\beta \Delta U) \rangle_0 . \tag{4.13}$$

This equation is not applicable if 1 and 0 are too dissimilar, such that the average is dominated by a few very large values of ΔU. One solution to this problem is given by the umbrella sampling method of Torrie and Valleau [156], which involves rewriting equation 4.13 as:

$$\langle \exp(-\beta \Delta U) \rangle_0 = \frac{\langle \exp(-\beta U_1)/\pi \rangle_\pi}{\langle \exp(-\beta U_0)/\pi \rangle_\pi}. \tag{4.14}$$

Here, π is a weighting function that determines the probability at which points are sampled in phase space. In practice, the weight function is adjusted such that there is sufficient overlap between regions 1 and 0. Often, the free energy calculation is carried out in several stages or windows, such that there is always sufficient overlap between neighboring windows in phase space [140]. We apply a unique umbrella sampling scheme, originally formulated

by Frenkel and coworkers [57, 62], to measure the free energy of a growing quasicrystalline nucleus as a function of its size in Chapter 5.

4.4 Methods for Characterizing Rare Events

In many cases, transitions are dominated by "rare events." Rare events are activated processes, which require that the system overcome a free energy barrier to transform from an initial state A to a final state B. In the context of our studies, some examples of rare events include the formation of a critical nucleus in homogeneous nucleation and local structural transitions in glassy supercooled liquids. Other standard examples include chemical reactions and conformal changes in macromolecules [157, 158]. Rare events are fluctuation-driven; therefore the average waiting time between events may be orders of magnitude longer than the timescale of the event itself. Conventional brute force molecular simulation is therefore not an acceptable tool for studying rare events, since most of the computational effort is dedicated to simulating the long waiting time between events.

MD simulations to calculate the rates of rare events were first carried out by Bennett in the context of diffusion in solids [159]. Chandler generalized and extended this method to the calculation of reaction rates [160]. Today, the most commonly used scheme to simulate rare events is known as the Bennett-Chandler method, which involves the calculation of the free energy along a pre-defined reaction coordinate [140]. A reaction coordinate is a generalized one-dimensional coordinate that represents progress along a reaction pathway, i.e. a path in phase space from the initial (reactant) to final (product) state. Sampling is typically implemented via a biasing scheme such as constrained dynamics, multicanonical, or umbrella sampling methods that preferentially sample regions of phase space important to the transition [140]. Using the Bennett-Chandler method, it is possible to calculate the reaction rate provided that the transition state (the configuration(s) at the top of the free energy barrier) can be determined. The primary disadvantage to using the Bennett-Chandler method for certain problems is that it requires an a priori definition of the reaction coordinate. For complex systems, such as the nucleation of a quasicrystal from the liquid state, the reaction coordinate may be a complicated multidimensional surface, and a poor choice for a reaction coordinate can lead to an inefficient or incorrect rate-constant calculation.

Chandler and coworkers have since developed Transition Path Sampling (TPS) [60], a scheme for simulating rare events that does not require the definition of a reaction coordinate. TPS is actually a straightforward MMC algorithm; however, whereas a typical MC simulation involves the perturbation of initial particle positions to converge to an equilibrium

ensemble of configurations, TPS involves the perturbation of an initial transition pathway (i.e., a series of phase space configurations that connects A and B) to converge to an equilibrium ensemble of transition pathways (the Transition Path Ensemble (TPE)). A perturbation may be, e.g., a change in momenta for deterministic trajectories generated with molecular dynamics (MD), or a change in the random number seed for stochastic trajectories generated via MC. The equilibrium probability distribution of a path within the TPE is given by:

$$P[x(\tau)] = \frac{h_A(x_0)h_B(x_\tau)}{Z_{AB}(\tau)}\rho(x_0)\prod_{i=0}^{\tau/\Delta t-1} p(x_{i\Delta t} \rightarrow x_{(i+1).\Delta t}) \tag{4.15}$$

Here, τ is the path length, Δt is the discretization step, $\rho(x_0)$ is the probability density of an initial point on the path, Z_{AB} is a transition path partition function. The binary functions h_A and h_B ensure that the paths satisfy the constraint of beginning in state A and ending in state B. We apply a novel TPS scheme for studying rare structural transitions in glassy liquids in Chapter 6.

4.5 Characterizing Structure

As outlined in chapter 2, there is a fundamental connection between particle packing and thermodynamics. In many cases, particle structure is also related to physical, chemical and optical properties of the system as well. Thus, it is natural to gain insight into the stability and properties of condensed systems by characterizing the structural patterns that they form. In this section, we review some standard structural measurements used to characterize liquids and solids. In chapter 7, we introduce more advanced structural characterization techniques.

4.5.1 Radial Distribution Function

Perhaps the most standard measure of structure in condensed systems is the radial distribution function $g(r)$:

$$g(\mathbf{r}) = \frac{V}{N^2}\left\langle \sum_i^N \sum_{j\neq i}^N \delta(\mathbf{r}-\mathbf{r}_i+\mathbf{r}_j)\right\rangle. \tag{4.16}$$

Physically, $g(\mathbf{r})$ measures the probability of observing a particle at a radial distance $\mathbf{r}$ away from a particle at the origin relative to the analogous probability in an ideal gas of the same density. If particular particle spacings are more probable than others, $g(\mathbf{r})$ will exhibit

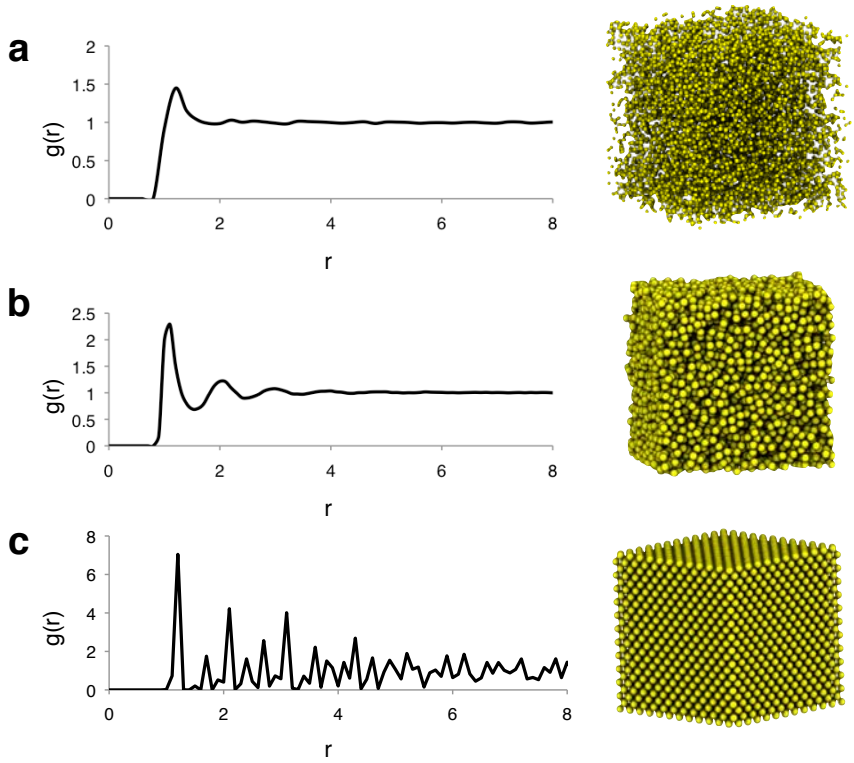


Figure 4.1 The radial distribution function (RDF). (a) RDF for a gas-like fluid. (b) RDF for a dense liquid. (c) RDF for a crystal.

maxima. Thus, $g(\mathbf{r})$ is a measure of translational ordering. For isotropic systems, such as liquids and gasses, all spatial directions are equivalent and we often measure an isotropic $g(r)$ with scalar r in the place of vector $\mathbf{r}$:

$$g(r) = \frac{V}{N^2} \left\langle \sum_i^N \sum_{j \neq i}^N \delta(r - |\mathbf{r}_i - \mathbf{r}_j|) \right\rangle . \tag{4.17}$$

Examples of the isotropic $g(r)$ function for a gas, a liquid, and a crystal formed in the monatomic LJ system are shown in Fig. 4.1a,b and c, respectively. Notice that, by definition, for an ideal gas, $g(r) = 1$. A pseudo code for computing $g(r)$ is given below:

```
//algorithm for computing g(r)
//1. compute probability of r:
for i in N
   for j in N
      set r = distance(x(i), x(j))
      //give weight of 2, since we are looping over pairs
      set rbin = numbins*floor(r/rmax)
      set hist(rbin) = radial_hist(rbin) + 1
   end for
end for
```

```
//2: normalize the rdf
set binsize = rmax / nbins
set delta = binsize / 2.0
if dim == 3
    set pre = 4.0/3.0*M_PI
endif
if dim == 2
    set pre = M_PI
endif
for i in numbins
    set r = (i+0.5)*binsize
    set vol_shell = pre*pow(r+delta, dim) - pre*pow(r-delta, dim);
    set np_ideal_gas = density*vol_shell;
    set normalization(i) = normalization(i) + np_ideal_gas;
    set probability(i) = probability(i) + hist(i)/N;
    gofr(i) = probability(i)/normalization(i);
end for
```

There are several subtleties that should be accounted for when computing $g(r)$. First, to compute an ensemble average over $g(r)$ it makes more sense to compute $\langle P(r)\rangle \,/\, \langle n(r)\rangle$ than $\langle g(r)\rangle$. Also, for systems with periodic boundary conditions, $g(r)$ should be truncated at a distance $L_{max} \leq L/2$, since otherwise self-correlations will exist. If the desired range of $g(r)$ is short relative to the box size, a cell-list can be used to optimize distance computations.

4.5.2 Static Structure Factor

Another structural quantity that is commonly measured for condensed systems is the structure factor $S(\mathbf{q})$. The structure factor can be considered a mathematical description of how a system scatters light or subatomic particles. The structure factor is given by the discrete Fourier transform (DFT) of particle positions $\mathbf{x}$:

$$S(\mathbf{q}) = \frac{1}{N}\sum_{j}^{N} \exp(-i\mathbf{q}\cdot\mathbf{r}_j). \tag{4.18}$$

The intensity of the signal obtained from scattering is given by the magnitude of $S(\mathbf{q})$: $I(\mathbf{q}) \propto [\Re(S\mathbf{q}) + \Im(S(\mathbf{q})]^{1/2}$. The full d-dimensional Fourier space representation of the system can be analyze directly. One way to simplify the problem is to reduce the dimensionality of the structure factor to 1d, $S(\mathbf{q}) \rightarrow S(q)$, by taking the length of the wavevector $q = |\mathbf{q}| = (q_i^2 + q_j^2 + q_k^2)^{1/2}$. In other words, we take the average intensity for a given shell of radius q in the full 3d diffraction pattern. This representation is particularly useful for isotropic structures, such as liquids, since the intensity only depends on $|\mathbf{q}|$ (see Fig. 4.3).

The isotropic structure factor is given by:

$$S(q) = \frac{1}{N}\sum_{j}^{N}\exp(-iqr_j). \tag{4.19}$$

Fig. 4.2 shows the isotropic structure factor for the KA binary LJ mixture [133] (see section 3.3. A pseudo code for computing the isotropic structure factor is given below:

```
//algorithm for computing the isotropic structure factor:
for Li=Limin; Li<=Limax; Li+=Listep
   for Lj=Ljmin; Lj<=Ljmax; Lj+=Ljstep
      for Lk=Lkmin; Lk<=Lkmax; Lk+=Lkstep
         set Scos = 0.0
         set Ssin = 0.0
         //change from box units to Fourier units (note L is the box period -> 2pi)
         set q = 2.0*pi*sqrt(Li*Li*invLx*invLx + Lj*Lj*invLy*invLy + Lk*Lk*invLz*invLz)
         for int i=0; i<NA; i++
            for j=0; j<NB; j++
               set r = distancevec(x(i), x(j))
               set temp = (Li*r(0)*invLx + Lj*r(1)*invLy + Lk*r(2)*invLz)*2.0*M_PI
               set Scos = Scos + cos(temp)
               set Ssin = Ssin + sin(temp)
            end for
         end for
         set S = (Scos*Scos + Ssin*Ssin)/sqrt(NA*NB)
         set qbin = floor(q/qmax)*nbins
         set Sq_total(qbin) = Sq_total(qbin) + S
         set Sq_normalization(qbin) = Sq_normalization(qbin) + 1
      end for
   end for
end for
return Sq_total/Sq_normalization
```

Computing the full 3d structure factor and then performing radial binning in Fourier space can be computationally expensive. An alternative, simpler way of computing the isotropic structure factor $S(q)$ is to compute the Fourier transform for a range of wave vectors q and then reweight the transform according to the relative probability of observing interparticle distance in the system corresponding to each wave vector. The latter quantity is precisely the information contained in $g(r)$. Specifically, $S(q)$ is related to $g(r)$ by [161]:

$$S(q) - 1 = \frac{4\pi\rho}{q}\int_0^\infty [g(r) - 1]r\sin(qr)dr \tag{4.20}$$

And conversely, $g(r)$ is related to $S(q)$ by:

$$g(r) - 1 = \frac{1}{2\pi^2\rho r}\int_0^\infty [S(q) - 1]q\sin(qr)dq \tag{4.21}$$

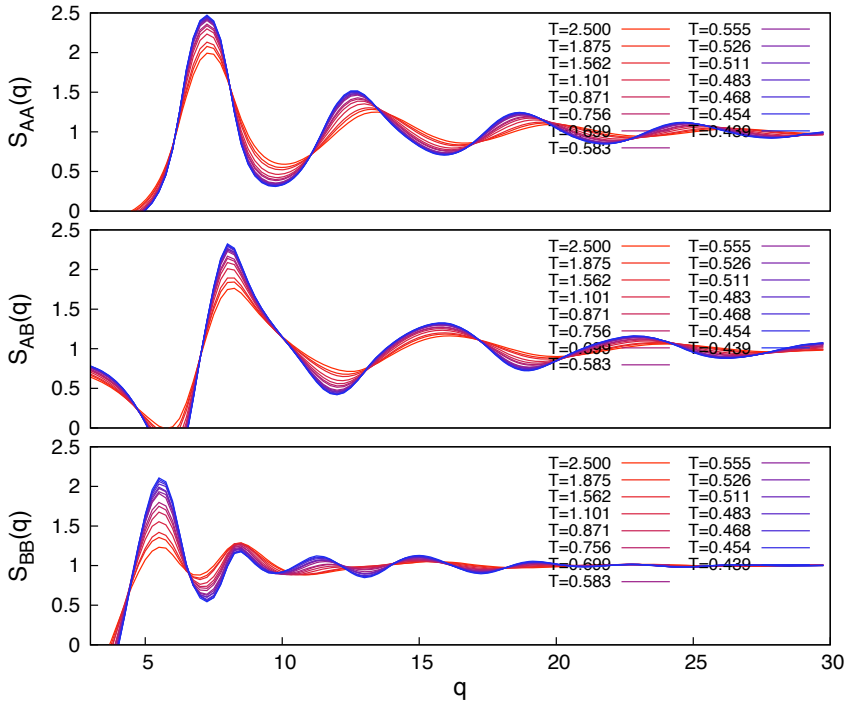


Figure 4.2 Isotropic structure factor for the KA binary LJ liquid [133] for a range of temperatures.

Thus, $S(q)$ and $g(r)$ are Fourier transform pairs. A pseudo-code for computing $S(q)$ from $g(r)$ given below:

```
//algorithm for computing the isotropic structure factor using g(r):
for q = qmin; q<=qmax; q += qstep
   set sum = 0.0;
   for i in gofrnumbins
      set r = (i+0.5)*dr
      set gofrm1 = gofr[i] - 1.0
      set sum += 4.0*M_PI*info.rho*grm1*r*r*sin(q*r)/(q*r)*dr
      set s(q) = 1.0 + sum
   end for
```

4.5.3 Diffraction Image

Another useful structural quantity to measure for crystal structures is the simulated diffraction pattern. The diffraction pattern arises in light scattering experiments where light waves (often X-rays for atomic systems) with wavelengths comparable to the interatomic spacings in a crystal are scattered and undergo constructive interference giving rise to maxima or known as Bragg peaks [162]. We can simulate the diffraction pattern by computing the 2d discrete Fourier transform (DFT) in a particular plane for system of particles. A simulated

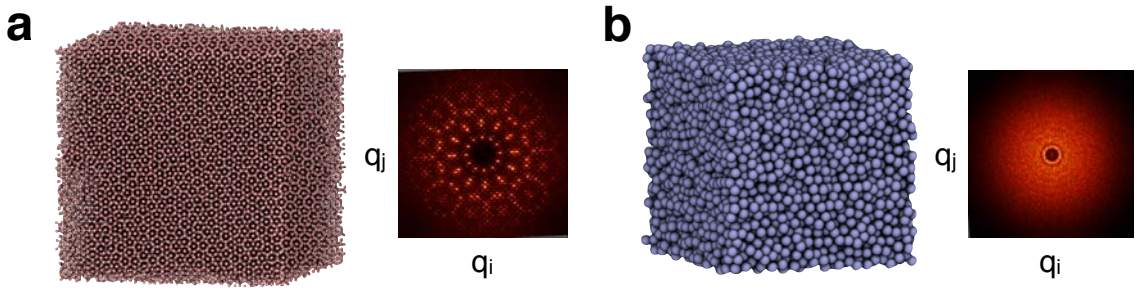


Figure 4.3 Simulated diffraction patterns. (a) A dodecagonal phase, formed in the Dzugutov system [127] with N=100,000 at $\rho = 0.85$. (b) A binary LJ liquid [163] with N=10,000, $\rho = 1.0$.

diffraction image for an ordered structure is shown in Fig. 4.3a. The pattern exhibits distinct maxima. A simulated diffraction image for a liquid is shown in Fig. 4.3b. Notice that the liquid does not exhibit Bragg peaks, since it has no long range ordering.

The diffraction pattern contains several useful pieces of information. First, the presence of Bragg peaks indicates that the structure is ordered. Second, the symmetry of the diffraction image reflects the rotational symmetry of the crystal structure. For example, the diffraction image for the quasicrystal-like structure in Fig. 4.3a has pseudo 12-fold symmetry, indicating the structure itself has pseudo 12-fold rotational symmetry. (Notice that the 12-fold symmetry is not perfect, as indicated by slight differences in the diffraction pattern upon applying a 12-fold rotation, and thus this particular structure is not a true quasicrystal). The diffraction image is trivial to calculate and a pseudocode is given below:

```
//algorithm for computing the simulated diffraction image:
set nq = 2*qmax/qstep+1
for i=0; i<N; i++
   for ii=0; ii<nq; ii++
      set qi = (ii-nq/2)*qstep
      for jj=0; jj<nq; jj++
         set qj = (jj-nq/2)*qstep
         set complex_image[ii][jj] = set complex_image[ii][jj] + ...
              complex(cos(qi*x[i][0]+qj*x[i][1]), -sin(qi*x[i][0]+qj*x[i][1]))
      end for
   end for
end for
//Now plot the pixels; abs(complex_image[i][j]) gives the intensity of pixel[i][j]
```

The diffraction image is plotted by taking the absolute value of each complex component as the pixel intensity. The code can be optimized by computing the fast Fourier transform (FFT) rather than the DFT [164].

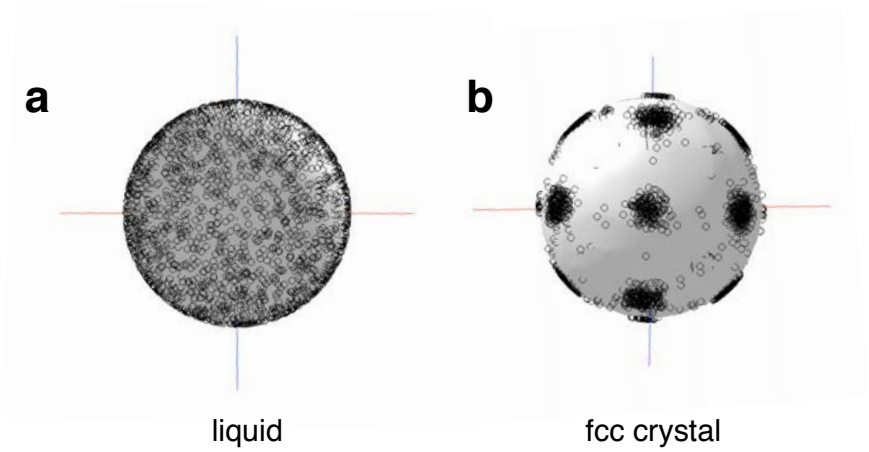


Figure 4.4 Bond Order Diagrams (BODs) [163, 165]. (a) BOD for a liquid system. (b) BOD for an fcc crystal. Notice that while the BOD for the liquid shows a uniform coverage, the BOD for the crystal shows distinct maxima.

4.5.4 Bond Order Diagram

A lesser known, but very useful quantity for studying ordered structures is the bond order diagram (BOD) [52]. The BOD is constructed by projecting the $[\theta, \phi]$ neighbor directions of each particle onto the surface of a sphere. Physically, the BOD represents the probability of observing a particle in a direction $[\theta, \phi]$, given that a particle sits at the origin. The BOD is depicted for a liquid and an face-centered cubic (fcc) crystal in Fig. 4.4a and Fig. 4.4b, respectively.

The patterns detected by the BOD have been characterized mathematically in the context of bond-order parameters (BOPs) [48, 132]. BOPs can give a quantitative measure of the degree of ordering in the system by detecting the degree of symmetry in the bond order diagram. Additionally, they can be used in some cases to identify structures based on their BOD. In Chapter 7, we describe BOPs in mathematical detail, and extend their applicability to more complex structures, such as those encountered in self-assembly studies. A pseudo code for computing the BOD is given below:

```
//algorithm for computing the bond order diagram:
for i in N
   for j in N
      if i != j
         set dx = distancevec(x(i), x(j))
         set r = sqrt(dx(0)*dx(0) + dx(1)*dx(1) + dx(2)*dx(2))
         if (r < cut)
            set dx = dx / r
            set bod(index) = dx
            index = index + 1
         end if
```

```
      end if
   end for
end for
//now plot the BOD coordinates in your favorite visualizer
```

4.6 Characterizing Dynamics

In many cases, we are interested in characterizing the dynamics of particles as well as their structure (i.e., how particles relax structurally as a function of time). This is particularly important when studying supercooled liquids, which are known to exhibit anomalous dynamical properties [63]. In this section, we review some standard functions used to measure particle dynamics, as well as some highly specialized dynamical measurements contrived specifically for studying dynamical heterogeneity (DH) in supercooled liquids.

4.6.1 General Considerations for Dynamical Measurements

When measuring dynamical properties, we are typically interested in how a system changes over different values of a time window t. One subtle consideration is that many measurements are best observed over logarithmically spaced time intervals. Therefore, in the interest of minimizing the computer memory required for an analysis routine, it is optimal to save data at logarithmic intervals . A generic psuedo-code for computing a dynamical measurement with logarithmic time intervals is given below:

```
//algorithm for computing a dynamical quantity with log-spaced time windows:
for p=0; p<=logmaxtime; p+=logstep
   set timewindows(index) = (int)pow(10.0, p)
   set index = index+1
end for

for t in timewindows
   call dynamicalmeasurement(x(0), x(t))
end for
```

In practice, we should obtain several measurements at each time interval t for averaging. Thus, we should choose several different statistically-independent reference times $t = 0$. Since most dynamical measurements decorrelate over a timescale of about 1 α relaxation time (see section 4.6.4 below), we should first estimate the α relaxation time, and then ensure that the reference times $t = 0$ are separated by at least this timescale. The α relaxation time can be roughly approximated by the time window t for which the mean-squared displacement (defined in the following section) reaches about 1 particle diameter.

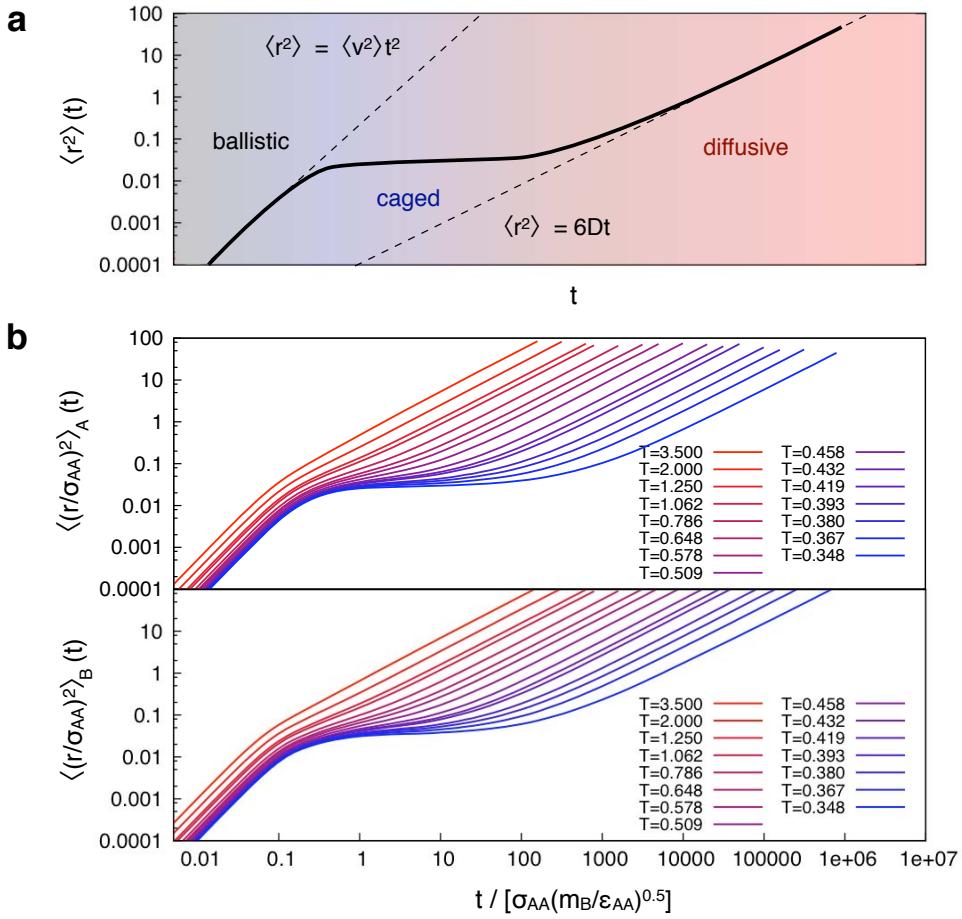


Figure 4.5 Mean-squared displacement (MSD). (a) Depiction of the MSD for a glassy supercooled liquid. The MSD exhibits three distinct regimes as a function of time, indicated by the different colors. (b) MSD for A and B particles for the WCA version of the Wahnström binary mixture [134] (see section 3.3). Note that for this model, $\sigma_{AA} = \varepsilon_{AA} = m_B = 1$.

4.6.2 Mean Squared Displacement

The mean squared displacement (MSD) is perhaps the most elementary measure of particle dynamics. The MSD measures how far particles move, on average, as a function of the time window t. Mathematically, the MSD is defined as :

$$\left\langle r(t)^2 \right\rangle = \left\langle |\mathbf{r}(t) - \mathbf{r}(0)|^2 \right\rangle . \tag{4.22}$$

For supercooled liquids, the mean squared displacement exhibits three distinct regimes (Fig. 4.7)a. Over very short timescales particles move without colliding with other particles and he mean squared displacement is given trivially by $\left\langle r^2(t) \right\rangle = \left\langle v^2 \right\rangle t^2$. This short timescale region of the MSD is known as the "ballistic regime." Over slightly longer times,

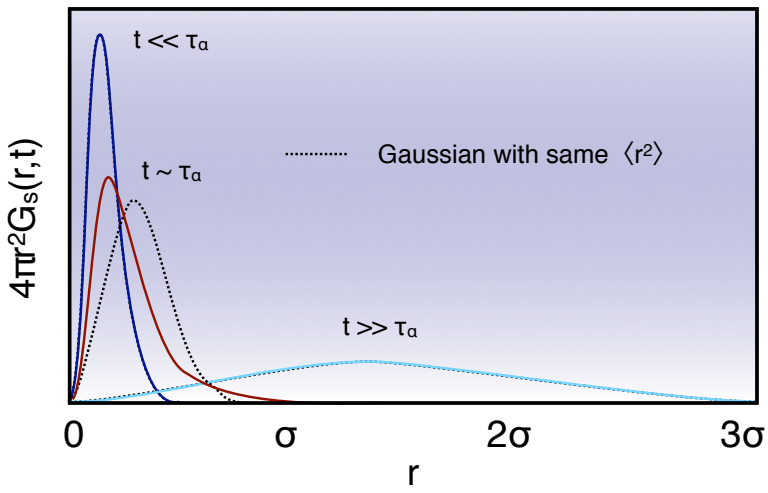


Figure 4.6 Self contribution to the Van Hove correlation function. The panel depicts $G_s(r,t)$ at three different times for a typical supercooled liquid. The times $t << \tau_\alpha$, $t >> \tau_\alpha$, and $t \sim \tau_\alpha$ corresponds to the ballistic regime, sub-diffusive regime, and the early crossover regime between caged and diffusive behavior, respectively.

particles collide with their nearest neighbors. For deeply supercooled liquids, glasses or solids, diffusion is rare or impossible and the MSD exhibits a caged regime where the MSD remains constant. Fig. 4.7b shows the MSD for the binary KA supercooled liquid as a function of temperature. As T decreases, the length of the caged regime increases dramatically. For liquids, particles eventually break out of their cages and the MSD enters a "diffusive regime." In this regime, the MSD is described by the Einstein relation $\left\langle r^2(t) \right\rangle = 2dDt$, where D is the diffusion coefficient and d is the dimensionality of space. The algorithm for measuring the mean squared displacement over a time window t is given below. Notice that this pseudo code can be used as a function within the pseudocode given in the previous section

```
//algorithm for computing the mean squared displacement:
rsq = 0.0;
for i in N
   r = distance(xt(i), x0(i))
   rsq += r*r
end for
return rsq / N;
```

4.6.3 Van Hove Correlation Function

The MSD is an average value of the displacement r, and thus it gives no information about the distribution of r that contributes to this average. In some cases, it is useful to analyze this

probability distribution (i.e., the distribution of r values that gives rise to each datapoint in the MSD). This is given by the self-contribution to a more general density-density correlation function known as the Van Hove correlation function $G(\mathbf{r},t)$:

$$G(\mathbf{r},t) = \frac{1}{\rho}\left\langle \rho(\mathbf{r},t)\rho(\mathbf{0},0)\right\rangle . \tag{4.23}$$

The bracketed term is proportional to the probability of finding a particle at a distance $\mathbf{r}$ from the origin $\mathbf{0}$ at time t, given that a particle sits at the origin at time 0. Notice that this is similar to a time-dependent generalization of the radial distribution function $g(r)$. Since we are interested in measuring the Van Hove correlation function for liquids, the system is isotropic and thus we can replace $\mathbf{r}$ with r. The "self" contribution to the Van Hove correlation function $G_s(r,t)$ is given by exclusively considering correlations between a given particle i at two times, rather hat than pairs of distinct particles i and j:

$$G_s(r,t) = \frac{1}{N}\left\langle \sum_{i=1}^{N} \delta[r - |\mathbf{r}_i(t) - \mathbf{r}_i(0)]\right\rangle . \tag{4.24}$$

In other words, $G_s(r,t)$ is the probability of a particle i moving a distance r over a time window t. In most cases, $G_s(r,t)$ is a Gaussian distribution, given by:

$$G_0(r,t) = \left(\frac{3}{2\pi\left\langle r^2(t)\right\rangle}\right)^{3/2} \exp\left(-\frac{3r^2}{2\left\langle r^2(t)\right\rangle}\right). \tag{4.25}$$

However, if the system exhibits spatially heterogenous dynamics (SHD), then $G_s(r,t)$ will deviate from Gaussian behavior [166]. This is exemplified by Fig. 4.6, which depicts $G_s(r,t)$ for a supercooled liquid over different time windows. The window denoted $t << \tau_\alpha$ corresponds to the ballistic regime of the MSD, and the time $t >> \tau_\alpha$ corresponds to the sub-diffusive regime, where the system has begun to crossover to diffusive behavior. The time $t \sim \tau_\alpha$ is near the cage breakout time of the MSD, where the particles begin to first leave their cages. (A more rigorous definition of τ_α is given in the following section. Notice that for $t >> \tau_\alpha$ and $t << \tau_\alpha$, the self-Van Hove exhibits Gaussian behavior. However, near $t \sim \tau_\alpha$, the dynamics are highly non-Gaussian, indicating the presence of SHD. A pseudo code for computing the self Van Hove is given below:

```
//algorithm for computing the self van hove:
nbins = rmax / rstep
for i in N
  r = distance(xt(i), x0(i))
  bin = (int)(r/rmax)*nbins
  vanhoveself[bin]  = vanhoveself[bin] + 1;
end for
```

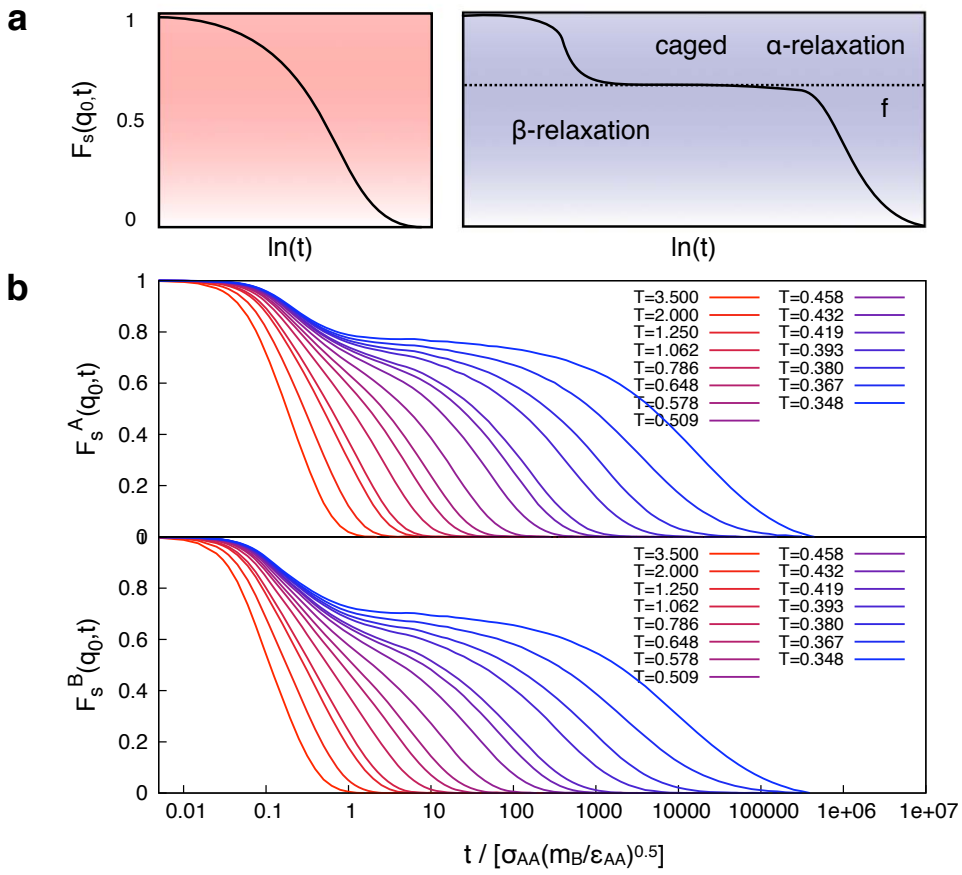


Figure 4.7 Self intermediate scattering function, $F_s(q_0,t)$. (a) Depiction of the $F_s(q_0,t)$ for a glassy supercooled liquid for high (left) and low (right) temperatures. For low temperatures $F_s(q_0,t)$ exhibits an initial β relaxation, followed by caged behavior, and then a primary α relaxation. (b) $F_s(q_0,t)$ for A and B particles for the WCA version of the Wahnström binary mixture [134] (see section 3.3. The scattering vectors $q_0^A = 7.5$ and $q_0^B = 8.0$ are determined by the peak of the structure factors $S_{AA}(q)$ and $S_{BB}(q)$. The MSD is shown for the same system above. Note that for this model, $\sigma_{AA} = \varepsilon_{AA} = m_B = 1$.

4.6.4 Intermediate Scattering Function

In many experimental systems, obtaining the particle positions directly is not feasible. However, as mentioned in section 4.5.2, the structure factor $S(q)$ can often be measured. Thus, in experimental systems we typically consider a Fourier space analogue of the Van Hove density-density correlation function known as the intermediate (i.e., time-dependent) scattering function $F(\mathbf{q},t)$, defined by:

$$F(\mathbf{q},t) = \frac{1}{N}\langle \rho(\mathbf{q},t)\rho(-\mathbf{q},0)\rangle\,. \tag{4.26}$$

Here, q is the wave vector, proportional to the wavelength of the incident light in a scattering experiment, and $\rho(\mathbf{q})$ is the Fourier transform of the density $\rho(\mathbf{r})$. Again, we are typically interested in the self-contribution to the correlation function, and we can write $\mathbf{q}$ as q, since we are interested in quantifying dynamics for isotropic systems. In practice, we typically measure $F_s(\mathbf{q},t)$ for a particular wavelength q_0, which is usually chosen to be the first peak in the structure factor $S(q)$. The self intermediate scattering function is given by:

$$F_s(q,t) = \frac{1}{N}\left\langle \sum_j^N \exp\left[-iq\left|\mathbf{r}_j(t) - \mathbf{r}_j(0)\right|\right]\right\rangle. \tag{4.27}$$

Here, we have inserted the definition of the structure factor (section 4.5.2) for $\rho(q)$. As depicted in Fig. 4.7a, for a normal liquid, $F_s(q,t)$ exhibits exponential decay: $F_s(q,t) = \exp(-t/\tau)$. For a deeply supercooled liquid, $F_s(q,t)$, like the MSD, exhibits three regimes: a regime of initial short lengthscale motion, a plateau regime in which particles are "caged" by their neighbors, and a regime of diffusive relaxation. The initial β relaxation is due to the vibrational motion of particles within their cages. The decay to the plateau regime can be described by $F_s(q,t) = f + At^{-}a$ [93]. The decay in the α relaxation regime can be described by the Kohlrausch-Williams-Watts (KWW) stretched exponential function [133]:

$$F_s(q,t) \propto \exp(-t/\tau)^{\beta}. \tag{4.28}$$

The variable τ in this equation is the α-relaxation time that is commonly to characterize the relaxation of glassy systems. The exponent β is bounded by $0 < \beta leq 1$. Notice that, perhaps confusingly, the exponent β applies to the α-relaxation regime. An algorithm for computing $F_s(q,t)$ is given in pseudo-code below:

```
//algorithm for computing the self intermediate scattering function:
//note: only use real scatterers (cos)
set sum = 0.0
for i in N
   set dr = distancevec(xt(i), x0(i))
   for k in dim
      set sum = sum + cos(q0*dr(k));
   end for
end for
set sum = sum/n;
set sum = sum/dim;
return sum;
```

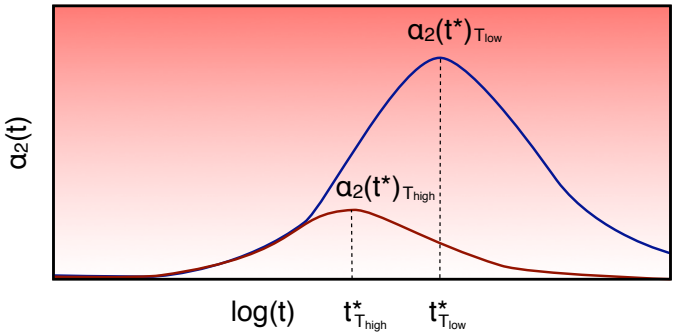


Figure 4.8 Non-Gaussian parameter. The panel shows a depiction of the non-Gaussian parameter as a function of the time window $\alpha_2(t)$. Particle motion deviates the most from Gaussian behavior at $t^*_{\alpha_2}$.

4.6.5 Heterogeneous Dynamics

As mentioned in section 4.6.5, supercooled liquids exhibit SHD, which becomes increasingly pronounced upon cooling towards a glass transition [67]. SHD is a universal characteristic of glassy materials, and has been observed in atomistic [167, 76], molecular [72, 67], colloidal [64, 168], and granular systems [31, 65, 66], both simulated and experimental. SHD itself exhibits universal characteristics, observable across diverse systems, such as string-like motion among mobile particles that grow as the glass transition is approached [73, 70]. Several measurements have been contrived to characterized SHD, and are the subject of our review in this section.

4.6.6 Non-Gaussian Parameter

As mentioned in section 4.6.3, the self Van Hove correlation function $G_s(r,t)$ is a simple measurement that can detect SHD. Specifically, $G_s(r,t)$ deviates from Gaussian behavior for a range of relaxation times, thus indicating the presence of SHD. Although viewing the probability histograms in Fig. 4.6 gives an intuitive sense of the degree of deviation from Gaussian behavior, a more direct measure is given by the non-Gaussian parameter α_2, defined as [169]:

$$\alpha_2(t) = \frac{3\left\langle |\mathbf{r}_i(t) - \mathbf{r}_i(0)|^4 \right\rangle}{5\left\langle |\mathbf{r}_i(t) - \mathbf{r}_i(0)|^2 \right\rangle^2} - 1. \tag{4.29}$$

Notice that the term in the denominator is the square of the standard MSD. As depicted in Fig. 4.8, the non-Guassuan parameter $\alpha_2(t)$ goes through a maximum $\alpha_2(t^*)$ at a time $t^*_{\alpha_2}$, which corresponds to the late-β / early-α relaxation regime of $F_s(q,t)$ (i.e., the time at

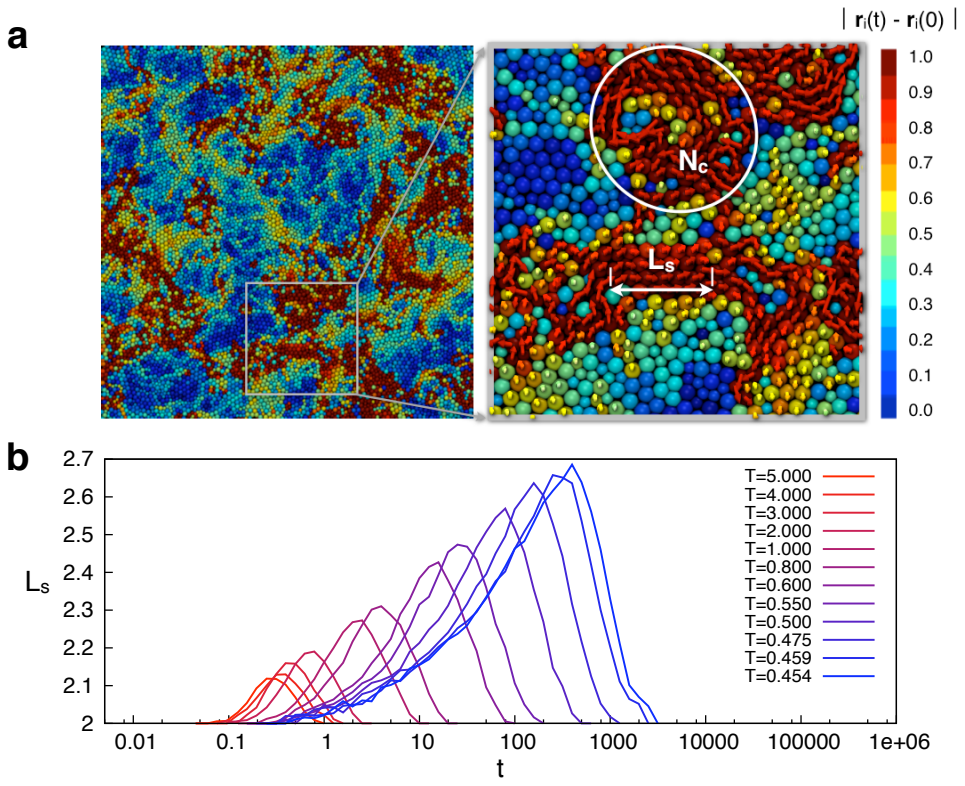


Figure 4.9 Mobile clusters and strings. (a) Depiction of particle motion showing mobile clusters (red) and strings (red, arrows). (b) Mean string length $L_s(t)$ for the KA system at $\rho = 1.2$ for a range of temperatures. For each T, $L_s(t)$ peaks on a time window t^* that is similar to $t^*_{\alpha_2}$.

which the α decay begins to occur). Both $\alpha_2(t^*)$ and $t^*_{\alpha_2}$ grow with decreasing temperature.

4.6.7 Mobile Clusters and Strings

The growth of the non-Gaussian parameter with decreasing T results from an increased degree of dynamical heterogeneity. This is depicted in Fig. 2.8; here, the regions of mobile and immobile particles become larger as temperature decreases. When viewing the dynamics in more detail (Fig. 4.9a), several dynamical quantities become apparent. One direct measurement of the degree of SHD is given by size of the regions of high mobility (red). This can be quantified by measuring the mean number of particles N_c within a "mobile cluster [74, 170, 71, 72, 64]." This requires that we make a distinction regarding which particles are highly mobile. To do so, we take the fN particles with the highest displacements in the system [74], where f is a fraction $f \in [0,1]$. The value of f is chosen to maximize the ratio $N_c(t^*)/N_c(t^m)$, where $N_c(t^*)$ is the maximum value of $N_c(t)$ and $N_c(t^m)$ is the minimum,

which occurs on timescales over which the system does not exhibit SHD. Typical values of f are $0.05 \leq f \leq 0.1$ [74, 170, 71]. Mobile particles i and j belong to the same clusters if they are within a cutoff range r_{nbr}, which is typically taken to encompass the first peak of $g(r)$.

Particle motion within the mobile clusters, shown by the displacement vectors in Fig. 4.9a, tends to be "string-like," where particles follow one another in quasi-1d paths [73, 71, 70]. Thus another intuitive measure of SHD is the mean string length, $L_s(t)$, defined by clusters of particles that, within a time interval t, replace the initial position of a neighboring particle. In practice, quantifying replacement requires that we introduce a threshold value for replacement r_{cut} such that $\left|\mathbf{r}_i(t) - \mathbf{r}_j(0)\right| < r_{cut}$. For spherical particles, we typically take r_{cut} in the range $0.3\sigma \leq r_{cut} \leq 0.5\sigma$, were σ is the particle diameter [73]. Mobile particles i and j belong to the same string if i replaces j over a time window t and i and j are neighbors at $t = 0$ (i.e., $\left|\mathbf{r}_i(0) - \mathbf{r}_j(0)\right| < r_{nbr}$). Since particles can only replace one another by moving in a pseudo 1d line, the string length goes as the number of particles in the string: $N_s \propto L_s$.

As shown in Fig. 4.9, the mean string length L_s grows as a function of decreasing temperature. The timescale for the maximum string length is identical to $t^*_{\alpha_2}$, the peak time of the non-Gaussian parameter [73, 71]. Although not shown here, the mean cluster size N_c shows an identical trend [71]. It has also been shown that at the peak time $t^*_{\alpha_2}$, the string length distribution is exponential while the cluster size distribution is a power-law [73, 71, 31]. The relationship between the mean string and cluster size, and the string and cluster size distributions are shown for our study of glassy dynamics in a granular system in Chapter 6.

4.6.8 Space-Time Susceptibilities

The growth of strings and mobile clusters with decreasing T described in the previous section is indicative of a growing lengthscale that is purely dynamical in nature (i.e., the domains are constructed solely on the basis of particle mobility). Similar growing lengthscales are observed for 'static' systems where the clusters are identified by an order parameter, for example, regions of space with different densities or different particle ordering [14]. Such growing static lengthscales are often associated with an underlying phase transition; for example lengthscales tend to diverge near a critical point when the distinction vanishes between the different phases [14]. Growing lengthscales are also associated with first order transitions, where, near equilibrium, large spatial fluctuations in the order parameter occur [14].

It has been proposed that the growing dynamical lengthscale associated with the glass

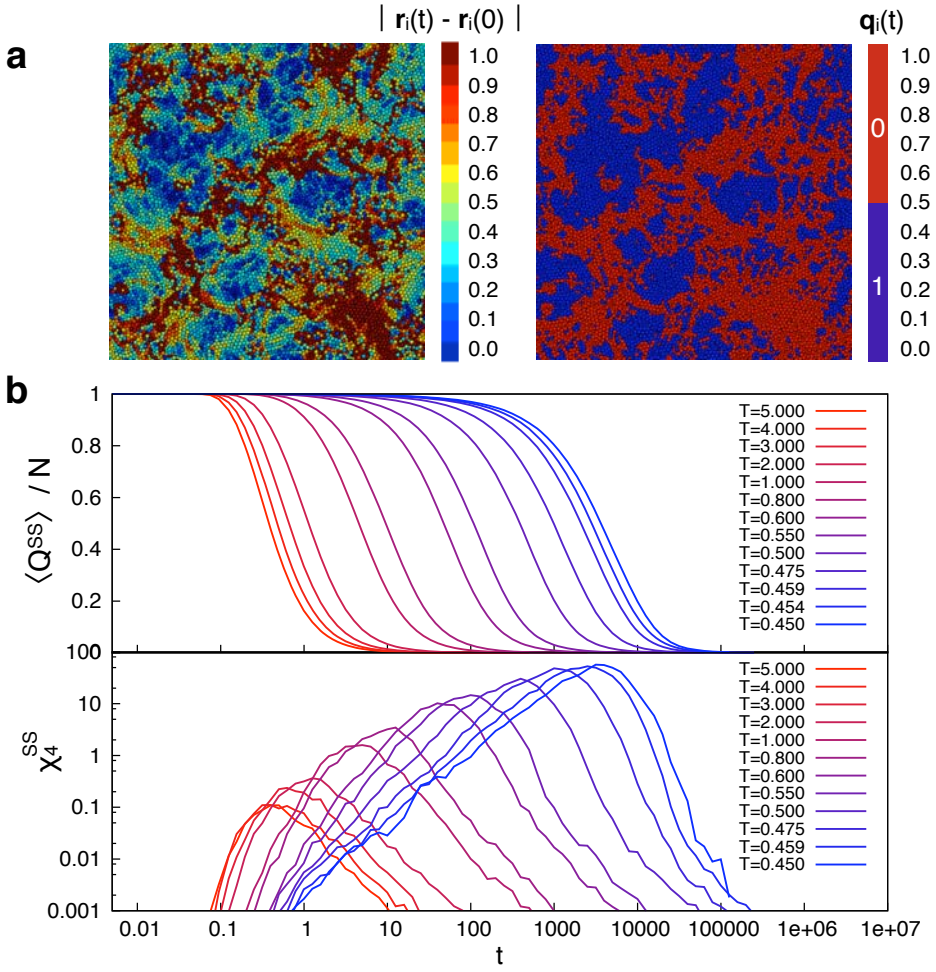


Figure 4.10 Space-time susceptibilities. (a) Depiction of particle dynamics for a 2d system. On the left, particles are colored continuously based on their mobility. On the right, particles are given a binary value of mobility based on whether or not they overlap with their initial position. (b) The mean value of the system self-overlap parameter $Q^{SS}(t)$, and the four point susceptibility $\chi_4^{SS}(t)$ as a function of temperature for the KA system at $\rho = 1.2$.

transition may also be associated within a unique underlying thermodynamic phase transition [171, 94, 109]. Since such a dynamical transition involves heterogeneous regions that are defined by mobility (i.e., the can only be measured by considering both space and time), the transition is said to occur in "space-time [108, 172, 109]" As proposed by Benneman *et al.* [167] and Donati, Glotzer, and Poole [173], many standard thermodynamic measurements can be generalized to fit within a space-time thermodynamic framework. That is, by idealizing the dynamical objects that arise in glassy systems as their analogous static quantities, we can treat them in the same thermodynamic framework that has already been developed for static systems.

One hallmark feature of a standard phase transition is a marked increase in a linear response function, or susceptibility near the transition [14]. For example, returning to the Ising model of section 2.1.1, we can write a magnetic susceptibility as:

$$\chi = \frac{1}{N}\left(\frac{\delta \langle M \rangle}{\delta \beta B}\right)_{\beta}. \tag{4.30}$$

Here, M is the magnetization and B is the magnetic field. Thus, the susceptibility physically represents the degree to which the average magnetization of the system changes upon applying a small change to an external field. One way that a susceptibility can be measured is in terms of spatial fluctuations in a local order parameter. In the case of the Ising model, the order parameter is given by the magnetization M of a subsystem of size N:

$$\chi = \frac{1}{N}\left\langle (\delta M)^2 \right\rangle = \frac{1}{N}\left(\left\langle M^2 \right\rangle - \langle M \rangle^2\right) \tag{4.31}$$

The susceptibility χ diverges at either a first order phase transition or a critical point [14]. In the case of a first-order transition, large spatial fluctuations in the order parameter will occur when the two phases are near equilibrium. Near a second-order critical point, the distinction between the phases disappears and the order parameter becomes correlated over very long ranges, resulting in very large spatial fluctuations.

Analogous susceptibilities can be defined to characterize the thermodynamics of space-time. An order parameter for a glassy liquid with SHD is given by the local values of the particle mobility (Fig. 4.10a). A susceptibility based on the total mobility $U(t) = \sum_i |\mathbf{r}_i(t) - \mathbf{r}_i(0)|$ is given by [167, 173]:

$$\chi_U(t) = \frac{\beta V}{N^2}\left[\left\langle U(t)^2 \right\rangle - \langle U(t) \rangle^2\right]. \tag{4.32}$$

A similar susceptibility can be defined by taking a binary value of the mobility, based on

whether particles escape their cage over a window t:

$$q_i(t) = \begin{cases} 1 & |\mathbf{r}_i(t) - \mathbf{r}_i(0)| < r_{overlap} \\ 0 & |\mathbf{r}_i(t) - \mathbf{r}_i(0)| \geq r_{overlap}. \end{cases} \tag{4.33}$$

The function $q_i(t)$ is a self-overlap order parameter. The susceptibility is then given by [76, 69]:

$$\chi_4^{SS}(t) = \frac{\beta V}{N^2} \left[\left\langle Q(t)^2 \right\rangle - \left\langle Q(t) \right\rangle^2 \right]. \tag{4.34}$$

The notation "SS" indicates that we consider self-overlaps in measuring Q_i, rather than self-distinct (SD) overlaps between particles i and j. Typically, we take a value of $r_{overlap}$ in the range $0.3\sigma \leq r_{overlap} \leq 0.6\sigma$, where σ is a particle diameter [69]. This captures correlations in particles breaking out of their cages during the α relaxation regime [76, 69]. However, different behavior can be captured by choosing different values for $r_{overlap}$ [174]. As depicted in Fig. 4.10b, χ_4^{SS} grows as the glass transition is approached. A similar trend is observed for $\chi_U(t)$ [167]. It has been speculated that this growing susceptibility may indicate proximity to a dynamical critical point [171, 94]. In contrast, a recent study indicates that the lengthscales associated with SHD may arise due to proximity to a first order transition in space-time [109].

Chapter 5
Forming Ordered Solids

As mentioned in the opening chapters, a sizable proportion of the solid materials that we interact with everyday are thermodynamically metastable. Thus, a comprehensive study of metastable ordered solids is well beyond the scope of the present study. Rather, we narrow our focus to a few more manageable case studies, designed to highlight some important aspects of metastable ordered solid formation. To do so, we focus specifically on the formation of quasicrystals, which exhibit desirable optical properties on the micro and nanoscales, and often form as metastable states. These attributes make quasicrystals ideal prototype structures for our studies.

In this chapter, we present work from two different articles on the topic of quasicrystal formation [30, 35]. In the first article, [A.S. Keys and S.C. Glotzer, "How Do Quasicrystals Grow?" *Physical Review Letters* (2007)] [30], we explore the unique growth mechanism that allows some systems to form quasicrystals rather than periodic approximants, which are often more thermodynamically stable. We find that, in contrast to the usual case of metastable solid formation, which is dictated by nucleation kinetics (see section 2.3.3), quasicrystal formation is determined by growth kinetics. An insightful perspective regarding the implications of this article is given in reference [175]. In the second article presented in this chapter [C.R. Iacovella, A.S. Keys, and S.C. Glotzer, "Self-Assembly of Soft Matter Quasicrystals," (2010)] [35] (currently pending publication), we address the problem of forming quasicrystals via self-assembly in diverse micro and nanoscale systems. To do so, we apply our knowledge of the thermodynamics of metastable solids, combined with our insights into quasicrystal growth obtained from the previous study, to design quasicrystal-forming systems. We demonstrate that quasicrystals can be consistently assembled from systems of micelles that exhibit shape polydispersity and have smaller mobile particles attached to their surface. Collectively, our results highlight some of the important and unique thermodynamic properties of quasicrystals, and present several exciting new possibilities for forming quasicrystals in experimental systems.

5.1 How Do Quasicrystals Grow?

Using molecular simulations, we show that the aperiodic growth of quasicrystals is controlled by the ability of the growing quasicrystal 'nucleus' to incorporate kinetically trapped atoms into the solid phase with minimal rearrangement. In the system under investigation, which forms a dodecagonal quasicrystal, we show that this process occurs through the assimilation of stable icosahedral clusters by the growing quasicrystal. Our results demonstrate how local atomic interactions give rise to the long-range aperiodicity of quasicrystals.

5.1.1 Introduction

Quasicrystals [122] are a unique class of ordered solids that display long-range aperiodicity, which distinguishes them from ordinary crystals. It is not known what 'special' qualities systems must possess in order to form quasicrystals versus crystals. Quasicrystals, like crystals, form via nucleation and growth [122], where a microscopic 'nucleus' of the solid phase spontaneously arises in the supercooled liquid and spreads outward, converting the system from liquid to solid [176]. A fundamental puzzle in quasicrystal physics is to understand how the growth phase of nucleation and growth can lead to a structure with long-range aperiodicity. Quasicrystals cannot grow like crystals, where the nucleus surface acts as a template for copying a unit cell via local interactions. Rather, quasicrystals, require specialized "growth rules" that dictate their formation [177].

Quasicrystal (QC) growth rules fall into two categories: energy-driven quasiperiodic tiling models [178, 179] and entropy-driven random tiling models [180, 181]. While energy-driven models rely on "matching rules" to dictate how atomic clusters or tiles attach to the nucleus, entropic models allow tiles to attach randomly to the nucleus with some probability. Although these models provide important insight into how QCs might form, the physical driving force underlying QC growth, and whether it is based on local interactions or long-range correlations, is not well understood.

In this section, we elucidate the physical mechanism underlying QC growth by studying the post-critical irreversible growth of a metastable dodecagonal QC from a simulated supercooled liquid. We show that QC growth is facilitated by structurally persistent atoms in low energy motifs that become kinetically trapped in their local configurations in the region surrounding the solid nucleus. As the nucleus grows, it incorporates these atoms in a way that minimizes expensive rearrangements and hastens solidification, allowing the QC to form instead of the stable crystalline approximant phase. In the system under investigation, we find that structurally persistent atoms were in icosahedral clusters prior to attaching to

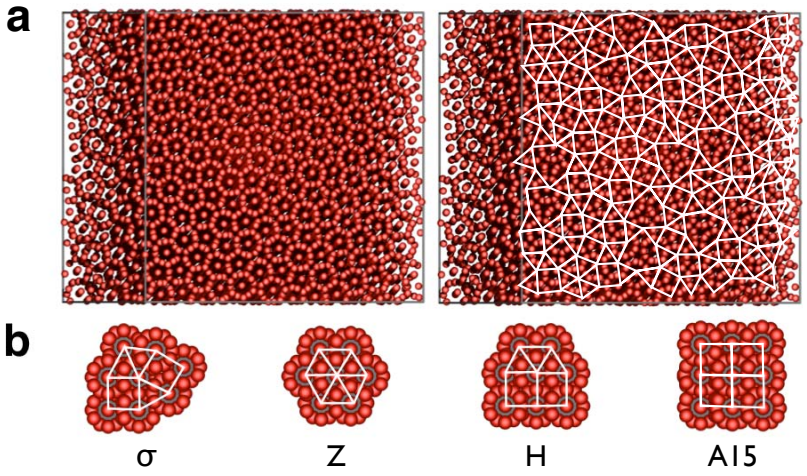


Figure 5.1 Dodecagonal QC and approximants. (a) 17,576 atom dodecagonal QC formed by the Dzugutov system using molecular dynamics at T=0.42 and ρ=0.85, instantaneously quenched to T=0. The image on the right shows the aperiodic tiles formed by connecting the centers of the dodecagonal rings of atoms. (b) Unit cells of various QC approximants.

the nucleus. Our results demonstrate how the long-range aperiodicity of QCs arises from local atomic interactions, thus providing a significant step forward in understanding the origin of the QC state.

5.1.2 Methods

To obtain these results, we perform three distinct sets of computer simulations. First, we use canonical (NVT) Monte Carlo (MC) to observe the growth of the QC from a static seed nucleus. We then use isothermal-isobaric (NPT) MC to observe the growth of large QC nuclei, generated via umbrella sampling [156]. Finally, we use umbrella sampling to generate many configurations containing nuclei to study the relationship between QC nuclei and icosahedral clusters. All simulations contain 3375 atoms with pair interactions modeled via the Dzugutov potential [127]. The form of the Dzugutov potential is identical to the 12-6 Lennard-Jones potential up to a distance at which an additional repulsive term dominates, suppressing the formation of BCC, FCC, and HCP crystals and favoring polytetrahedral ordering, where the 13-atom icosahedron is the ideal local structure. In the Dzugutov supercooled liquid, atoms are known to organize into local energy-minimizing icosahedral clusters comprised of face-sharing and interpenetrating icosahedra [182, 128, 70], which exhibit lower mobility than the bulk [128, 70]. The number of atoms that participate in

icosahedral clusters at any time increases with the degree of supercooling [129, 70]. At certain state points, the system forms a dodecagonal QC from the melt, which exhibits long-range polytetrahedral ordering [183, 52] (see Fig. 5.1a). Although the QC is physically stable over the timescale of a typical simulation, it is thermodynamically metastable with respect to the σ-phase periodic approximant [52] (see Fig. 5.1b). Here, we run simulations at temperature T=0.55, pressure P=3.5 and density ρ=0.85, which is slightly below the degree of supercooling ($T/T_m \sim 0.7$) at which the system forms a QC in the absence of a seed nucleus or specialized simulation techniques. At this state point, the growth of the solid phase occurs from a single nucleus, although under deeper supercooling many nuclei may grow simultaneously [184].

To observe the growing nucleus in our simulations, we define an order parameter to detect QC local ordering. Our order parameter is a modification of the $\mathbf{q}_6(i) \cdot \mathbf{q}_6(j)$ scheme of reference [57]. There, the nearest-neighbor directions of an atom i are expanded in spherical harmonics $Y_\ell(\theta, \phi)$ (with $\ell = 6$) to construct a $2\ell+1$ dimensional complex vector $\mathbf{q}_6(i)$, which can be thought of as a cluster "shape-descriptor" containing information regarding the shape and orientation of the cluster. An atom i forms a solid-like connection with neighbor j if the vector dot product $\mathbf{q}_6(i) \cdot \mathbf{q}_6(j)$ exceeds a certain value, and atoms with many solid-like connections are defined as being solid-like, reflecting the fact that in simple crystals all atoms have identical coordination shells. This scheme must be modified for QCs and approximants, since neighboring atoms have non-identical coordination shells corresponding to different Frank-Kasper polyhedra [44]. For dodecagonal QCs, we increase the range of the neighbor cutoff to $r_{cut} = 2.31\sigma$, corresponding to the first $\sim$2.5 neighbor shells. Also, we modify the set of harmonics from $\ell = 6$ to $\ell = 12$, since we find that $\mathbf{q}_{12}$ is sensitive to the symmetry of the dodecagonal QC, whereas $\mathbf{q}_6$ produces no signal. Pairs of atoms form a solid-like connection if $\mathbf{q}_{12}(i) \cdot \mathbf{q}_{12}(j) \geq 0.45$, with $\mathbf{q}_{12}(i) \cdot \mathbf{q}_{12}(j)$ normalized on the interval [0,1]. Atoms with $\geq 50\%$ solid-like connections are solid-like, otherwise they are liquid-like. These cutoffs are chosen so as to maximize the distinction between liquid and QC; however, we note that the distinction becomes ambiguous near the liquid-solid interface where atoms exhibit properties that are intermediate between liquid and solid. Therefore, for a diffuse nucleus, the solid-like atoms identified using this scheme represent only the nucleus core.

We next define $q_6(t) \equiv \mathbf{q}_6(i;t_0) \cdot \mathbf{q}_6(i;t)$, autocorrelation function that measures how correlated atomic configurations are at time t to their configurations at an earlier or later time t_0. We base our scheme on $\mathbf{q}_6$ rather than $\mathbf{q}_{12}$, since our goal is to quantify how closely clusters match in terms of shape orientation, rather than to detect quasicrystalline correlations between non-identical neighbor shells. We define $r_{cut} = 1.65$ to include the first

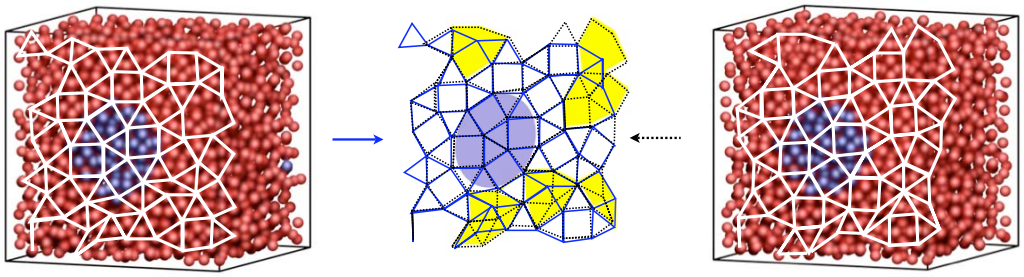


Figure 5.2 Dependence of QC tiling arrangement on liquid structure. The images show characteristic results from MC runs with the same quasicrystalline seed (blue) but with a different random number sequence. At certain points in space, highlighted in yellow in the overlay, the tiling arrangements differ.

neighbor shell in our analysis. We normalize $q_6(t)$ such that 1 is the maximum value and 0 represents the value for random correlations. Configurations that are less correlated than the average random value have $q_6(t) < 0$.

5.1.3 Effect of Randomness

We begin by considering the growth of the solid phase from a small static seed nucleus in the form of a periodic approximant [126] that is inserted into the MC simulation cell (see Fig. 5.2). Approximants are crystals with identical local ordering to QCs; therefore, for small nuclei, QCs and approximants are identical and the difference in long-range ordering results from a different growth mechanism. Constraining the seed in the form of an approximant allows us to determine whether the system requires a seed with a 'special' structure to grow a QC. We randomize our MC simulations at high temperature starting at time t_r before quenching to T=0.55 at t_q, at which point atoms begin to attach to the seed, causing rapid solidification. We observe that the system consistently forms a QC for all seed sizes, positions, and approximant structures, indicating that the system does not copy the seed, but rather incorporates atoms into the solid via a different paradigm.

Energy-driven QC growth models suggest that atomic attachment to the nucleus is deterministic, whereas entropy-driven models suggest that it is stochastic. We test the applicability of these models for our system by modifying the random number sequence (RNS) used during the simulation, holding all else constant. As depicted in Fig. 5.2, for the same seed nucleus (blue), we consistently obtain distinguishable QC tiling arrangements, indicating that QC growth has a stochastic element. It is clear that the growth is energetically

constrained as well, since most of the tiling discrepancies (yellow) represent "phasons [185]," tiling arrangements with nearly identical local energy. Thus elements of both growth models appear relevant to QC growth.

Although the growth of the QC is affected by the RNS, the attachment of tiles to the nucleus is not random. For random attachment, a change to the RNS causes an immediate change in the growth pathway, resulting in different tiling arrangements. In contrast, our system exhibits an appreciable lag time between changes to the RNS and the appearance of tiling discrepancies. For example, if we change the RNS at t_q, we observe fewer tiling discrepancies in the area immediately surrounding the nucleus than if we make a change at t_r. (Note that in both cases, the nucleus is identical since the solid does not begin to grow until t_q). This implies that QC growth is affected by stochasticity only insofar as it engenders differences in the local arrangement of atoms around the nucleus.

5.1.4 Effect of Liquid Structure

We can test this idea quantitatively by using $q_6(t)$ to detect structural correlations between atoms surrounding the nucleus and the QC tiles that they subsequently form. First, we generate many independent nucleation events in which the system grows a QC. Previously, we used a seed to initiate nucleation; here we use umbrella sampling to generate many configurations with growing nuclei. Our NPT MC runs are biased according to the harmonic weight function $w = \frac{1}{2}k(N-N_0)^2$ [57, 62]. Here, $k = 0.075$, N is the number of atoms comprising the nucleus (measured by $\mathbf{q}_{12}(i) \cdot \mathbf{q}_{12}(j)$), and N_0 is specified such that nucleus sizes near N_0 are sampled selectively. We slowly increase the bias from $N_0 = 10, 20, ...90$ so that nuclei reach $N = 80-100$. We then use these microstates as starting points for unbiased NPT MC runs. We observe that nuclei with $N > 75$ atoms tend to grow, although factors other than size (e.g., shape, structure, etc.) may affect nucleus stability as well [19]. We run MC simulations of growing nuclei for 75,000 MC cycles, the time it takes for nuclei to grow from $N \sim$100 to $N \sim$500.

We measure $\langle q_6(t) \rangle$ versus t in the non-equilibrium nucleating system described above for atoms that attach to the growing QC nucleus at $t_0 = 0$, which we refer to hereafter as "attaching atoms." For $t < 0$, attaching atoms are in the region surrounding the nucleus, and for $t \geq 0$, attaching atoms are in the solid nucleus (see Fig. 5.3a, middle curve). We include only the atoms that attach permanently to the nucleus in our analysis, to ensure that we measure correlations between atoms in the QC and their former (non-solid) configurations rather than correlated reattachments of solid atoms. Specifically, we exclude atoms that recross the 50% threshold for solid-like connections (defined above) after fluctuations are

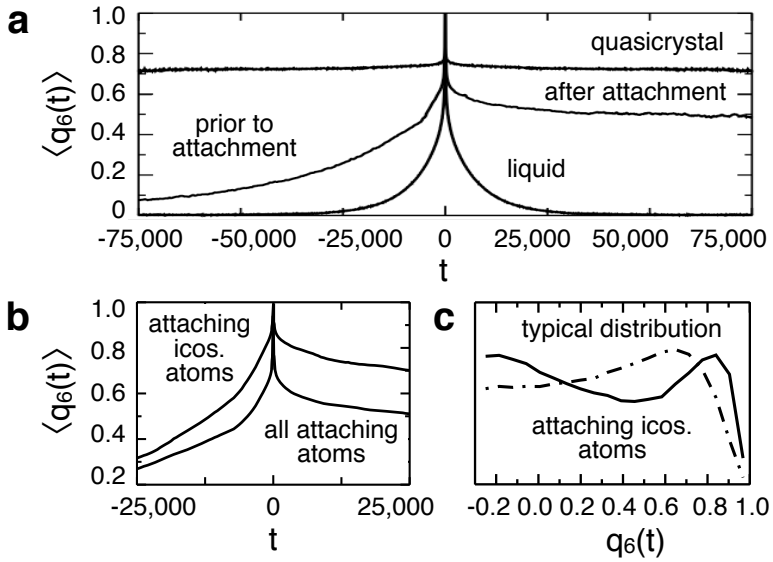


Figure 5.3 Structural correlations. (a) Average value of $q_6(t)$ versus t (MC steps). From top to bottom: atoms in the dodecagonal QC, atoms in the non-equilibrium nucleating system that attach to the nucleus at $t = 0$, atoms in the liquid. For all runs, the reference time $t_0 = 0$. (b) Average value of $q_6(t)$ versus t for attaching atoms. Top: attaching atoms in icosahedral configurations. Bottom: all attaching atoms. (c) Probability distribution of $q_6(t)$ at $< q_6(t) = 0.35 >$. Dotted line: the typical distribution for $< q_6(t) = 0.35 >$, calculated from atoms in the supercooled liquid. Solid line: attaching atoms in icosahedral configurations.

averaged out. Roughly 60% to 70% of the atoms attach without ever detaching.

We compare $\langle q_6(t) \rangle$ for attaching atoms to atoms in the bulk QC and the bulk supercooled liquid at the same state point (Fig. 5.3a). The value of $\langle q_6(t) \rangle$ is proportional to the degree of correlation to the reference structure at $t = 0$. This is exemplified by the high, constant value of $\langle q_6(t) \rangle$ observed for attaching atoms ($t > 0$) and bulk QC atoms, which indicates a solid-like environment. (The initial drop is due to thermal fluctuations). For $t < 0$, attaching atoms exhibit relatively high $\langle q_6(t) \rangle$, indicating that atoms joining the nucleus at $t = 0$ are highly correlated to their former (pre-solidification) configurations.

We can dissect the $\langle q_6(t) \rangle$ curve for attaching atoms into components based on local structure. Overall, the dodecagonal QC consists of atoms in four different types of coordination shells: icosahedral, Z13 , Z14, and Z15 configurations, where, 'Zn' stands for the Frank-Kasper polyhedron [44] with coordination number 'n.' We find that icosahedral atoms exhibit high $\langle q_6(t) \rangle$ (Fig. 5.3b), whereas other motifs do not deviate significantly from the average. We rationalize the high value of $\langle q_6(t) \rangle$ for icosahedral atoms by considering the probability distribution of $q_6(t)$ at each point on the $\langle q_6(t) \rangle$ curve (Fig. 5.3c). We find that atoms in icosahedra, and, to a lesser extent, atoms in Z13 configurations (not shown),

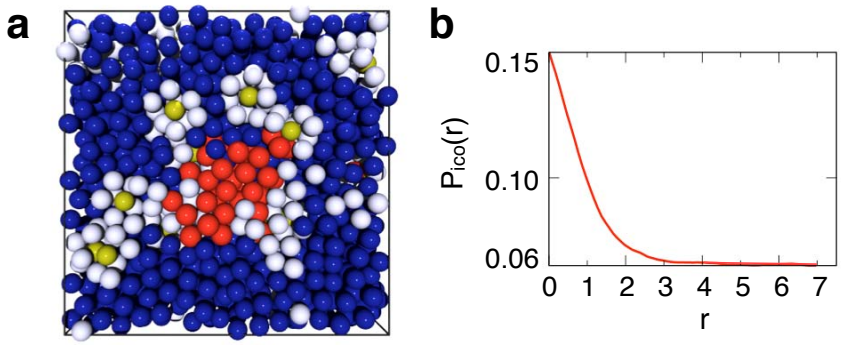


Figure 5.4 Icosahedral environment. (a) Simulation snapshot showing a QC nucleus (red) together with icosahedral clusters (yellow for icosahedral centers, white for surface atoms) in the liquid (blue). (b) The average probability of observing an atom at the center of an icosahedron versus r, the distance from the nucleus surface.

exhibit an unusually high proportion of strong correlations. This indicates that as the nucleus grows, it incorporates a certain subset of icosahedral and Z13 atoms with minimal structural rearrangement. Interestingly, Z14 atoms do not exhibit either high $\langle q_6(t)\rangle$ or a skewed $q_6(t)$ distribution, which indicates that although the icosahedral glass formed by the Dzugutov system has vibrational modes similar to the thermodynamically stable σ-phase [186] (25% icosahedra and 75% Z14), the most correlated atoms do not exhibit σ-like character. Rather, the high degree of icosahedrality and the presence of correlated Z13 atoms (which do not appear in the approximants but are highly present in the supercooled liquid) indicate that atoms in liquid-like icosahedral clusters surrounding the nucleus tend to retain their configurations during incorporation into the nucleus.

5.1.5 Icosahedral Wetting Layer

We can obtain a more intuitive picture of the role of icosahedral clusters by considering their spatial arrangement in relation to the growing QC nucleus. We generate a large number of nuclei using the umbrella sampling scheme outlined above. To expedite sampling, we allow configuration swapping between simulations via parallel tempering [62]. In all, we run 10 simultaneous MC simulations for 3.5 million MC steps, where each simulation has a unique biasing potential minimum $N_0 = 10, 20, \ldots 100$ for a given simulation. We save configurations every 100 MC steps, giving us 35,000 total microstates containing nuclei of sizes N=10-110 for analysis. We identify icosahedral clusters in our microstates using the method of reference [187], an extension of the method of reference [188].

As depicted in Fig. 5.4a, we find that icosahedral clusters (yellow, white) "wet" the core of the QC nucleus (red), a mechanism that may reduce interfacial tension [176]. We quantify the tendency for icosahedral clusters to aggregate around the nucleus by calculating $P_{ico}(r)$, the average probability of observing an atom at the center of an icosahedron a distance r away from the nucleus surface (see Fig. 5.4b). For nuclei of all sizes, we observe that $P_{ico}(r)$ starts with a value of 0.15 near the nucleus surface and decreases to the liquid value of 0.06 over a range of about three particle diameters, indicating that there is an increased presence of icosahedral clusters in the region surrounding the nucleus. As the nucleus grows, it must change the connectivity of these clusters from liquid-like local-energy minimizing arrangements to ordered quasicrystalline arrangements. The tendency to retain the configurations of some of the clusters rather than copying the nucleus surface template is the "growth rule" underlying the formation of the QC.

5.1.6 Conclusions

Our results demonstrate how QCs provide a 'path of least resistance' for solid phase growth versus crystals. In this case, whereas the stable σ-phase approximant must rearrange kinetically trapped atoms into a crystal lattice, the less constrained QC is able to reach a 'structural compromise' with the surrounding atoms to grow more rapidly. Our results explain why QCs often form in rapidly quenched metallic alloys, as these systems produce rapidly growing nuclei as well as low-energy icosahedral clusters. In terms of QC growth models, our results give physical insight into how the nucleus 'decides' to form a particular tile as it grows. We note that although icosahedral clusters are not the energy-minimizing structural motif for all QCs, the basic mechanism at hand − the tendency for certain atoms to retain their liquid configuration when incorporated into the growing solid nucleus − should hold generally for QC-forming systems.

5.1.7 Supplementary Information: Order Parameters

The $\mathbf{q}_{12}(i) \cdot \mathbf{q}_{12}(j)$ order parameter that we employ for distinguishing particles in quasicrystal-like and liquid-like local configurations is largely based on the $\mathbf{q}_6(i) \cdot \mathbf{q}_6(j)$ method for detecting fcc, bcc, and hcp crystal grains of reference [57]. To adapt the scheme for dodecagonal quasicrystals, we require several modifications. Our methods are best described within the context of the shape matching framework outlined in section 7, where we provide the quasicrystal order parameter as an example problem (see section 7.2.15).

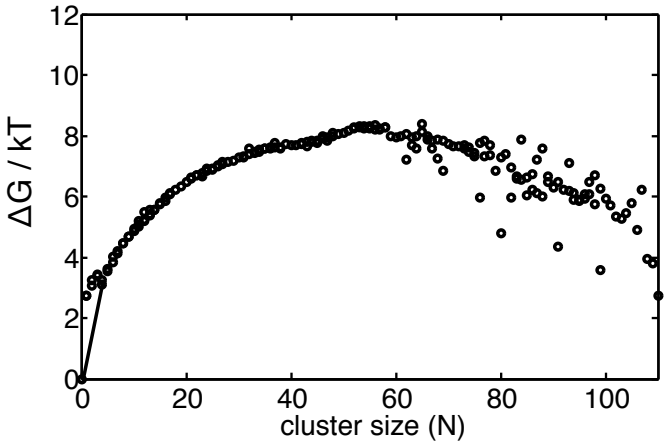


Figure 5.5 Average Gibbs free energy as a function of the nucleus size N, computed via umbrella sampling [57]. The nucleus size is not the only reaction coordinate for this system, which is evidenced by the spread in the committer P_B as a function of N [60] (not shown).

The $q_6(t)$ autocorrelation function that we employ is a relatively straightforward application of the shape matching framework to a time correlation function. Several similar examples are given in section 7.3.5. We use the $\mathbf{q}_6$ shape descriptor because it sufficient for detecting structural changes in the types of local clusters that arise in our system.

5.1.8 Supplementary Information: Free Energy Calculation

In the main text, we state that nuclei of a given size tend to grow on average, whereas smaller nuclei tend to shrink. This tendency is quantified by the average Gibbs free energy as a function of the nucleus size N, computed using the umbrella sampling [156] scheme of reference [57] (Fig. 5.5). We see that, on average, nuclei of size $N \geq\sim 60$ become increasingly stable with N and therefore tend to grow, whereas smaller nuclei are unstable and tend to shrink. The curve shown in Fig. 5.5 considers N exclusively, and thus only represents an averaged free energy measurement; reaction coordinates other than N may be important for determining the stability of a given nucleus. Nonetheless, N is a sufficiently important reaction coordinate that the larger nuclei that we choose for our runs of size $N \sim 75$ tend to grow for the runs that we perform.

5.2 Self-Assembly of Soft Matter Quasicrystals

We propose a two-part mechanism for forming 3d dodecagonal quasicrystals and their approximants by self-assembly. First, we attach small mobile particles to the surface of spherical particles to encourage icosahedral packing. Second, we allow a subset of particles to deviate from the spherical shape, to discourage close-packing. We find that model soft-matter systems with both of these properties consistently assemble dodecagonal quasicrystals and/or approximants. We predict that many systems may assemble quasicrystals by this mechanism, including micelle-forming and nano-satellite systems.

5.2.1 Introduction

Quasicrystals are a class of ordered solids with long-range orientational ordering but no long-range periodicity [189]. Their unique structure gives rise to favorable properties such as low surface tension and high wear resistance on the atomic scale and a full photonic bandgap on the micron and nanometer scales [190, 191, 28, 27]. Despite their useful properties, their potential applications are limited by their scarcity. Only recently have quascrystals and their approximants been observed in non-atomistic systems. Examples include holographically trapped [192] and laser-field induced [193] quasicrystals made of micron sized spheres, as well as self-assembled quasicrystals and approximants formed by binary nanoparticle superlattices [125], phase-separated star-triblock copolymers [194], spherical micelles of phase-separated diblock copoylymers [195], and spherical dendrimer micelles [191, 196]. Computer simulations predict that quasicrystals can be assembled from systems of spherical particles with unusual pair potentials [183, 52, 131], however, such potentials have not been realized experimentally, even for highly tunable systems such as colloidal suspensions. Recent simulations have shown that quasicrystals can form in systems with solely excluded-volume interactions for tetrahedrally-shaped particles [33].

Here, we propose and test a method for self-assembling dodecagonal quasicrystals (DQCs) and their periodic approximants [126] in bulk 3d systems of (roughly) spherical particles, that, we predict, is both widely applicable and experimentally feasible. Our mechanism is inspired by the basic ideas underlying the Dzugutov (DZ) [183] and Lennard-Jones-Gauss (LJG) [131] potentials that are both known to form DQCs in systems of point particles. These potentials, like the standard Lennard-Jones (LJ) potential, promote icosahedral packing locally, but, unlike the LJ potential, include a relative energy penalty for long-ranged fcc and hcp packing, causing DQCs to form instead. We exploit relatively common features of soft-matter-like systems to achieve the same effect without turning to

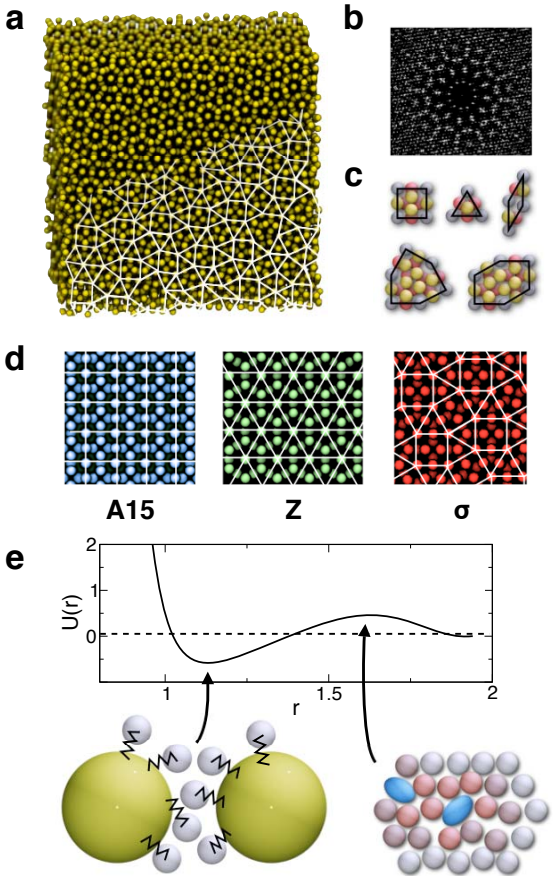


Figure 5.6 Using packing constraints alone to mimic the effect of a complex pair potential. (a) A dodecagonal quasicrystal formed in the Dzugutov system at $T = 0.42$, $\rho = 0.85$. The particles are energy minimized to remove thermal vibrations. (b) Diffraction pattern for the structure shown in (a). (c) Valid tiles for the dodecagonal quasicrystal model. Left to right, top to bottom: square, triangle, rhomb, shield and hexagon. (d) Three simple quasicrystal approximants, formed by periodic arrangement of the tiles. Left to right: A15 (squares), Z (triangles), σ (squares and triangles). (e) Important characteristics of the Dzugutov potential and their packing counterparts. The potential well, which favors local icosahedral ordering, is replaced by attaching mobile particles to the surface of the larger particles. The potential repulsion, which hinders the formation of fcc, hcp, and bcc crystals, is replaced by adding particle shape polydispersity to the system.

complex potentials (see Fig. 5.6). Instead, the promotion of local icosahedral ordering is achieved by functionalizing particles with mobile entities connected to their surface. We show that, as predicted by Ziherl and Kamien [197] for spherical dendrimers, this creates an entropic driving force that promotes local polytetrahedral ordering. The suppression of crystalline ordering is accomplished by introducing shape polydispersity (in the form of asphericity) into the system. We show that DQCs and their approximants are better able to incorporate aspherical particles than standard fcc, hcp, and bcc crystals. Using molecular

simulations, we show that three different model systems that exhibit both of these properties reproducibly assemble DQCs and/or their approximants. These models, given by a minimal model micelle and two micelle-forming systems composed of tethered nanosphere building blocks [198, 199], represent the only simulated soft matter systems currently known to form 3d DQCs or approximants. The minimal micelle model is one of only two model systems that form quasicrystals or appoximants without attractive interparticle interactions [33]. Due to their similarities, the models and mechanism that we study may provide additional insight into how DQCs and/or approximants form for systems composed of spherical dendrimers, as reported by Zeng, *et al.* [191], and spherical block copolymer micelles, as reported by Lee, *et al.* [195]. Finally, our study serves as a representative example of how complex model potentials can be effectively achieved through particle packing and highlights the importance of studying these models.

5.2.2 Minimal Micelle Model

We begin by exploring an explicit minimal model of a spherical micelle (MSM) that only considers excluded volume interactions between terminal groups on the micelle surface (Fig. 5.7a). The MSM consists of a non-interacting rigid scaffolding with 42 lattice points on the surface of a sphere, given by the vertex points of a 2-frequency icosahedral geodesic with diameter = 5σ. Each lattice point anchors a small spherical particle with diameter σ attached by a harmonic spring (defined as $U(r) = kr^2$) with spring constant k, which controls the degree of surface particle mobility. Excluded volume interactions for the surface spheres are modeled by the purely-repulsive Weeks-Chandler-Andersen (WCA) potential [118]. We perform NVT MD simulations using LAMMPS [148] at $T = 1.0$ and effective volume fraction, $\phi \approx 0.8$ (computed assuming a micelle radius of 3σ). Roughly speaking, the MSM can represent many different nanoscopic objects, including core-satellite nanoparticles [200, 201], or spherical micelles composed of dendrimers [197, 191], block copolymers [195], or tethered nanoparticles [198, 199].

In the absence of shape polydispersity, the MSMs tends to form close-packed (fcc or hcp) arrangements for $k > 5$, and bcc structures for higher surface particle mobility, i.e., $k \leq 5$. A bcc ordered structure of 60 MSMs is shown in Fig. 5.7b for k=5. We find a dramatic change in the structural arrangement of the MSMs when shape polydispersity is incorporated into the system in the form of aspherical "dimer" micelles (see Fig 5.7c). We allow dimers to form naturally by taken advantage of the fact that at low k some of the MSMs will overlap and become locked together if k is increased. By starting from $k = 2$ and slowly increasing k, we create systems with dimer fraction $0.20 \leq f_{dimer} < 0.40$ and an average aspect ratio

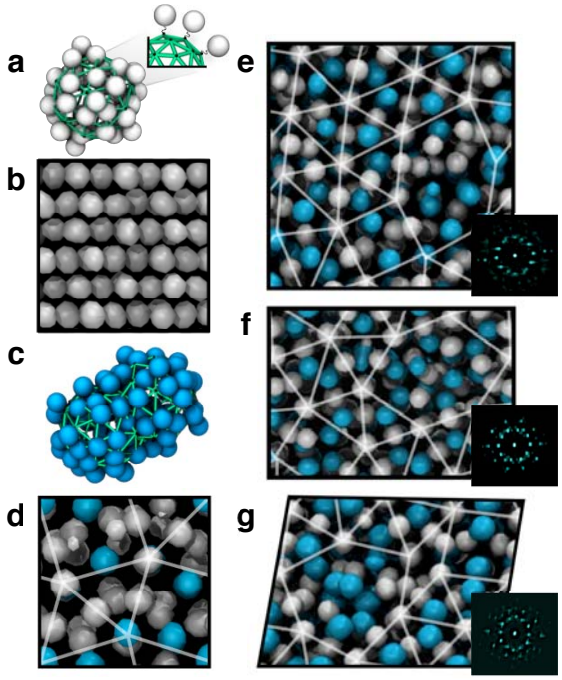


Figure 5.7 Minimal spherical micelle (MSM) model with mobile surface particles. (a) MSM monomer (white) extracted from a simulation. (b) 60 MSMs at $k = 5$ with no dimers in a bcc configuration. (c) MSM dimer (cyan) extracted from a simulation. (d) 60 MSMs with $k = 5$ and $f_{dimer} \approx 0.24$ in a sigma structure. (e-g) Systems with 360 MSMs and: (e) k =4 and $f_{dimer} \approx 0.39$, (f) k =4.75 and $f_{dimer} \approx 0.37$, and (g) k =5 and $f_{dimer} \approx 0.36$. In all cases, we plot time-averaged isosurfaces of the centers of mass of the micelles/dimers, rather than micelles themselves. Systems are viewed along the axis with 12-fold symmetry, as calculated using the diffraction pattern, shown to the right. Note, (e) only appears as a parallelogram due to the projection.

of 1.45.

Three representative independent simulations, each composed of 360 MSMs in rectangular boxes with aspect ratio 1.28:1.28:1.00, are plotted in Figs. 5.7e-g. Figs. 5.7e,f, and g show systems at k =4, 4.75, and 5, with f_{dimer} =0.39, 0.37, and 0.36, respectively. In all cases, we observe DQC-like FK structures, which are viewed by orienting the systems along the pseudo 12-fold axis. While the number of MSMs used here is small compared to the number of atoms in typical atomistic simulations [183, 131], the structures in Fig. 5.7e-g exhibit the hallmarks of quasicrystalline ordering. The systems form unique "tilings" with different configurations rather than any particular approximant structure. We note that we are limited to smaller systems than typical point-particle models due to the fact that we must resolve the timescale on the motion of the surface particles that comprise the MSM, rather than the centroid. Since DQCs grow more easily than approximants [30], it is possible that the DQC-like tilings are thermodynamically metastable relative to a more

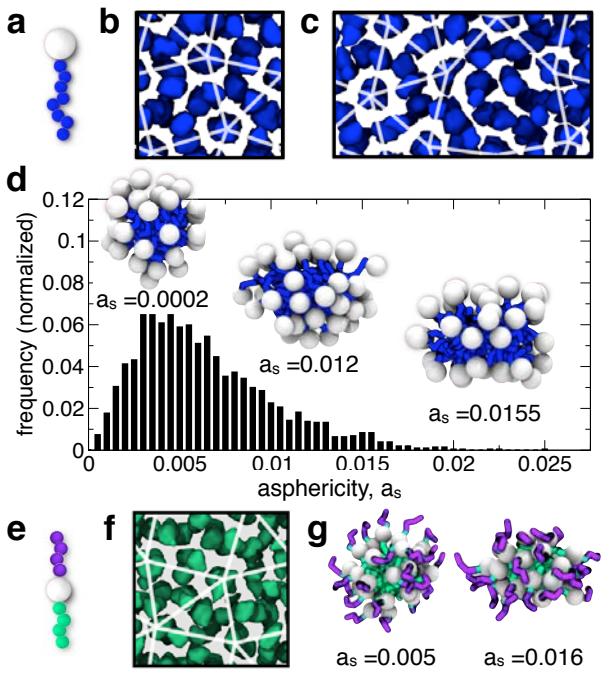


Figure 5.8 Tethered nano sphere (TNS) systems. (a) Schematic of a mono-TNS building block, where the 8 tether beads (blue) of size σ aggregate, that self-assemble spherical micelles with a soft core surrounded by relatively hard "satellite" nanoparticles (white) of size 2.5σ that act as mobile surface entities. (b)$\sim$60 micelles formed by mono-TNS that arrange into a sigma approximant, and (c) $\sim$120 mono-TNS micelles that form a DQC-like arrangement; both systems at $\phi = 0.275$ and T=1.1. (d) Histogram of asphericity, a_s, of the mono-TNS micelles shown in (b); representative micelles are overlaid on the figure. (e) Schematic of the di-TNS building block where the 4 tether beads (green) of size σ aggregate and nanoparticles (white) of size 2σ are also attractive; 4 tether beads (blue) of size σ that do not aggregate coat the outside of the micelle. (f) $\sim$60 di-TNS micelles arrange into a sigma approximant at $\phi = 0.2$ and T = 1.2. (g) Representative di-TNS micelles at with different a_s. All results were simulated using an in-house Brownian dynamics code [198, 199] and in all cases we show density isosurfaces of the aggregating polymer tethers.

stable approximant structure; however, the relative stability of DQC versus approximants is difficult to determine even for trivial point-particle models and we note that spherical dendrimers have been reported to form both DQCs [191] and sigma approximants [196] in experiment.

5.2.3 Tethered Sphere Model

We can further test the mechanism for a system where we do not have explicit control over surface particle mobility or shape polydispersity. We consider two model tethered nanosphere (TNS) systems, mono-TNS [198] and di-TNS [199], both of which form roughly

spherical micelles with mobile surface entities. Schematics of the building blocks are shown in Figs. 5.8a,e respectively, and micelles they form are shown in Figs. 5.8d,g respectively. The mono-TNS micelles have an outer shell of mobile nanospheres that closely match the MSM model, while the di-TNS micelles have a shell of short polymers, similar to spherical micelles formed by block copolymers [195] and dendrimers [191, 197] . Details of the models and methods are given in Refs. [198, 199]. These models are computationally expensive, and thus only relatively small systems are explored. Fig. 5.8b depicts density isosurfaces of the aggregating tethers for a system of 2500 mono-TNS building blocks that assemble into $\sim$60 spherical micelles arranged in a sigma appoximant and Fig. 5.8c depicts isosurfaces for a system of 5000 mono-TNS that self-assemble into $\sim$120 spherical micelles that form a DQC-like arrangement of tiles. The mono-TNS micelles naturally exhibit shape polydispersity. Fig. 5.8d shows a histogram of the asphericity, a_s, computed from the principle radii of gyration [202] of the micelles, with representative micelles at various values of a_s inset ; for reference, the MSM dimer shown in Fig. 5.7c has $a_s = 0.02$. Fig. 5.8f shows a sigma structure formed from 2000 di-TNS building blocks that self-assemble into $\sim$60 micelles. The distribution of a_s for the di-TNS is similar to the mono-TNS. Two representative di-TNS micelles at low and high a_s are depicted in Fig. 5.8g. Overall, DQC-like structures assembled from TNS micelles were observed in 20 independent simulations.

5.2.4 Stabilization Mechanism

We now explore the thermodynamic basis underlying both aspects of our mechanism. The first aspect, the functionalization of particles with mobile surface entities, is inspired by the theoretical work of Ziherl and Kamien [197]. The authors argue that the stability of the A15 structure, a low order DQC approximant, in systems of dendrimeric micelles can be traced to the tendency for neighboring micelles to minimize surface contact area [197]. This mechanism states that entropy is maximized when steric interactions between terminal polymers on the micelle surface are minimized [197]. The space-filling arrangement of cells that best minimizes surface contact area is the Weaire-Phelan structure, which corresponds to the A15 structure [203], and serves as a counter-example to the famous conjecture by Kelvin that the bcc lattice minimizes surface area [203]. A similar mechanism is thought to stabilize the related dendrimer DQC observed in experiments [191]; however it is not clear why this mechanism alone might stabilize a DQC or a higher-order approximant rather than the A15 structure.

We calculate the Helmholtz free energy F as a function of k for a system of monodisperse MSMs (i.e., without dimers), described previously, relative to the hcp crystal, taken as a

convenient reference state. Here, we calculate the relative free energy for a given structure using free energy perturbation [204] (FEP), where the reversible work required to change k is given by $F(k_j) - F(k_i) = -k_B T \ln\langle \exp\left[-(U_{k_j} - U_{k_i})/k_B T\right]\rangle_{k_i}$. The calculation is performed in multiple stages (8 total) to avoid asymmetric bias [204]. The offset between the curves for different structures is computed using an FEP variant of the standard Frenkel-Ladd method for molecular systems [205] for systems with $k = 5$. FEP is used to adapt the method for our complex molecules; although this method is non-standard, it gives a reasonable estimate of F consistent with our simulation results. This only effects the vertical offset of the curves and not how F changes as a function of k. Details of this scheme are given in the SI below.

In Fig. 5.9a we see that as k decreases (i.e. surface particle mobility increases), F decreases more rapidly for the A15, dod, and bcc structures than for the fcc and hcp structures (note, the "dod" phase is the average value calculated from several higher-order DQC approximants [206]). For low k, bcc appears to be the stable state, similar to our simulations results. The change in F is the strongest for the A15 structure, which minimizes surface contact area, followed by the dod and bcc structures respectively. This change in F is entropically driven, since the difference in average potential energy $\langle U \rangle$ changes little with k, and does not decrease with F (Fig. 5.9a, inset); this serves as a direct verification of the predictions of Ziherl and Kamien [197].

We explore the second aspect of our mechanism, the addition of shape polydispersity, by simulating binary mixtures of soft spheres and short, pill-shaped dimers, modeled by the WCA potential (see Fig. 5.9b). The dimers are modeled by a short rigid body of length 1.5σ consisting of two overlapping soft spheres 0.5σ apart. Fig 5.9b shows the Helmholtz free energy F as a function of the dimer fraction f_{dimer} for several structures at a statepoint $\rho = 0.9$, $T = 0.25$. The free energy is computed based on the standard Einstein crystal thermodynamic integration (TI) method [140], with an additional TI step to compute the free energy required to transform a given fraction of spheres into dimers. Details of this scheme are given in the SI below. As f_{dimer} increases, A15 and dod structures become increasingly stable relative to close packed crystals, and, to a lesser extent, bcc. This difference in stability can be traced to the tendency for dimers to adopt larger, more aspherical neighbor shells, which are present in FK structures but not fcc, hcp or bcc. For example, in the dod phase, dimers tend to sit at large Z15 FK coordination shells [44] in a $\sim 7:1$ ratio over smaller Z14 shells.

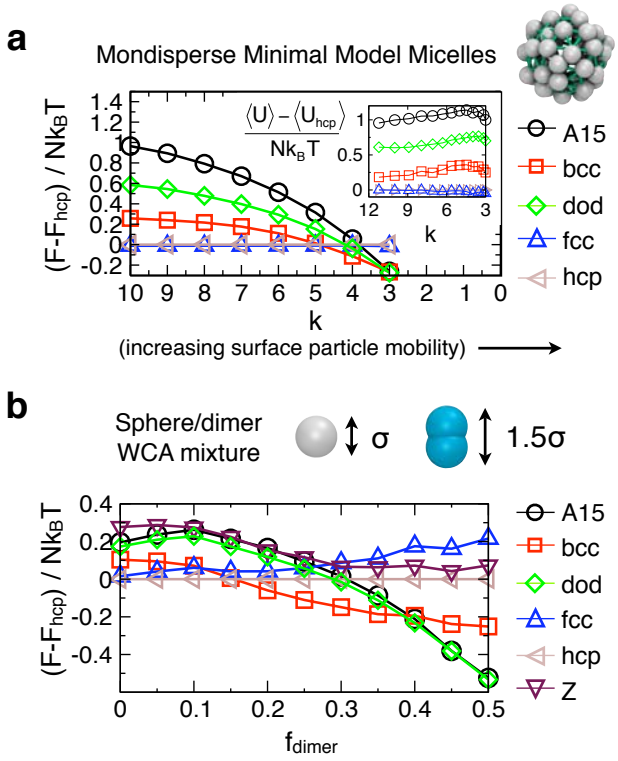


Figure 5.9 Effect of mobile surface particles and particle shape polydispersity. (a) Helmholtz free energy per molecule, F, as a function of surface particle mobility (controlled by the spring stiffness k) for monodisperse MSMs. The inset shows the potential energy per molecule U. (b) F as a function of f_{dimer} for the WCA sphere/dimer mixture. For parts (a) and (b), the energies are reported with respect to the hcp crystal for convenience. For all datapoints, error bars are smaller than the data labels.

5.2.5 Conclusions

For $f_{dimer} > 0.4$, the FK phases (A15 and dod) are more stable than close-packed and bcc crystals. This implies that mixtures of spherical and pill-shaped colloids or nanoparticles [207, 208] might produce DQCs or approximants. However, since many dimers are required to destabilize crystal structures, in practice such mixtures may remain liquid-like or form other structures not considered here. Along this same line, it is possible that, in specific cases, systems may form DQCs based on mobile surface particles alone; the entropic effect may be stronger for terminal polymers that are longer than the one-bead model tested here. However, since asphericity is common in many micellar systems that also have mobile surface entities, it is likely that, as our results suggest, both effects play a role in the formation DQCs and approximants [191, 196, 195]. Overall, we find that the combination surface particle mobility and asphericity tends to result in DQC-like phases for

a broad range of systems.

5.2.6 Supplementary Information: Simulation Methods

In this section, we provide a detailed description of the methods used to produce the results reported in this study. These methods can be used to exactly reproduce our results, which may aid in future studies of these systems. Note in all cases, we use Lennard-Jones reduced units [140].

Self-Assembly Simulations of Model Micelles

As previously described in the main text, spherical micelles are modeled as mobile spheres tethered to the surface a large spherical surface. The spherical surface is modeled as a rigid scaffolding of 42 lattice points, described by the vertex points of a 2-frequency icosahedral geodesic with diameter = 5σ. Each lattice point anchors a small spherical particle with diameter=σ, attached by a harmonic spring. The energy of the harmonic spring is defined following the convention in LAMMPS [148],

$$U_{harmonic} = k(r - r_0)^2 \tag{5.1}$$

where k is the spring constant, r is the separation between the particle and its attachment point, and r_0 is the equilibrium separation, set to 0 in this case. The rigid lattices are non-interacting and are used simply to provide attachment points for the mobile surface particles. Surface particles interactions are modeled using the Weeks-Chandler-Andersen [118] potential,

$$U_{WCA} = \begin{cases} 4\varepsilon \left(\dfrac{\sigma^{12}}{r^{12}} - \dfrac{\sigma^6}{r^6} \right) + \varepsilon \, , & r < r_{cutoff} \\ 0 \, , & r \geq r_{cutoff} \end{cases} \tag{5.2}$$

where ε = 1.0, σ = 1.0, and $r_{cutoff} = 2^{1/6}$. The model micelles were simulated using NVT molecular dynamics (MD) with the LAMMPS software package [148]. The Nose-Hoover thermostat was used with T=1.0 and timestep = 0.005. The volume fraction, $\phi \approx 0.8$, assuming a characteristic diameter of 6σ of the micelles.

The following general simulation procedure was employed. The systems were started as a disordered arrangement of micelles at $k = 2$. The spring constant was then incrementally increased by 0.25 until the final value was reached, typically k = 4 - 5. At each value of k,

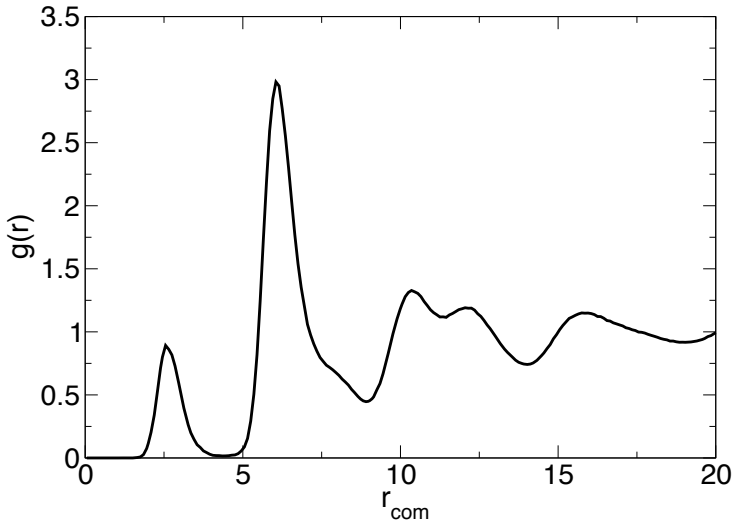


Figure 5.10 $g(r)$ calculated for the centers of mass of the model micelles in a FK structure at $k = 4$. The formation of dimers is evidenced by the first peak at r$\sim$ 2.6 which corresponds to overlap between the micelles.

the systems were typically run for between 10 to 50 million timesteps, depending on the system size (i.e. larger systems were run for longer).

As previously discussed in the main text, simulations performed at low k have a natural tendency to overlap. Incrementally increasing k results in the formation of dimer micelles. The formation of dimers is evidenced in plotting the radial distribution function, $g(r)$, of the centers of mass of the micelles, as shown in Fig. 5.10. The first peak in $g(r)$ occurs at approximately 2.6σ and corresponds to the overlap of micelles. Assuming a characteristic diameter of micelles of 6σ, the average dimer-to-monomer size ratio is 1.44. We note that the first peak in $g(r)$ is relatively symmetric, indicating a narrow distribution of dimer sizes around the average.

Self-Assembly Simulations of Tethered Nanospheres

The coarse-grained simulation model and methodology used to study the self-assembly of tethered nanospheres (TNS) (both mono- and di-TNS) is discussed briefly below and in greater detail in references [198, 209, 210]. Our general model consists of bead spring representations of the TNS, similar to what is typically used to study block copolymers and surfactants.

Mono-TNS

The mono-TNS building block consists of a chain composed of 8 spherical beads of diameter σ, connected via finitely extensible non-linear elastic (FENE) springs. Each chain is in turn attached via a FENE spring to a nanosphere of diameter D = 2.5σ. The potential energy of the FENE spring is given by:

$$U_{FENE}(r) = -\frac{1}{2}kR_o^2 \ln\left[1 - \left(\frac{r}{R_o}\right)^2\right] \tag{5.3}$$

where k is the spring constant, r is the separation between the points, and R_o is the maximum allowable separation. Here k=30 and R_o= 1.5. Tethers are treated as "solvent-phobic" and thus aggregate at sufficiently low T. To model this aggregation, the attractive Lennard-Jones (LJ) potential is used:

$$U_{LJ} = \begin{cases} 4\varepsilon\left(\dfrac{\sigma^{12}}{r^{12}} - \dfrac{\sigma^6}{r^6}\right) - U_{shift}\ , & r < r_{cutoff} \\ 0\ , & r \geq r_{cutoff} \end{cases} \tag{5.4}$$

where U_{shift} is the energy at the $r = r_{cutoff}$ and $r_{cutoff} = 2.5$. All other interactions are treated with the purely repulsive WCA potential (Eqn. 5.2), appropriately radially-shifted to account for excluded volume. Simulations were performed using Brownian dynamics, where the volume fraction of individual beads was varied between $0.25 \leq \phi \geq 0.30$. Systems were found to order at approximately $T \leq 1$.

The general simulation procedure used is as follows. We start with a disordered mixture of mono-TNS above the order-disorder temperature, where little-to-no aggregation occurs. We then incrementally cool the system, allowing our simulations to run for several million timesteps at each temperature, monitoring system potential energy to ensure we have reached a steady state before subsequent cooling.

Simulations were typically run for approximately 40 million timesteps. Multiple independent cooling sequences (i.e. different cooling rates) were performed to ensure reproducibility of results. Simulations were performed in systems of 2500 building blocks in cubic boxes (22500 total beads) and 5000 building blocks in boxes with aspect ratio 2:2:1 (45000 total beads). While still relatively coarse-grained in their detail level, the mono-TNS model is more computationally expensive than the model micelle system; the model micelle system only considers the outermost surface layer of the micelles, whereas the mono-TNS model additionally incorporates the particles that make up the micelle core.

Di-TNS

Di-TNS are modeled in much the same way as mono-TNS described above. Chains composed of 4 beads of diameter σ are connected via FENE springs (Eqn. 5.3). Two chains are connected to a single nanosphere of diameter D = 2.0σ, diametrically opposed. This planar angle of 180 degrees between the chains is maintained by the use of a standard harmonic spring (Eqn. 5.1) between the first beads of the two polymers. The two polymer chains are chemically distinct. One chain is considered to be solvent-phobic (i.e. attractive), and thus is treated with the LJ potential (Eqn. 5.4). The other chain is considered to be solvent-philic (i.e non-attractive) and treated with the WCA potential (Eqn. 5.2). Nanosphere-nanosphere interactions are modeled with the LJ potential, appropriately radially-shifted to account for excluded volume. All other interactions are modeled by the WCA potential, appropriately radially-shifted. Simulations were performed using Brownian dynamics, where the volume fraction of the individual beads was set to $\phi = 0.20$. The systems were found to order at $T \approx 1.2$. The general di-TNS simulation procedure matches the mono-TNS procedure described above.

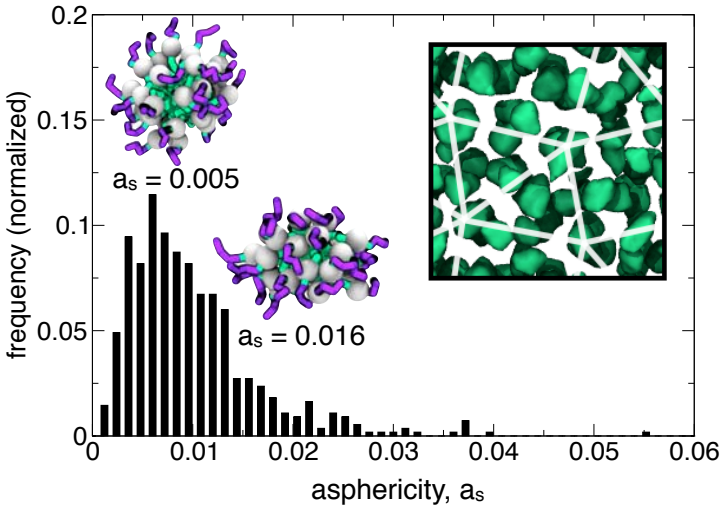


Figure 5.11 Histogram of asphericity, a_s, of the di-TNS micelles. Two representative micelles at different levels of a_s are inset in figure, along with a schematic of the sigma tiling, formed by the di-TNS.

The micelles formed by the di-TNS are roughly spherical in nature, however a subset of micelles possess distinct shape polydispersity. A representative histogram of the asphericity of the di-TNS micelles is plotted in Fig. 5.11, closely matching the mono-TNS histogram plotted in the main text.

Lattice Potential Energy Calculations

Some simple insight into the effect of shape polydispersity can be obtained by computing the lattice potential energies for several different crystal structures as a function of the fraction of dimer particles, f_{dimer}. For our study we use the A15, bcc, fcc, hcp, σ, and Z structures, as well as the two higher order quasicrystal approximants discussed below. Our calculations involve running a short canonical (NVT) MC simulation at constant density, where particles are allowed to rotate about their lattice positions, and dimers are allowed to swap with monomers. Particles are not permitted to translate. The system temperature is held constant at T=1.0 for all runs. For each calculation, we perform a long equilibration run to slowly step down the density from an initially sparse state of $\rho = 0.1$. This allows the dimers to rearrange to their equilibrium lattice positions and rotations so that the system does not become trapped in a metastable state.

Free Energy Calculations for Model Micelles

To study the effect of surface particle mobility, we compute the relative change in Helmholtz free energy F as a function of the strength of the harmonic springs k that tether particles to the surface of the model micelles, as defined in Eqn. 5.1. To do so, we use free energy perturbation with changing parameter k. For each crystal structure, we run 8 separate stages $k = [3, 4, ...10]$ where we calculate the ensemble average:

$$F_{k_j} - F_{k_i} = -k_B T \ln \left\langle \exp\left(-\frac{U_{k_j} - U_{k_i}}{k_B T} \right) \right\rangle_{k=k_i}. \tag{5.5}$$

Here, the per molecule potential energy difference $U_{k_j} - U_{k_i}$ is computed by evaluating the energy with different k using configurations in the ensemble $k = k_i$. The total free energy change is reported with respect to a reference point at $k_{ref} = 10$ chosen because all crystals remain physically stable up to this point. The free energy change for a given value of the spring strength k is then given by:

$$\Delta F_{fw}(k) = \sum_{k_i=k}^{k_{ref}-1} F_{k_i} - F_{k_{i+1}} \tag{5.6}$$

Equivalently, we can compute the free energy change by:

$$\Delta F_{bw}(k) = \sum_{k_i=k+1}^{k_{ref}} -(F_{k_i} - F_{k_{i-1}}) \tag{5.7}$$

In practice, we ensure that both the forward and reverse calculations are equivalent to rule out the possibility of asymmetric bias [204]:

$$\Delta F(k) = \Delta F_{fw}(k) = \Delta F_{bw}(k) \tag{5.8}$$

The calculations are carried out for NVT MD trajectories created using LAMMPS [148], as outlined above. For each lattice structure, independent MD simulations are carried out for each stage, all of which have a different value of k. The lattice positions are given by the A15, bcc, fcc, hcp, σ, and Z lattices, as well as the two higher order quasicrystal approximants discussed above. All lattices contain between 1000 and 2000 particles. For each simulation, a sparse system is slowly compressed to a target density with particles fixed to their lattice positions with strong springs. The springs are then released, and the system is equilibrated for 2.5 million MD steps, followed by a production run of length 5 million MD steps. For each window we compute 10,000 values of the energy difference $\frac{U_{k_j}-U_{k_i}}{k_B T}$.

Free Energy Calculations for Sphere-Dimer Mixtures

To study the effect of mixing dimers with spherical particles, we calculate the Helmholtz free energy F as a function of dimer fraction f_{dimer}. How we compute F depends on the particular system under investigation. For systems that are relatively stable in FK structures without any dimers, such as the Dzugutov (DZ) system, we use a two-step thermodynamic integration (TI) scheme.

In step I, we compute the free energy difference between a non-interacting harmonic (Einstein) crystal and a system of spherical particles interacting with the chosen pair potential using the standard Frenkel-Ladd method [140]. To do so, we simulate a system with a potential energy function given by:

$$U(\lambda)_{\mathrm{I}} = U_0 + (1-\lambda)(U-U_0) + \lambda \sum_{i=1}^{N} \alpha(\mathbf{r}_i - \mathbf{r}_{0,i})^2 \tag{5.9}$$

Here, λ is a switching parameter $\lambda \in [0,1]$ that changes the system between the Einstein crystal and the system with pair interactions. The Einstein crystal spring constant is given by α. The variable U is the system energy evaluated for the pair potential, U_0 is the potential energy of the system when all particles sit at their lattice positions, given by $\mathbf{r}_{0,i}$. The lattice positions are given by the A15, bcc, fcc, hcp, σ, and Z lattices, as well as the two higher order quasicrystal approximants. All lattices contain between 1000 and 2000 atoms. We perform thermodynamic integration by integrating over the derivative of the energy function:

$$\Delta F_I = \int_{\lambda=1}^{\lambda=0} d\lambda \left\langle \frac{\delta U(\lambda)}{\delta \lambda} \right\rangle_\lambda = \left\langle \sum_i \alpha (r_i - r_{i,0})^2 - (U - U_0) \right\rangle_\lambda \tag{5.10}$$

Here, we run a series of NVT MC simulation with the energy functions $U(\lambda)$, with λ incrementing from 0 to 1, and at each point evaluate the derivative given in equation 5.10. Each simulation is started from an ideal lattice and then equilibrated before collecting statistics. The derivative is evaluated for 20 independent simulations total, with small increments $d\lambda$ at the interval endpoints 0, 1 where the derivative changes the fastest.

In the second step of our TI scheme, we compute the work required to change a given fraction, f_d of the spherical particles in the system into dimers. We consider a system with the energy function:

$$U(\lambda)_{\mathrm{II}} = (1-\lambda) U_{\mathrm{pure}} + \lambda U_{\mathrm{mix}} \tag{5.11}$$

The free energy required to change the pure system to the binary mixture is the integral over the derivative with respect to the switching parameter λ:

$$\Delta F_{II} = \int_{\lambda=1}^{\lambda=0} d\lambda \left\langle \frac{\delta U(\lambda)}{\delta \lambda} \right\rangle_\lambda = \left\langle U_{\mathrm{mix}} - U_{\mathrm{pure}} \right\rangle_\lambda \tag{5.12}$$

Again we run 20 independent MC simulations for different values of λ. For each simulation, we slowly compress systems to their target density while allowing particle rotations and swaps, similar to our lattice energy calculations above, to avoid falling into metastable traps. Each simulation is then equilibrated at the target density, and statistics are collected. Our total free energy for the binary mixture is the given by the free energy sum of the two stages, plus the reference Einstein crystal:

$$\Delta F = \Delta F_{Ein} + \Delta F_I + \Delta F_{II} \tag{5.13}$$

This method is accurate when the pure component system (i.e., all spherical particles) acts approximately as an Einstein crystal, but fails otherwise. In the latter case, the values of the derivative in step I become exorbitantly high, leading to significant numerical error in evaluating the derivative. This is indeed the case for the WCA and LJ systems; the lattices for the FK structures are highly distorted, and do not exhibit harmonic vibrations about the ideal lattice positions. To address this, we add an additional step to our TI scheme. We first carry out steps I and II as before, for an interaction potential that gives harmonic behavior, such as the DZ system. Then, we compute the work required to change the potential from DZ to the target potential, such as the WCA potential. This involves computing the energy

function:

$$U(\lambda)_{\rm III} = (1-\lambda)U_{\rm DZ} + \lambda U_{\rm WCA} \tag{5.14}$$

The free energy difference for changing potential is given by integrating over the derivative with respect to the switching parameter λ:

$$\Delta F_{\rm III} = \int_{\lambda=1}^{\lambda=0} d\lambda \left\langle \frac{\delta U(\lambda)}{\delta \lambda} \right\rangle_{\lambda} = \langle U_{\rm WCA} - U_{\rm DZ} \rangle_{\lambda} \tag{5.15}$$

Again we run 20 independent MC simulations for different values of λ. Since each simulation starts with a given fraction of dimers added during step II, we again slowly compress each system with particles constrained to their lattice positions allowing rotations and swaps before equilibrating at constant density. The total free energy for the binary mixture for which the pure system exhibits anharmonic behavior is given by:

$$\Delta F = \Delta F_{Ein} + \Delta F_{\rm I} + \Delta F_{\rm II} + \Delta F_{\rm III} \tag{5.16}$$

This formula is used to evaluate F for the WCA system, shown in Fig. 3 in the main text.

5.2.7 Supplementary Information: Finite Size Scaling For the Dodecagonal Quasicrystal

Quasicrystals cannot be explicitly simulated in finite molecular simulations, since structures with periodic boundaries inevitably repeat and are therefore technically crystalline. Some insight can be gained into properties of the quasicrystal, such as the free energy, by considering how properties change when we incrementally approach a quasicrystalline structure. This can be accomplished by constructing increasingly higher order approximants. This can be accomplished for the dodecagonal quasicrystal, for example, by the method of [206]. Fig. 5.12 shows a series of four approximants of increasing order used in this study, along with a dodecagonal quasicrystal formed by the Dzugutov system. Each subsequent inflation gives a structure closer to a quasicrystal. That is, each higher order approximant more closely approximates 12-fold rotational symmetry, as shown by the diffraction patterns.

In our studies of the lattice energy and free energy for different structures, we note how the quantities change as we increase the approximant order. All of the quantities that we study tend to converge to very similar values for the sigma structure and the two higher order approximants. We label these values the "dod" phase in the text, because we estimate that their values are close to what we would obtain for an infinite square-triangle quasicrystal. Of course, quasicrystals can have many different types of local isomorphism (LI) groups, and

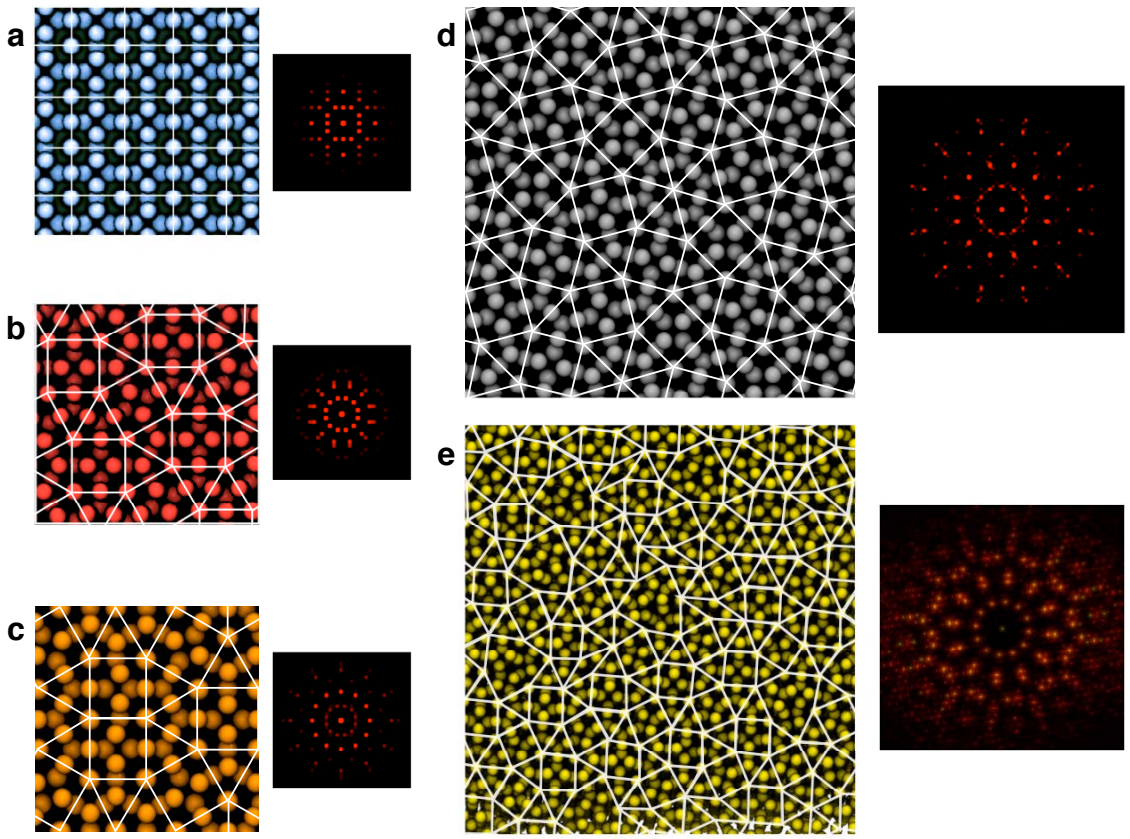


Figure 5.12 Dodecagonal quasicrystal approximants of increasing order, with associated diffraction patterns shown to the right. (a) A15 structure, (b) σ structure, (c) a higher order approximant, (d) an increasingly higher order approximant, and (e) a dodecagonal quasicrystal structure formed by the Dzugutov potential.

the quasicrystal-like structures obtained from assembly simulation contain different tiles, such as rhombs, shields and asymmetric hexagons, as shown in Fig. 5.12e. The investigation of such structures is beyond the scope of this study, but future studies could implement a similar methods as those described here to estimate their properties.

5.2.8 Supplementary Information: Stabilization of FK Phases in the Sphere-Dimer Mixture

For our primary analysis, we chose a particular system and set of conditions for studying the sphere/dimer binary mixture. We used the WCA system with $T = 0.25$, $\rho = 0.9$, and dimer size 1.5σ, where monomers making up the dimer are separated by $r_{sep} = 0.5\sigma$. These choices are largely arbitrary; $T = 0.25$ and $\rho = 0.9$ provide a reasonable statepoint where

the structures are stable for all dimer fractions $0 < f_{dimer} < 0.5$, and $r_{sep} = 0.5\sigma$ is chosen to make the model intuitive; other parameter choices and systems can also be studied. In this section, we study the effect of changing parameters on the results of our analysis, and see that the main result that we obtain, that adding aspherical particles to systems of spheres greatly enhances the stability of FK phases, is highly general.

Effect of Particle Interactions

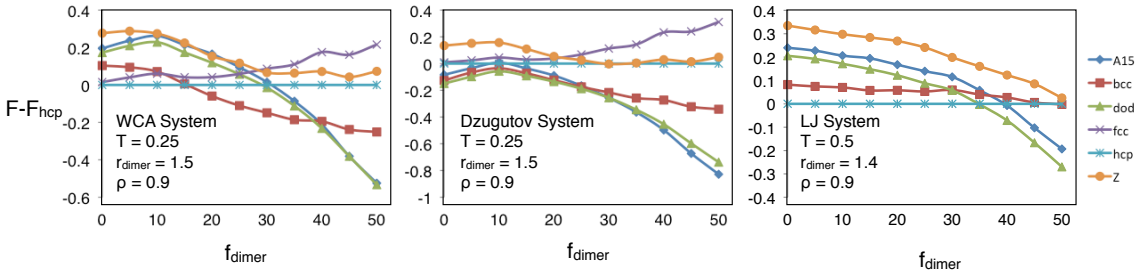


Figure 5.13 Effect of particle interactions. The graphs show the Helmholtz free energy F compared the hcp crystal for three systems with three different sets of particle interactions. Notice that the addition of dimers stabilizes FK structures relative to the fcc,hcp and bcc structures for all of the systems.

The first parameter that we consider is the effect of the pair potential used to model the particle interactions. In our representative test system from the main text, we used the WCA soft-sphere potential, and showed that adding dimers enhances the stability of FK structures relative to the more usual fcc, hcp, and bcc crystal structures. Fig. 5.13 shows that we obtain similar results for two other potentials: the Dzugutov system, for which FK phases are already stable without dimers, and the LJ system, for which FK phases have never been observed in the monatomic system.

These results support out conclusion in the main text that the mechanism by which dimers stabilize FK phases over ordinary crystals is based on packing and excluded volume interactions, since our results are not impacted by longer-range differences in the interaction potential. Short range differences, such as the hardness or softness of the excluded volume component of the potential, may indeed play a role. For the three potentials shown, the excluded volume term is equal to or similar to a r^{-12} repulsive interaction.

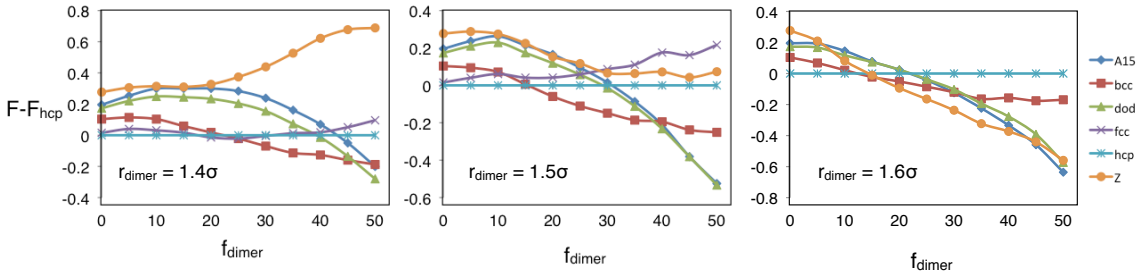


Figure 5.14 Effect of dimer size. The graphs show the Helmholtz free energy F compared the hcp crystal for the WCA system at $T = 0.25$, $\rho = 0.9$ with three different dimer sizes: 1.4σ, 1.5σ, 1.6σ. Notice that although all dimer sizes stabilize FK structures relative to ordinary crystals, the size of the dimer does effect the relative stability between different FK structures.

Effect of Dimer Size

Another parameter that we consider is the effect of the aspect ratio of the dimers. In our representative test system from the main text, we use a dimer-to-sphere aspect ratio of $1.5:1$. Thus, the dimer length r_{dimer} is 1.5σ. Fig. 5.14 shows the Helmholtz free energy F as a function of the dimer fraction for three different values of the dimer size r_{dimer} for the WCA system at $T = 0.25$, $\rho = 0.9$. We observe that all dimer sizes stabilize FK structures relative to ordinary crystals; however, the larger dimers tend to favor the Z structure and, to a lesser extent the A15 structure.

The stabilization of the Z structure for larger dimers can be explained by the structural composition of the different FK crystals shown in Fig. 3 of the main text. For larger dimer aspect ratios, the dimers have a strong preference for the larger Z15 coordination polyhedra, which are only present in systems with triangle tiles, such as the Z and σ structures. For smaller aspect ratios, dimers don't have a strong preference between Z14 and Z15 polyhedra, and the relative energy of square and triangle tiles depends on other factors. There is not an obvious explanation for why the A15 structure increases slightly in stability relative to the σ structure for larger dimer sizes, since the σ structure contains more triangle tiles, and therefore more Z15 polyhedra. The answer may require an analysis of second-order packing effects of the dimers, spanning over multiple neighbor shells and is beyond the scope of our study. Our results imply that it may be possible to control the formation of different FK structures by adding dimers of different sizes in particular ratios. We leave this topic for future investigation.

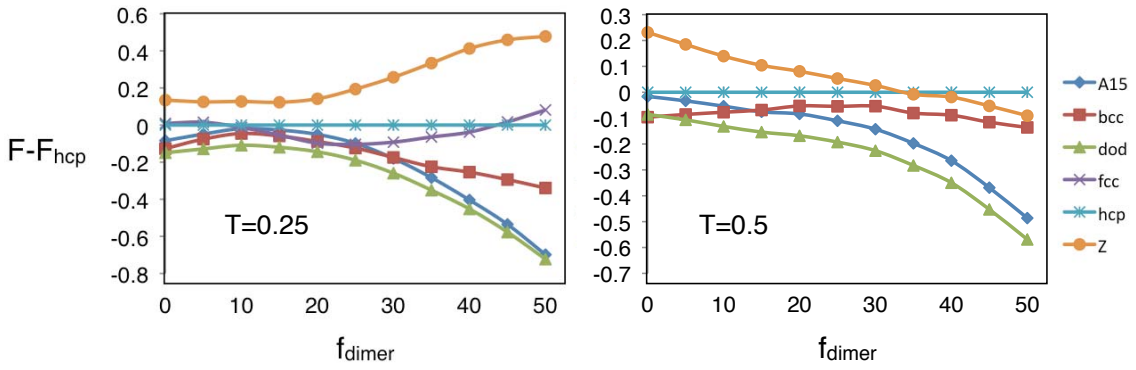


Figure 5.15 Effect of temperature. The graphs show the Helmholtz free energy F compared the hcp crystal for the Dzugutov system for $\rho = 0.9$ and $r_{dimer} = 1.4$ at a low T, $T = 0.25$ and a moderate T, $T = 0.5$. Notice that the σ and Z structures become increasingly stable at higher T, indicating that they are more entropically favorable than the A15 structure.

Effect of Changing Temperature

Another parameter that we consider is the effect of the temperature. Fig. 5.15 shows the free energy F as a function of the dimer fraction for the Dzugutov system at $\rho = 0.9$ for two different temperatures. The plots shows similar trends at both temperatures, but the σ and Z structures become relatively more stable than the bcc and A15 structures at higher T.

Our results indicate that the σ and Z structures may be more entropically favorable than the A15 structure, which may be relatively favorable energetically. This can again be explained by the structural composition of the different FK crystals shown in Fig. 3 of the main text. For the σ and Z structures with large Z15 coordination polyhedra, dimers may have more room to sample configuration space thus increasing the entropy relative to the A15 structure.

Dzugutov Stability Range

In our study of the model micelle system, the mobile surface particles on the spherical micelles lower the stability gap between FK phases and the more usual fcc,hcp and bcc crystals. However, this effect is not sufficient to stabilize FK structures outright for the statepoints that we considered. We can create a similar situation for the Dzugutov system by increasing the density above the point at which FK structures are typically stable. Fig. 5.16a shows the potential energy U as a function of density for the Dzugutov system without dimers. In terms of energy, the ideal density for FK structures is around $\rho = 0.7 - 0.8$. For higher densities, the bcc structure becomes more stable, followed by close-packed

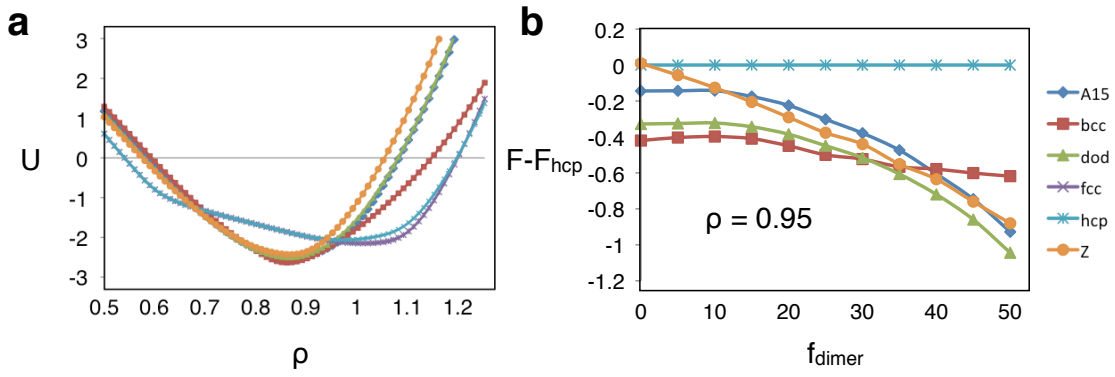


Figure 5.16 Extending the FK stability range by adding dimers for the Dzugutov system. (a) The potential energy as a function of density for the Dzugutov system without dimers for different crystal structures. (b) The Helmholtz free energy F compared the hcp crystal for the Dzugutov system at $T = 0.5$, $\rho = 0.95$ and $r_{dimer} = 1.5$ as a function of dimer fraction. Notice that the dimers stabilize FK structures, even though they increase the packing fraction, which, as shown in part (a), tends to destabilize FK structures if no dimers are present.

fcc/hcp structures at even higher densities. Fig. 5.16b shows the Helmholtz free energy F compared the hcp crystal for the Dzugutov system at $T = 0.5$, $\rho = 0.95$ and $r_{dimer} = 1.5$ as a function of dimer fraction. We observe that the dimers stabilize FK structures, even though they increase the packing fraction, which, as shown in Fig. 5.16a, tends to destabilize FK structures if no dimers are present.

Chapter 6
Forming Amorphous Solids

The glass transition has been studied for over 50 years; however, exciting new findings regarding glassy supercooled liquids are presented each year. In this chapter, we present two case studies that contribute to the vast body of research on this topic. In the first study, [A.S. Keys, A.R. Abate, S.C. Glotzer, and D.J. Durian, "Measurement of Growing Dynamical Length Scales and Prediction of the Jamming Transition in a Granular Material," *Nature Physics* (2007)] [31], we investigate the interesting connection between SHD in supercooled liquids, and heterogeneous dynamics observed in dense granular materials near the onset of "jamming." Our study demonstrates that in both a qualitative and quantitative sense, particle dynamics in granular systems are indistinguishable from particle dynamics in supercooled liquids approaching a glass transition. A perspective on this study is given in reference [211]. Our second study, [A.S. Keys, L.O. Hedges, J.P. Garrahan, S.C. Glotzer, and D. Chandler, "Structure of Localized Excitations and Relaxation in Supercooled Glass-Forming Liquids" (2011)] [36] (pending publication), addresses the problem of extracting the elementary particle motions from atomistic systems. In doing so, we find a detailed physical connection between atomistic systems and the theory of kinetically constrained models (KCMs) (see section 2.4.2). We find that theoretical arguments developed by studying (KCMs) can be accurately applied to describe the particle dynamics and bulk relaxation properties of fully-atomistic model glass formers. We note that this study is still evolving, and the work that we present in section 6.2 is not yet finalized.

6.1 Measurement of Growing Dynamical Lengthscales and Prediction of the Jamming Transition in a Granular Material

Supercooled liquids and dense colloids exhibit anomalous behaviour known as "spatially heterogeneous dynamics" (SHD), which becomes increasingly pronounced with approach

to the glass transition [67, 212, 64]. Recently, SHD has been observed in confined granular packings under slow shear near the onset of jamming, bolstering speculation that the two transitions are related [213, 66, 65]. Here, we report measurements of SHD in a system of air-driven granular beads, as a function of both density and effective temperature. On approach to jamming, the dynamics become progressively slower and more spatially heterogeneous. The rapid growth of dynamical time and length scales characterizing the heterogeneities can be described both by mode-coupling theory [92] and the Vogel-Tammann-Fulcher (VTF) equation [79], in analogy with glass-forming liquids. The value of the control variable at the VTF transition coincides with point-J [214, 215] the random close-packed jamming density at which all motion ceases, indicating analogy with a zero temperature ideal glass transition. Our findings demonstrate further universality of the jamming concept and provide a significant step forward in the quest for a unified theory of "jamming" in disparate systems.

6.1.1 Introduction and Motivation

At low temperature, high density, and low driving, the constituent particles in supercooled liquids [67], dense colloids [212, 64], and granular packings [213, 66, 65], respectively, are nearly locked into a single disordered configuration in which motion is heterogeneous in space and time. Dynamics in these systems may be governed by proximity to a generic "jamming transition [216]", beyond which rearrangements cease and the viscosity diverges. Key features of SHD on approach to the transition include unusual correlations [167] in which particles move in one-dimensional paths ("strings [73]") that aggregate into clusters [74], and dynamical correlations as measured by a dynamic four-point susceptibility χ_4 [75, 76, 174]. Clusters of strings arise naturally in dynamic facilitation [105, 106] theory and the random first-order theory of glasses [101]; their shape reflects the fractal nature of dynamical motion in these systems [217]. Strings are also a crucial ingredient in a recent theory of liquid dynamics near the glass transition [102].

Recent studies demonstrate that close-packed granular systems under slow shear exhibit SHD as well [213, 66, 65], bolstering speculation that liquids and granular matter share dynamical similarities on approach to the jammed state. However, the universality of the jamming hypothesis has not yet been tested in terms of variation in the hallmark dynamical heterogeneities as a function of the control parameter. Here, we present the first simultaneous measurements in any experimental system of the growth of the cluster correlation length, string length, four-point correlation length, and their characteristic timescales by varying the control parameter. We show that the SHD observed in a far-from-equilibrium, athermal system of air-fluidized granular beads is essentially indistinguishable from that observed in

thermal systems like supercooled liquids and dense colloidal suspensions. Moreover, we show that theoretical models developed for the glass transition can be used to describe our granular system, and predict a mode-coupling like transition and, more importantly, the jamming transition packing fraction, known as point-J [214, 215] from quantities characterizing SHD.

6.1.2 Experimental Setup

We characterize the spatiotemporally heterogeneous nature of dynamics in an athermal, far-from-equilibrium system of air-driven steel spheres on approach to jamming. Compared to sheared or shaken granular systems, in which energy is injected at the boundaries, in air-driven systems the energy input is uniform in space and time. Our granular system consists of a 1:1 bidisperse mixture of steel beads of diameters $d_s = 0.318$ cm and $d_l = 0.397$ cm, with respective masses of 0.130 gm and 0.266 gm, confined to a circular region of diameter 17.7 cm. The packing density is varied from an area fraction of $\phi = 0.597$ to $\phi = 0.773$ by changing the total number of beads from 1470 to 1904. Bead motion is restricted to rolling within a horizontal plane, and is excited by an upward flow of air at a fixed superficial flow speed of 545 cm/s. Bead positions are identified by reflecting light from their chrome surface to a camera three feet above. The duration of experimental runs is 20 minutes. By contrast with the molecules in a supercooled liquid, here the particles are macroscopic objects driven at random by a continuous input of energy. Consequently the speed distributions are non-Maxwellian, and the average kinetic energies of the two bead species are unequal. Nevertheless, as reported previously [218], the system mimics a simple liquid for low ϕ and exhibits tell-tale changes in the average structure and dynamics at increasing packing densities.

6.1.3 Spatially Heterogeneous Dynamics

Dynamical characteristics for an example case, $\phi = 0.773$, are displayed in Fig. 6.1a-d. The mean-squared displacement versus time interval (delay time), Fig. 6.1a, averaged over all start times and all beads, is perhaps the most familiar quantity. It shows ballistic motion at short times, diffusive motion at late times, and a plateau of sub-diffusive motion at intermediate times that is indicative of caging. While informative, the mean-squared displacement cannot distinguish uniform from heterogeneous dynamics. For this, we perform three measurements developed for supercooled liquids. The first involves clusters of beads that are "mobile," i.e. which have displacements ranking among the top 10% of

all bead displacements in a given delay time [166] (in this system, 10% gives the largest distinction between mobile and immobile beads at all packing densities). The configuration displayed in Fig. 6.1e, where beads are colour-coded according to mobility, demonstrates that the mobile beads are not distributed at random; rather, mobile and immobile beads are clustered and spatially separated, indicating spatially heterogeneous dynamics. One measure of SHD is thus the average size of mobile clusters, $S_c(t)$, defined as the average number of neighbouring mobile beads for a given time interval t. The motion within a mobile cluster, shown by the displacement vectors in Fig. 6.1e, tends to be correlated into quasi-1d paths called "strings [73]." Thus a second measure of SHD is the average string length, $L_s(t)$, defined in terms of the number of beads that, within a time interval t, replace the initial positions of neighbouring beads to within a tolerance of $0.3d_s$ [73]. Yet a third measure of SHD may be constructed from a four-point susceptibility $\chi_4(t)$, which measures the extent to which the dynamics at any two points in space are correlated within a time interval [75]. The self-contribution $\chi_4^{SS}(t)$ dominates the general result [76] and is computed from the variance of the self overlap order parameter $q_S(t)$, which decays from one to zero:

$$\chi_4^{SS}(t) = N\left[\left\langle q_S(t)^2\right\rangle - \left\langle q_S(t)^2\right\rangle\right]. \tag{6.1}$$

Here, N is the total number of beads, and $q_S(t)$ is defined as:

$$q_S(t) = \frac{1}{N}\sum_{i=1}^{N} w\left(|\mathbf{r}_i(t) - \mathbf{r}_i(0)|\right). \tag{6.2}$$

The overlap parameter w is defined by:

$$w = \begin{cases} 1 & |\mathbf{r}_i(t) - \mathbf{r}_i(0)| < 0.5d_s \\ 0 & |\mathbf{r}_i(t) - \mathbf{r}_i(0)| \geq 0.5d_s. \end{cases} \tag{6.3}$$

Here, $\mathbf{r}_i(t)$ is the position of bead i at time t; averages are taken over all beads and over all start times.

The example results in Fig. 6.1b-d for the cluster size $S_c(t)$, the string length $L_s(t)$, and the four-point susceptibility $\chi_{ss}^4(t)$, all exhibit well-defined peaks as a function of time interval, as found in glass-forming liquids. The locations of the peaks indicate the time interval over which the dynamics are most heterogeneous, and the heights of the peaks indicate the spatial extent or "strength" of the heterogeneities. As with glass-forming liquids [74, 170, 71, 72] and colloids [64], the cluster size and string length are largest at the crossover between caged and diffusive motion, while $\chi_{ss}^4(t)$ (and $\chi^4(t)$) peaks later, in the

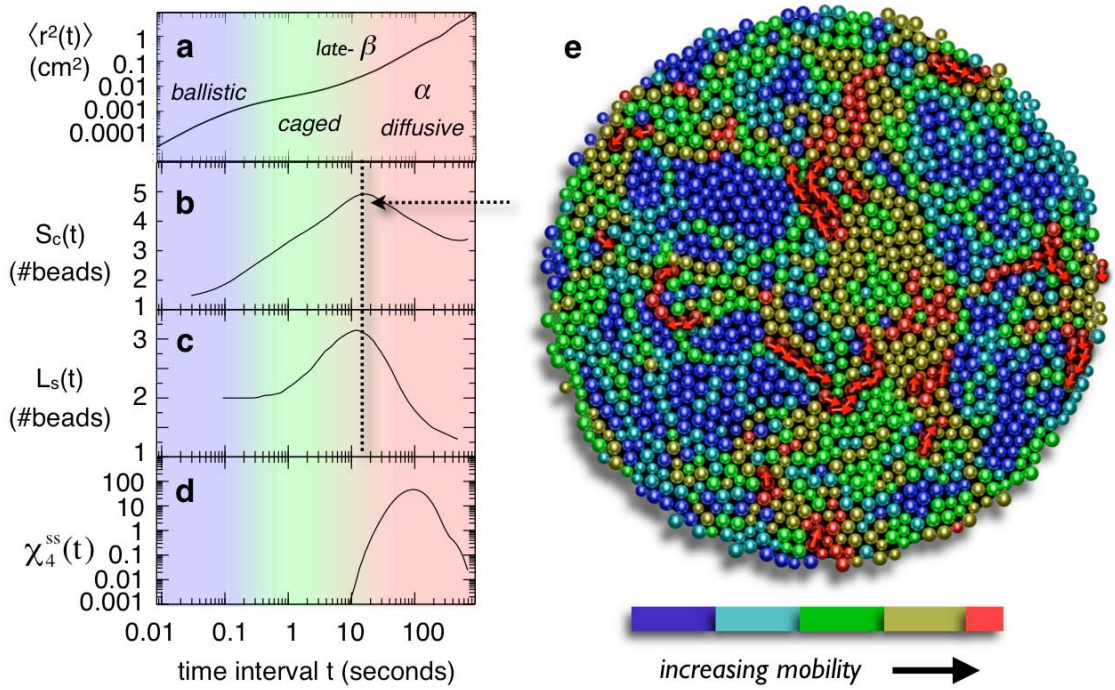


Figure 6.1 (a) Mean square displacement. (b) Number average mobile bead cluster size $S_c(t)$. c) Number average string length $L_s(t)$. (d) Self contribution to the four point susceptibility $\chi^4_{ss(t)}$. (e) An instantaneous bead configuration where the colouring of beads indicates the mobility over a time interval of 12s (the timescale for both maximum cluster size and string length). The 10% most mobile beads are red; note that they form clusters. Beads moving in strings have vectors superimposed to indicate their directional motion. Note that the dynamics are spatially heterogeneous.

so-called alpha or structural relaxation regime [76, 69]. The athermal air-fluidized beads therefore exhibit spatially heterogeneous dynamics that is identical to thermal glass-forming systems with respect to these three measures.

6.1.4 Dynamics as a Function of the Control Parameter

Now that spatially heterogeneous dynamics are established for gas-fluidized beads, we turn to their variation as a function of control parameter. The distribution of cluster sizes at the peak time interval, shown in Fig. 6.2a for three different packing densities ϕ, approaches a power-law as ϕ is increased. This is consistent with the percolation of mobile bead clusters;

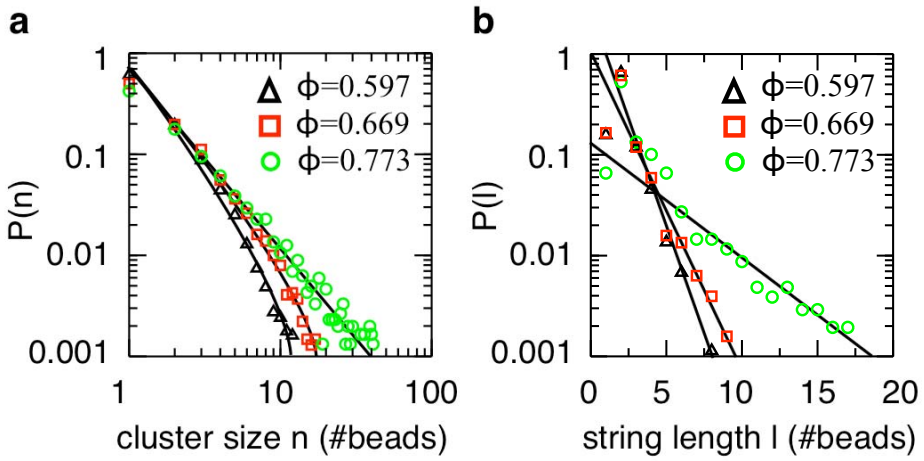


Figure 6.2 (a), (b) Distribution of (a) cluster sizes and (b) string lengths for three values of control variable ϕ. The cluster size distribution approaches a power law with increasing ϕ (solid lines indicate fit by power law multiplied by an exponential cutoff), while the string length distribution is exponential.

similar power-laws have been observed in colloids [64] and in simulations of supercooled liquids [74, 170, 71] near the mode-coupling temperature. Furthermore, the distribution of string lengths at the peak time interval, shown in Fig. 6.2b, is exponential, , at all values of ϕ , where l_0 is set by the average string length. This is consistent with behaviour reported in simulations of several supercooled liquids [73, 71, 219]. We note the average cluster size is not much larger than the average string length, although the largest clusters observed (100 particles) are substantially larger than the largest string observed (30 particles) (not shown).

Results for $S_c(t)$, $L_s(t)$, and $\chi_s^4 s(t)$ vs t are displayed in Fig.6.3a-c for a sequence of different packing densities ϕ. When beads are added to the system, the average effective temperature also decreases, resulting in a trajectory in the (ϕ ,T_{eff}) phase diagram that heads towards point-J, the zero-temperature jamming transition previously found for this system at $\phi = 0.83$, which is coincident with the packing density at which the system is random close-packed. As the motion becomes more restricted, the peaks in all three measures of SHD grow and move to later times. Therefore, the dynamics not only slow down but also become more heterogeneous on approach to point-J. Since the SHD functions have approximately the same shape when viewed on a log-log plot (see data collapse in insets of Fig. 6.3a,b), this behaviour is fully characterized by the ϕ-dependence of the characteristic or peak time scales $\{t_{Sc}^*, t_{Ls}^*$ and $t_{\chi 4}^*\}$ and length scales $\{\zeta_{Sc}(t_{Sc}^*), \zeta_{Ls}(t_{Ls}^*)$ and $\zeta_{\chi 4}(t_{\chi 4}^*)\}$. The length $\zeta_{Ls}(t_{Ls}^*) \propto L_s$ is a correlation length for stringlike motion, $\zeta_{Sc}(t_{Sc}^*) \propto S_c$ is a correlation length of mobile particle clusters, and is a correlation length [65] of clusters of caged

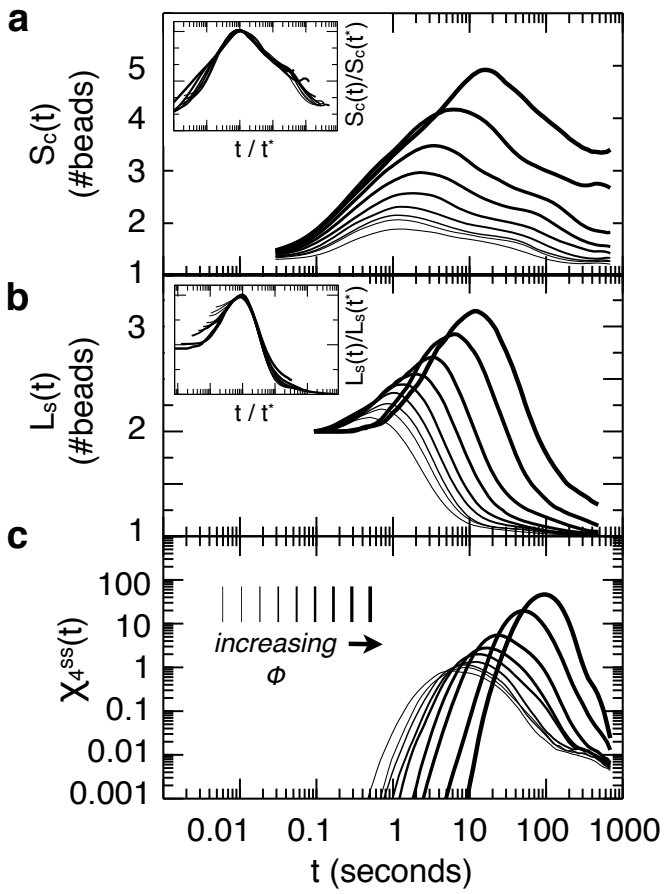


Figure 6.3 (a) Cluster size, (b) string length, and (c) self contribution to the four-point susceptibility, all as a function of time interval t, for a sequence of area fractions ϕ. In order of increasing peak height: $\phi = 0.597, 0.633, 0.647, 0.669, 0.693, 0.722, 0.742, 0.762, 0.773$. Insets in (a) and (b) show collapse of all data sets upon scaling by the peak heights and peak times.

particles. On approach to point-J, both the characteristic times and the correlation lengths appear from Fig. 6.3 to grow without bound. This is reminiscent of behavior for supercooled liquids as temperature is lowered. Though very different, both types of systems appear to approach an unusual critical point where the growing length scale is purely dynamical, such that there is no macroscopic change in instantaneous structure [67, 220, 221, 222].

6.1.5 Analogies with the Supercooled Liquids

To further quantify this analogy, the growth of the characteristic timescales and dynamical length scales is shown in Fig. 6.4a-d as a function of packing density. Motivated by recent studies [171, 94, 95] predicting a power law divergence of dynamical lengthscales from

mode coupling theory [92] (MCT), as well as earlier applications of MCT to liquids and colloids, we fit all data to a power-law of the form $1/|\phi/\phi_c|^x$, where both ϕ_c and x are adjustable parameters. As seen in Fig. 6.4a,b, excellent fits are obtained to all data for a single value $\phi_C = 0.79 \pm 0.02$. This value of ϕ lies well above the onset of caging and is less than the jamming packing fraction, in analogy with well-established findings that the mode-coupling temperature is below the caging transition but above the glass transition temperature [92, 93], and demonstrates for the first time a mode-coupling-like transition in a granular system. In addition to MCT, the glass transition can also be well-described by the Vogel-Tammann-Fulcher (VTF) equation [79]; therefore, we also fit the characteristic time and length scales to the form $\exp(E/|\phi/\phi_0 - 1|)$, where E and ϕ_0 are adjustable parameters. As seen in Fig. 6.4c,d, excellent fits are obtained to all data for a single value $\phi_0 = 0.84 \pm 0.02$, consistent with the value of random-close packing for the bead system and the value of point-J. Since random close packing is the densest random packing possible and the point at which all motion ceases, a VTF packing fraction of ϕ_{RCP} is analogous to an effectively zero-temperature ideal glass transition, consistent with the definition of point-J. This is the first prediction of point-J in a granular system from analysis of spatially heterogeneous dynamics.

6.1.6 Results and Conclusions

Our study implies that the behaviour of jammed systems, both thermal and athermal alike, may be understood using the theoretical tools developed for liquids. This, in turn, highlights the importance of packing in the underlying physics of the glass transition and jamming. Our results open the door to future theoretical insight into the relationship between granular materials and supercooled liquids, which might be united by a unified theory of jamming.

6.2 Structure of Localized Excitations and Relaxation in Supercooled Glass-Forming Liquids

For several atomistic models of glass formers, at conditions below the glassy dynamics onset temperatures, T_o, we use importance sampling of trajectory space to study the structure, statistics and dynamics of localized excitations, all of which we relate to dynamic heterogeneity. We define excitations in terms of irreversible particle displacements. At super-cooled conditions, we find that these excitations are associated with correlated particle motions that are sparse and localized, with an average radius that is temperature independent

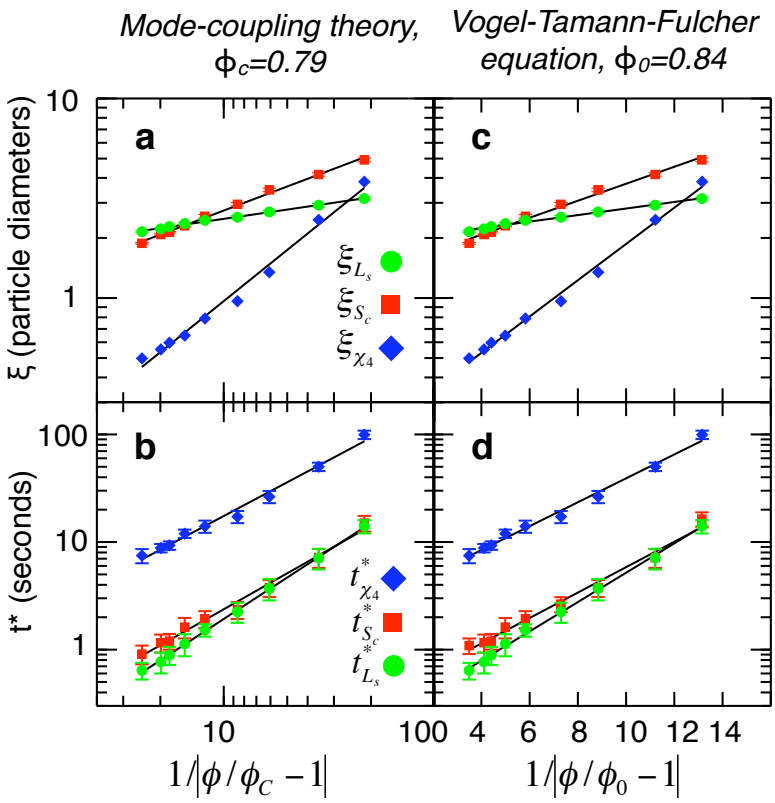


Figure 6.4 (a) Log-log plot of dynamical correlation lengths versus $1/|\phi/\phi_c - 1|$ fit with $x = 0.40$ for ζ_{Sc}, $x = 0.15$ for ζ_{Ls} and x=1.70 for $\zeta_{\chi 4}$. (b) Log-log plot of characteristic times t^* versus $1/|\phi/\phi_c - 1|$ with $x = 1.3$ for t^*_{Sc}, $x = 1.1$ for t^*_{Ls}, and $x = 1.0$ for $t^*_{\chi 4}$. (c) Semilog plot of dynamical correlation lengths versus $1/|\phi/\phi_0 - 1|$ fit with $E = 0.10$ for ζ_{Sc}, $E = 0.03$ for ζ_{Ls}, and $E = 0.42$ for $\zeta_{\chi 4}$. (d) Semilog plot of all t^*s vs $1/|\phi/\phi_0 - 1|$ fit with E=0.27. See text for fit expressions. Error bars for (a), (c) represent the standard error of the measurement. Error bars for (b), (d) represent the uncertainty in determining the time at which the function is maximum. .

and no larger than a few molecular diameters. As a function of temperature, $T < T_o$, the equilibrium concentration of these excitations is proportional to $\exp[-J_a(1/T - 1/T_o)]$, where a is the lengthscale of the irreversible displacement. The energy scale J_a grows logarithmically with a, in a way that is consistent with a scenario of hierarchical dynamics. In the low-temperature regime, excitation dynamics is facilitated by the presence of other excitations, causing dynamics to slow in a hierarchical way as temperature is lowered. The resulting mean equilibrium relaxation time is proportional to $\exp[J^2(1/T - 1/T_o)^2]$. The quantities J and T_o are material properties that depend upon molecular density

6.2.1 Introduction

Supercooled glass-forming liquids exhibit an assortment of behaviors indicative of highly correlated and complex microscopic dynamics [223, 224, 225]. These include spatial heterogeneity and super-Arrhenius growth of relaxation times with lowering temperature. Theoretical treatments fall largely into two classes. In one, originating with Adam and Gibbs [96], a super-cooled liquid is a mosaic of structured domains, and relaxation follows from reorganizing these domains [226]. In the other, the material contains a gas of point-like defects [103], and relaxation follows from hierarchical dynamics [227], where defects facilitate [228] the creation and destruction of neighboring defects. In the former case, dynamics slows because mosaic domains grow, while in the latter, dynamics slows because defect separations grow. While there is significant support for both perspectives, the most basic underlying features – the mosaic in the former and localized excitations in the latter – have yet to be demonstrated. Here, we use computer simulation and importance sampling to address this shortcoming in our understanding.

We have considered the molecular dynamics of three different atomistic models. For each, we ask the question: How does an atom move between distinct neighboring positions? We are able to answer this question with the method of transition path sampling [60]. In particular, molecular reorganization in a super-cooled liquid is a rare event, and transition path sampling harvests ensembles of trajectories exhibiting such events. It is an importance sampling that preserves an equilibrium distribution of trajectories, and it is free of preconceived notions of mechanisms. Applying this technique, we find three important results: (1) Irreversible particle movements – what we call "excitations" – are associated with the correlated displacements of only a handful of neighboring particles. These displacements are closely related to the micro-strings [229] discovered in earlier computer modeling studies of dynamic heterogeneity. (2) At super-cooled conditions, i.e., at temperatures below that of the onset temperature, $T_{\rm o}$, excitations are sparse. They arise in localized regions of relatively high mobility whose size is largely independent of temperature, and whose spatial distribution is that of an ideal gas. (3) For a given displacement length a, the concentration of excitations, c_a, has a Boltzmann temperature dependence

$$c_a \propto \exp\left[-J_a\left(\frac{1}{T}-\frac{1}{T_{\rm o}}\right)\right], \quad T < T_{\rm o} \tag{6.4}$$

where J_a, like $T_{\rm o}$, is a material property depending upon the model and the density of the model. The energy scale J_a varies logarithmically with a, in a way that is consistent with a scenario of hierarchical dynamics [227], where excitations on smaller lengthscales and

timescales facilitate excitations on larger lengthscales and timescales. For a given material, these constants J_a and T_o are directly related to those used to fit super-Arrhenius relaxation times according to a parabolic law [107]

$$\frac{\tau}{\tau_o} \approx \exp\left[J^2\left(\frac{1}{T}-\frac{1}{T_o}\right)^2\right], \quad T < T_o \tag{6.5}$$

where τ_o is the relaxation time at the onset temperature.

These results support the picture of facilitating localized defects or excitations [230]. In addition, by providing a microscopic recipe for computing c and thereby predicting τ, the results demonstrate a microscopic procedure for deriving the local excitation picture from atomistic models. These findings do not exclude the possibility of a correct mosaic model, but a link between mosaics and localized excitations is currently unknown. Our results show that the principal growing length associated with growing time scales is the mean distance between excitations, $l = c^{-1/d}$, where d is the system dimensionality. Relaxation slows as the scarcity of excitations grows. A connection between the local excitations and mosaics would need to relate l to the size of mosaic elements.

6.2.2 Definition of Excitations

The excitations that we consider are the dynamical events associated with microscopic particle displacements maintained for a significant period of time. These displacements coincide with transitions between two adjacent basins of a potential energy landscape, and they are distinct from fleeting intra-basin vibrations. Inherent structures [231] provide one way to distinguishing the former from the latter. The inherent structure of a configuration of an N-particle system at time t is obtained by steepest descent to the nearest minimum in the potential energy landscape [151]. The inherent structure evolves as dynamics progresses, but intra-basin vibrations are not apparent in that evolution. Excitations are present when the dynamics produces significant change in particle positions in the inherent structure.

We detect these motions from the behavior of the functional of path $h_{\mathrm{Ex}}[x(t)]$, where

$$h_{\mathrm{Ex}}[x(t)] = \begin{cases} 1 & \text{if} \quad |\bar{r}_1(t)-\bar{r}_1(0)| \approx 0, \quad 0 < t < t_s \\ & \text{and} \quad |\bar{r}_1(t)-\bar{r}_1(0)| \geq a, \quad \Delta t + t_s < t < \Delta t + 2t_s \\ 0 & \text{otherwise.} \end{cases} \tag{6.6}$$

Here, $x(t) = [r_1(t), r_2(t), \ldots, r_N(t)]$ refers to the net configuration of the system at time t,

the corresponding inherent structure is $\bar{x} = [\bar{r}_1(t), \bar{r}_2(t), \ldots, \bar{r}_N(t)]$, and $\bar{r}_1(t)$ refers to the position of a tagged particle, particle 1, in the inherent structure of $x(t)$ Employing $h_{\rm Ex}[x(t)]$ presumes that a trajectory has run for at least the time frame from $t = 0$ to $t = \Delta t + 2t_s$. The functional is 1 when an excitation appears during this time, and it is zero otherwise. An excitation is thereby characterized by three periods: During the first, $0 < t < t_s$, the inherent structure of the tagged particle remains at its initial point; following this sojourn, the system undergoes a transition that lasts up to a time Δt, after which the tagged particle is found in a distinctly different position for another relatively quiescent period.

This definition of an excitation introduces a displacement length, a, a transition time, Δt, and a sojourn time, t_s (see Fig. 6.5a). The length a is set to facilitate reasonable statistics and to ensure that excitations are associated with no more than one inter-basin crossing, and should be small enough so as to not exclude too many examples of excitations. At the same time, if a is too large trajectories leaving one basin will be captured for long periods of time in intermediate basins; separations in time scales will be obscured, and the average of $h_{\rm Ex}[x(t)]$ will fail to exhibit linear growth with time after a transient time [232]. For the systems that we consider here, taking a to be on the order of one particle diameter σ proves to be a satisfactory choice.

The transition time Δt is allowed to vary in transition path sampling [233] so that dynamics determines a distribution of these times without pre-conceived notions of how it should behave. In the parlance of rare-event dynamics, Δt is the plateau or commitment time [234, 235] – its mean being the average time scale for committing to one basin or the other when traversing the transition state separating two basins. The sojourn time t_s is set to a reasonable timescale over which the particle can be determined to be in a quiescent state. To properly enforce this criterion, this timescale must be at least of the order of the commitment time; thus a reasonable value for t_s is given by the mean commitment time $t_s \approx \langle \Delta t \rangle$. Since $\langle \Delta t \rangle$ is not typically known, *a priori*, t_s is determined by trial and error. For the systems we have studied for this paper, at super cooled conditions, the typical mean and variation of Δt is very much smaller than structural relaxation times. Therefore, over these timescales the time-coarse grained dynamics differs little from that of inherent structure dynamics,

$$\bar{x}(t) \approx \frac{1}{t_s} \int_{-\frac{1}{2}t_s}^{\frac{1}{2}t_s} dt' x(t + t'). \tag{6.7}$$

6.2.3 Transition Path Sampling

With excitations defined, we have used the methods of transition path sampling to harvest trajectories with the probability

$$P[x(t)] \propto P_0[x(t)]h_{\mathrm{Ex}}[x(t)] \tag{6.8}$$

where $P_0[x(t)]$ is the equilibrium probability for a trajectory, and $h_{\mathrm{Ex}}[x(t)]$ is the non-equilibrium factor that limits the ensemble to trajectories that exhibit an excitation. In practice, this is done by first equilibrating the system with a trajectory that runs for many structural relaxation times. A few short sections of this trajectory are then chosen to provide first examples of a transition or excitation associated with the tagged particle. Tagged particles are chosen randomly from the distribution of particles that move a distance a earliest along a given trajectory. In this way, the excitations that we sample involve particles, initially at rest, that undergo continuous displacements of length a without long-lived intermediates. A typical trajectory for a displacement length $a = \sigma$ is shown in Fig. 6.5a for a 2d WCA system [236]. From those first trajectories, a series of shooting and shifting moves are then performed [60] generating an ensemble of thousands of independent examples of excitations. Fig. 6.5b illustrates trajectories found from transition path sampling at different temperatures.

Because we do not, *a priori*, know how the timescale for excitations changes with temperature, we employ the variable-path-length shooting algorithm of reference [233]. In practice, we find that the commitment time Δt varies little with temperature. Fig. 6.6a shows a few typical examples of $P(\Delta t)$ over a temperature range $T_{\mathrm{o}} \leq T \leq T_{\mathrm{min}}$, where T_{min} is the lowest temperature that we can equilibrate in MD simulations. The distributions tend to exhibit a slightly higher proportion of shorter trajectories at lower temperatures, which implies that higher-temperature trajectories sometimes involve short-lived intermediates. The commitment time Δt is dependent on a, with longer timescales being required to sample longer particle displacements, as shown in Fig. 6.6b. When sampling excitations on different lengthscales a, the sojourn time t_s must be adjusted accordingly.

The temperature-independence of the commitment time allows us to, in some cases, sample excitations in very cold systems out of equilibrium. These systems are created from configurations of a warmer equilibrated system, but with temperatures chosen from a Maxwell-Boltzmann distribution at a significantly lower temperature. One such case is shown in the left-most panel of Fig. 6.5b. Whether modestly super-cooled or deeply super-cooled at non-equilibrium conditions, the motions associated with excitations appear to be localized. This impression can be quantified by computing the average of the displacement

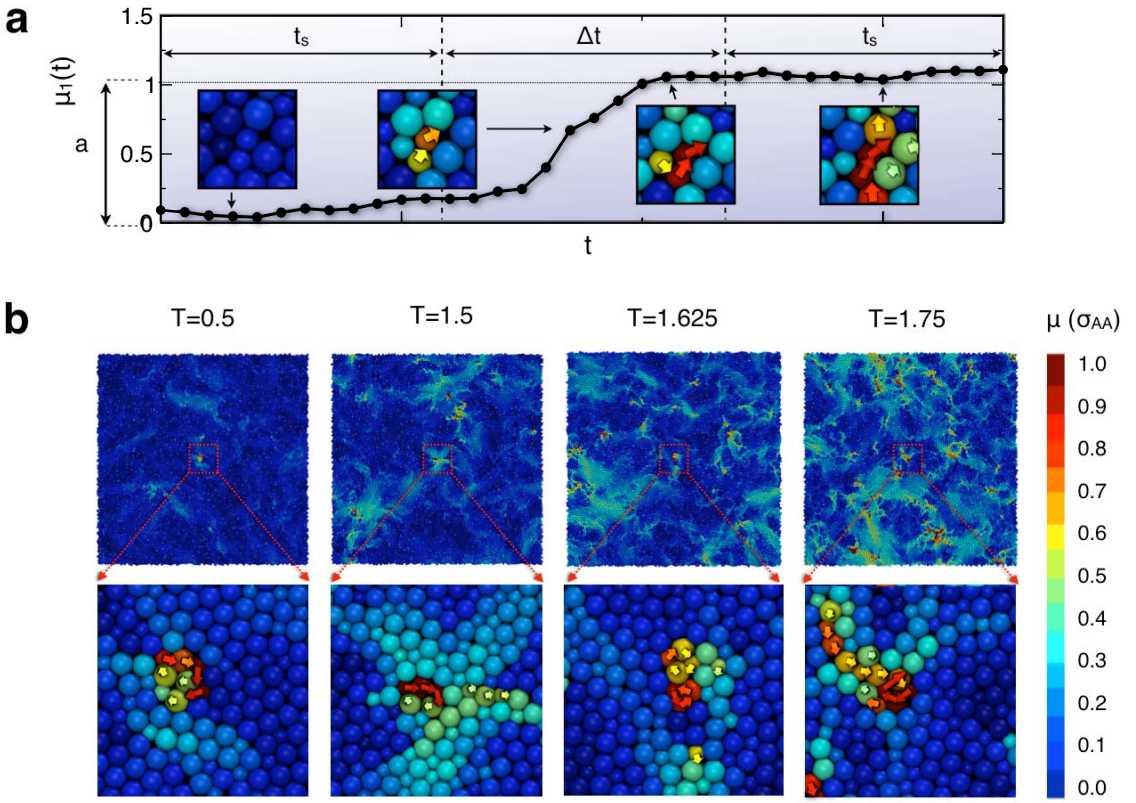


Figure 6.5 Localized excitations in a two-dimensional glass former harvested by transition path sampling. The particular system considered here is a a two-dimensional mixture of WCA potentials (WCA2D-1, defined in materials and methods), and the density $\rho = 0.75$. The onset temperature for this model at this density is $T_o = 2.6$. Temperatures at which the trajectories are harvested are indicated by the value of T. (a) The panel depicts a series of snapshots from a transition path sampling trajectory over which the tagged particle (center) undergoes a displacement that is slightly greater than 1 particle diameter. The functional of path $h_{\mathrm{Ex}}[x(t)]$ is set with $a = \sigma$. The plot of μ_1 versus r tracks the distance that the tagged particle has moved from its initial position at $t = 0$ as a function of time. The displacement is measured in terms of inherent structure coordinates. (b) The panels depict inherent structures at a time $\Delta t + 2t_s \approx 15$, with particles colored with reference to their inherent structure positions at time $t = 0$. Particles that overlap with their initial positions are blue; those that are displaced by at least one particle diameter are red. Intermediate displacements are depicted with a linear interpolation of this color code. The tagged particle, which necessarily moves at least one particle diameter, is near the center of the panel. Illustrated trajectories are representative.

field,

$$\mu(r) = \sum_{i>1} |r_i(t) - r_i(0)|\, \delta\left[r - r_i(0)\right] \tag{6.9}$$

subject to the constraint that the tagged particle, particle 1, is at the origin at time $t = 0$. The behavior of this averaged displacement field for $t = \Delta t + 2t_s$ is shown in Fig. 6.6c-f for trajectories sampled with $a = \sigma$ for three different standard 3d glass formers. For all

temperatures, the displacement field is similar to the radial distribution function of an gas, with no observable character outside of a few particle diameters from the exciting particle at the origin. The typical correlation length of about 2-5 particle diameters is the mean lengthscale of the motions associated with the excitations, which is dependent on the particle interactions and density, but not the temperature. Fig. 6.6c-f shows that the lengthscale for a given system is unchanging over the range of temperatures accessible to equilibrium molecular dynamics simulations, as well as lower temperature non-equilibrium simulations accessed via transition path sampling (Fig. 6.6f). As shown in the SI, we obtain equivalent results for displacements in the range $\sigma/4 \leq a \leq \sigma$. Our collective results give no indication of a growing dynamical lengthscale with decreasing temperature for irreversible particle displacements, in agreement with the localized excitations picture of glass-forming liquids.

6.2.4 Excitation Energy Scale

Although the timescales and cooperativity lengthscales of the excitations do not change with temperature, the probability with which they occur changes drastically. This is depicted in Fig. 6.5 and quantified in Fig. 6.6, which shows a decrease in the average of the mobility field $\mu(r)$ far from the tagged particle at the origin. The rate at which excitations disappear with decreasing T is related to the energy scale required to create excitations within the system. In glassy systems, particle rearrangements are rare thermally activated events; thus their probability p should follow an Arrhenius temperature dependence: $p_a \propto \exp(-J_a/T)$, where J_a is an activation energy for an excitation characterized by an irreversible particle displacement of length a. This physical feature of glass-forming liquids is captured by kinetically constrained models (KCMs), and is described by equation 6.4, which relates the concentration of excitations to their energy scale J.

Given an approximate timescale for particle motion on a lengthscale a, which we have already obtained from transition path sampling (see Fig 6.6e,f), the concentration exciting particles that have moved a length a, $c_a = p_a N/V$ can be measured in equilibrium MD simulations according to:

$$c_a = \left\langle \frac{1}{V} \sum_i^N h_{i,\mathrm{Ex}}[x(t)] \right\rangle. \tag{6.10}$$

The path functional $h_{i,\mathrm{Ex}}[x(t)]$ can be defined as in equation 6.8 above, where i identifies the tagged particle. In practice, for the purposes of estimating J_a we can apply a simpler

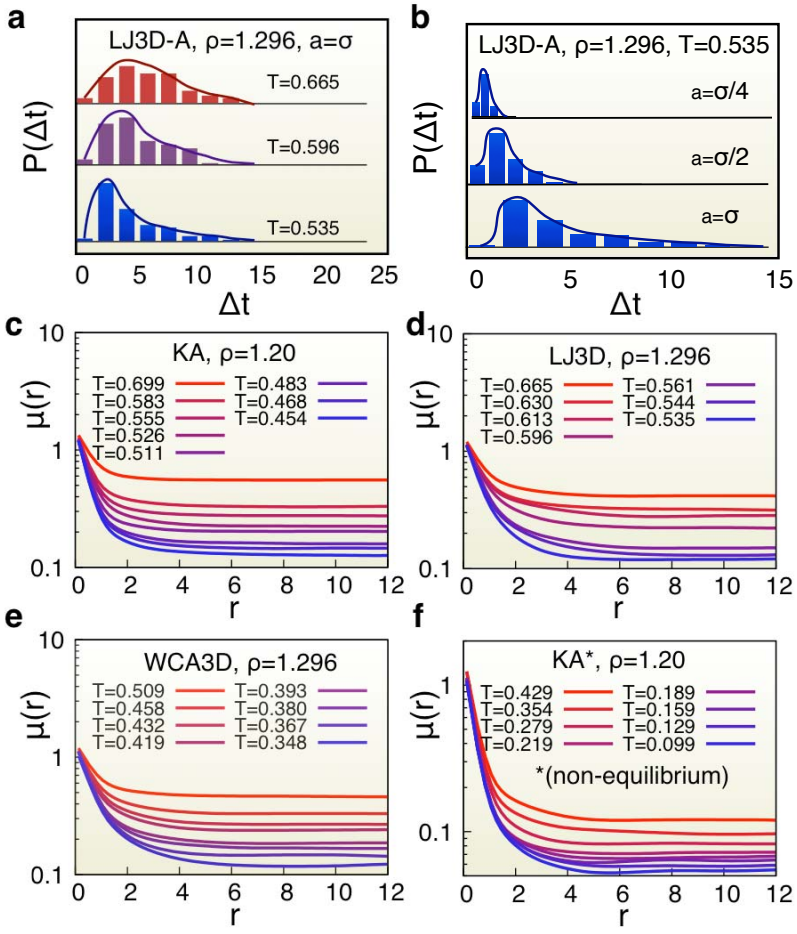


Figure 6.6 Localized excitations. (a) The probability distribution of commitment times Δt for type A particles in the LJ3D system as a function of temperature for a fixed displacement length $a = \sigma$. (f) The probability distribution of commitment times Δt for type A particles in the LJ3D system at the lowest temperature $T = 0.535$ as a function of the displacement length a. (c-f) Average displacement field, μ for $t = \Delta t + 2t_s$ at a distance r from the tagged (exciting) particle for (c), (d), (e), the KA, LJ3D and WCA3D systems under equilibrium conditions. (f) The average displacement field a distance r from the tagged (exciting) particle at very cold non-equilibrium conditions for the KA system.

functional that does not explicitly enforce irreversibility:

$$h_{i,\mathrm{a}}[x(t)] = \begin{cases} 1 & \text{and} \quad |\bar{r}_i(t) - \bar{r}_i(0)| \geq a, \quad t = t_a \\ 0 & \text{otherwise.} \end{cases} \tag{6.11}$$

Although the value of c_a is sensitive to the choice of $h[x(t)]$, the value of J_a (which depends on the relative rate at which c_a changes with temperature) is not noticeably affected. The advantage of employing $h_{\mathrm{a}}[x(t)]$ is that we obtain very similar results for J_a and T_{o} from pairs

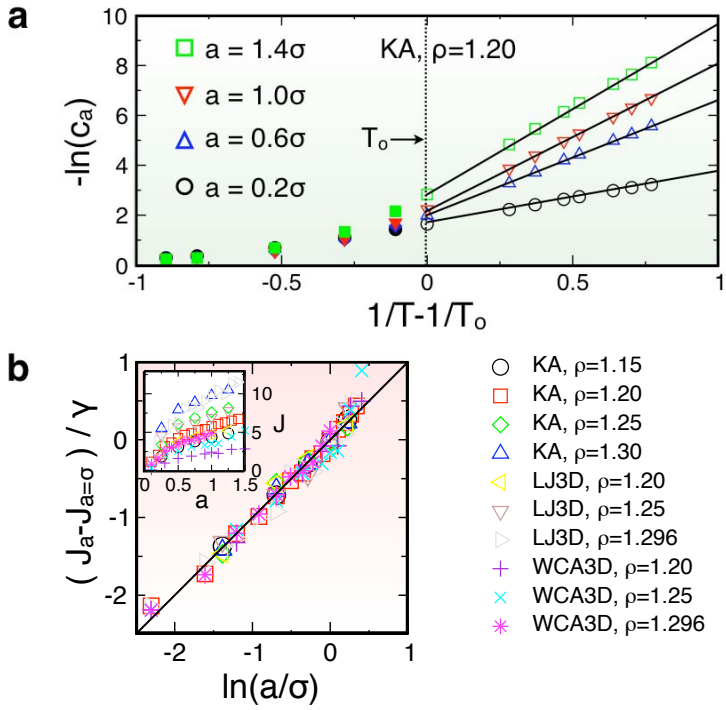


Figure 6.7 Excitation energy scale. (a) Concentration of exciting particles c that move a distance a for the KA system. Below an onset temperature T_o, the concentration obeys Boltzmann statistics (equation 6.4). (b) Data collapse demonstrating that J varies logarithmically with the lengthscale for particle motion; $J_a = \gamma \ln(a)$, where γ is a constant that depends on the model. Fit parameters are given in table 6.1.

.

Table 6.1 Fitting parameters for data collapse in Fig. 6.7b.

Model	ρ	γ	$J_{a=\sigma}$	1-R2
KA	1.15	1.85	4.4	8e-4
	1.2	2.16	5.8	5e-3
	1.25	2.90	7.7	9e-3
	1.3	3.04	9.8	2e-3
LJ3D	1.2	1.92	4.8	1e-2
	1.25	1.87	7.0	2e-2
	1.296	3.68	9.7	1e-2
WCA3D	1.2	1.01	2.3	9e-3
	1.25	1.58	3.8	3e-2
	1.296	1.73	4.6	5e-3

of snapshots t_a apart, rather than the detailed trajectories required for path sampling (see SI for further details). This type of calculation may be more applicable to experimental systems, such as granular matter or colloids. The timescale t_a should be greater than the typical

commitment time Δt to allow exciting particles sufficient time to move, but not so long that particles can excite multiple times. We find that t_a on the order of $t_a \sim 1-3\max(\Delta t)$ is typically sufficient, where $\max(\Delta t)$ is the maximum of the $P(\Delta t)$ distribution obtained from path sampling. In practice, we find that any timescale $t_a \sim O[\max(\Delta t)]$ gives similar results for J_a, when computed according to equation 6.4. Plots of c_a versus the reduced inverse temperature for several different displacement lengths a are shown in Fig. 6.7a for the KA binary LJ mixture [133]. Below an onset temperature $T_{o,a}$, the curves exhibit a nearly perfect Boltzmann temperature dependence. Above $T_{o,a}$ the Boltzmann dependence breaks down as the separation of timescales (between the excitation timescales and structural relaxation times) disappears, causing some particles to participate in multiple excitations on the timescale t_a. The value of $T_{o,a}$ is independent of the lengthscale for particle motion a for the range of a studied.

The energy scale J_a is given by the slope of the lines in Fig. 6.7a. We see that the energy scale is not constant, but rather grows with the displacement length a. This is reminiscent of the picture of hierarchical dynamics proposed by Palmer *et al.* [227], where excitations on smaller timescales and lengthscales give rise to higher-energy excitations on larger timescales and lengthscales. The East model, used to arrive at the universal fit (equation 6.5), is a simple KCM that displays hierarchical dynamics [237], for which the energy scale for structural relaxation obeys $J_a \propto \ln(a/a_0)$. The parameter a_0 is the lengthscale associated with an "elementary" excitation, the minimal building block of particle motion. As shown in Fig. 6.7b, J_a goes as the natural logarithm of a, $J_a \propto \ln(a)$ for the 10 3d model atomistic glass-formers studied here. This relationship holds closely for $a > \sigma/4$, which implies that displacements on the order of $\sigma/4$ might represent the smallest meaningful particle motions, or "elementary excitations." Such small displacements might correspond to, for example, cage escapes [238, 236] or neighborship changes [239]. It is also possible that the elementary excitations correspond to even smaller motions that we can not resolve using our methodology, and displacements on the order of $\sigma/4$ belong to a higher "energy level" within the hierarchy. In either scenario, a short flurry of elementary excitations occurring in rapid succession would give rise to the displacements on the order of $a = \sigma$ studied previously. This does not negate our earlier result that localized excitations are not associated with a growing dynamical lengthscale, but rather implies that the correlated motions that we observe for particle displacements of the order $a = \sigma$ represent the time-superposition of several elementary excitations correlated in space and time through dynamical facilitation. The distinction between elementary excitations and these larger excitations that occur on longer timescales is mostly inconsequential; within the hierarchical paradigm, dynamics on different scales are self-similar, and should exhibit the same overall behavior.

Table 6.2 Fitting parameters for Fig. 6.8.

Model	ρ	species	T_o	J	$\log(\tau_o)$	$stdev$	$T_{o,a\alpha}$	$J_{a\alpha}$	$c_{a\alpha,T_o}$	$\frac{c(T_o)}{c(T_{min})}$	$1-R2$
KA	1.15	A	0.619	1.319	0.782	0.0035	0.55	5.9	0.07	67	2e-4
		B	0.615	1.306	0.829	0.0035	0.55	3.8	0.02	39	1e-5
	1.2	A	0.690	1.875	0.993	0.0106	0.69	7.8	0.02	46	1e-4
		B	0.715	1.747	0.969	0.0235	0.69	4.9	0.05	74	3e-4
	1.25	A	0.817	2.463	1.078	0.0120	0.78	10.4	0.08	122	3e-4
		B	0.841	2.315	1.029	0.0103	0.78	6.7	0.03	39	1e-4
	1.3	A	1.145	2.546	0.634	0.0036	0.96	13.0	0.03	73	4e-4
		B	1.131	2.506	0.745	0.0050	0.96	8.2	0.03	38	2e-4
LJ3D	1.2	A	0.503	2.205	0.935	0.0349	0.49	4.8	0.23	49	1e-3
		B	0.517	2.125	0.614	0.1087	0.49	4.3	0.21	35	6e-4
	1.25	A	0.591	3.040	1.057	0.0085	0.55	7.0	0.04	14	5e-4
		B	0.586	3.180	0.880	0.0086	0.55	6.2	0.11	10	5e-4
	1.296	A	0.709	3.723	1.004	0.0141	0.67	9.2	0.05	29	2e-4
		B	0.712	3.754	0.756	0.0117	0.67	7.9	0.12	18	1e-3
WCA3D	1.2	A	0.305	0.650	1.200	0.0284	0.35	2.2	0.13	624	2e-4
		B	0.310	0.637	0.866	0.0251	0.27	1.7	0.05	43	3e-5
	1.25	A	0.402	1.176	1.241	0.0113	0.38	3.6	0.03	69	1e-4
		B	0.425	1.096	0.820	0.0060	0.38	2.8	0.10	28	2e-4
	1.296	A	0.531	1.739	1.160	0.0452	0.51	5.2	0.04	113	2e-4
		B	0.564	1.607	0.655	0.0495	0.51	4.1	0.12	40	4e-4

6.2.5 Excitations and Transport

As excitations occur throughout the system, the particles undergo structural relaxation. The energy scale of the microscopic excitations is related to the bulk relaxation behavior of the system through the "universal fit," given by equation 6.5. This quadratic relationship between the relaxation time τ and the inverse temperature arises as a theoretical prediction based on kinetically constrained models, and has been verified for a wide range of experimental and simulated glass-forming liquids [107]. The parameter J in equation 6.5 is proportional to the energy scale for excitations of the type that produce bulk relaxation within the system. In analogy with the definition of the standard α-relaxation time τ_α, we define a relaxing particle as having moved a distance $a_\alpha = 2\pi/S(q_0)$, where q_0 is the wavevector for which the isotropic structure factor $S(q)$ is at a maximum [240]. Boltzmann fits for J_{a_α} computed from the concentration of excitations c_{a_α} are shown in Fig 6.8a for the 10 3d systems studied.

The relationship between J and J_{a_α} depends on specific factors of the model, but should roughly correspond for a given class of models (i.e., J_{a_α}/J, is a constant > 1) [241]. The values of J and T_o are measured by fitting structural relaxation data to equation 6.5 as outlined in reference [107] (see also, SI). Fig. 6.8b and table 6.2 show the fits to the structural relaxation data for the 10 3d systems studied.

Fig. 6.9a,b shows a comparison of energy scales J and onset temperatures T_o obtained from fitting the transport and excitation data. We observe that the onset temperatures T_o typically correspond well for the two independent measurements. The large error bars in

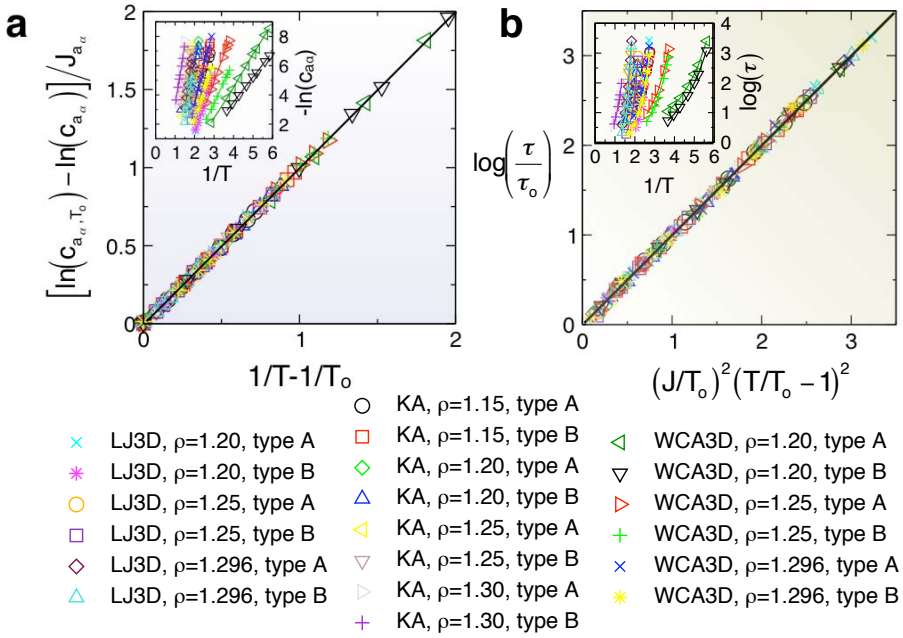


Figure 6.8 Comparison between excitation and transport energy scale J. (a) Data collapse to equation 6.4 for 3 different models at 10 different statepoints for excitations on a lengthscale a_α. (b) The transport data collapse to a quadratic master curve given by equation 6.5. The fit parameter J is physically related to the excitation energy scale J_{a_α}. Fit parameters are given in table 6.2

Fig. 6.9b reflect the fact that the onset of facilitated dynamics (i.e., the crossover to the regime in which dynamics are rare and activated) is smooth rather than abrupt. As predicted, the energy scales J for excitation and transport data have a consistent relationship, with $\alpha = J_{a_\alpha}/J \approx 12/5$, as shown in Fig. 6.9a. An exception is posed by the type A (larger) particles in the KA system (not shown in Fig. 6.9, but shown in table 6.2), which have a much larger energy scale for excitations than than the transport data suggests. These particles rarely undergo excitations at low T, but rather tend to relax mostly as a secondary effect of excitations involving smaller B particles. Thus J corresponds much more closely to J_{a_α} for the type B particles (see SI for a further discussion of this topic). For the other systems with more comparable particle sizes, the effect is not as pronounced. Our collective results demonstrate a direct quantitative connection between the microscopic excitations and the bulk relaxation dynamics.

6.2.6 Dynamical Facilitation and Dynamical Lengthscales

At any particular time, localized excitations are distributed randomly throughout the system, as evidenced by Fig. 6.6. However, excitations do not appear randomly; rather their

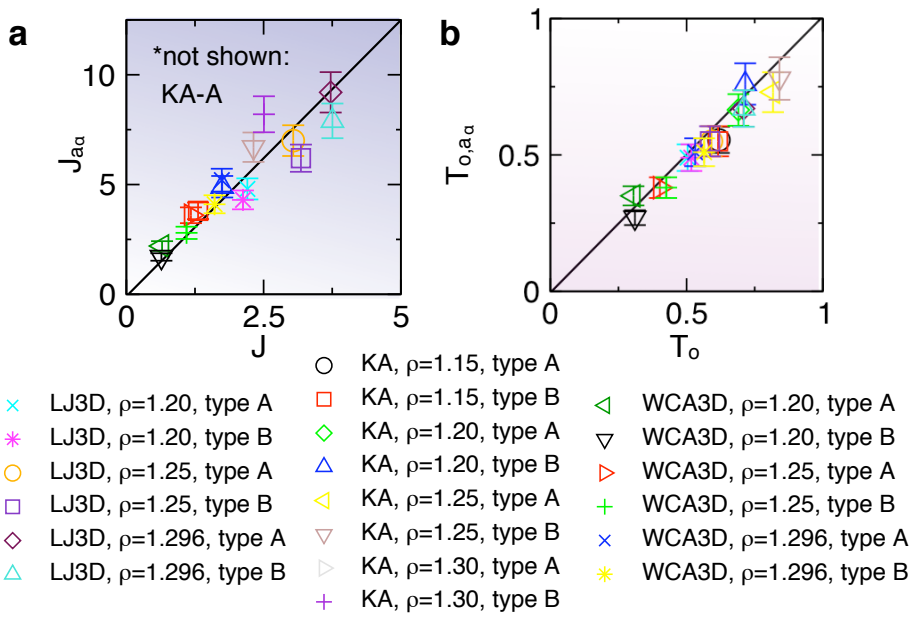


Figure 6.9 Comparison between the energy scale and onset temperature obtained from excitation and transport data. (a) Comparison of energy scales J. The excitation energy scale J_a is measured for a lengthscale a_α, but we obtain equivalent qualitative results for other values of a with a different average value of J_a/J. As described in detail in the main text, the J values for the type A particles for the KA system are outliers and are not shown. (b) Comparison between T_o obtained from excitation and transport data.

occurrence is correlated with other excitations in space and time. Excitations "facilitate" the formation of nearby excitations at a later time through a process known as dynamical facilitation [103, 230, 242, 243]. Fig. 6.10 depicts a simulation trajectory for the 2d WCA-1 system and gives some insight into how dynamical facilitation occurs on an atomistic scale. The top panel Fig. 6.10a depicts the time evolution of particle displacements. The bottom panel depicts a related quantity $\Delta nbrs$, that represents the fraction of a particle's initial nearest neighbors that are lost. Initially, localized excitations appear randomly throughout the system with concentration that depends on temperature. The occurrence of an excitation gives rise to "surges" of high-frequency, low-amplitude motion in the surrounding regions. These stringy surges are characterized by correlated displacements of less than a cage size in magnitude (depicted by light blue particles in Fig. 6.10a). These back and forth surges are ubiquitous and reversible, and their lengthscale is governed by the distance between the initial excitations $l_{surge} \sim c^{-1/3}$. In some cases, rare fluctuations produce strong surges that cannot easily be reversed, which gives rise to new excitations. These new excitations tend to fill the unrelaxed regions between the initial excitations, such that the initial excitations "connect" on timescales on the order of τ_α.

A small subset of the exciting particles exhibit large displacements a, on the order of

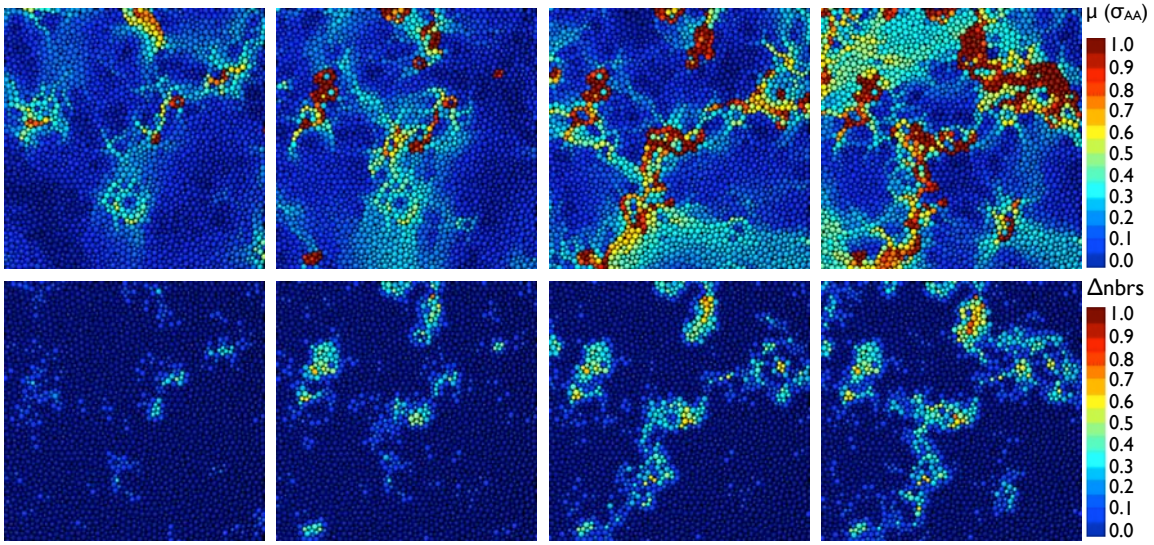


Figure 6.10 Excitations, dynamical facilitation and dynamical heterogeneity. (a) Time evolution of particle displacements for subsection of a 10,000 particle WCA2D-1 system at T=1.1, $\rho = 0.75$. (b) Time evolution of $\Delta nbrs$, the fraction of initial neighbors lost, for the same system as in (a). The timescales shown are approximately (left to right): $\tau_\alpha/10$, $\tau_\alpha/2$, τ_α, $3\tau_\alpha/2$. Movies corresponding to these trajectories are given in the SI.

about 1 to 1.5 particle diameters σ and thus tend to fully replace the initial position of their neighbors. These particles represent the "microstrings" described in reference [229] that make up the fundamental building blocks of string-like motion. String-like motion is a standard feature of glassy systems that has been observed in model atomistic [244, 229] and molecular [242] liquids, colloids [245] and granular systems [31]. String-like motion is characterized by highly-mobile particles that follow one another in quasi-1d paths [244]. The length of the strings grows with increased supercooling [244, 229, 245, 31]; however, as described in detail by Glotzer and coworkers [229, 246, 70] particles within strings do not jump simultaneously. Rather, the strings are made up of smaller microstrings that move at different times [229]. Although the mean string length grows with supercooling, the length of the microstrings remains constant [229], in direct analogy with the excitations described here. That is, the microstings are mobile subset of the hierarchy of excitations characterized here; for large excitations of the order $a = \sigma$, the excitations and the microstrings are equivalent.

The growth of strings of mobile particles (red) is depicted in Fig. 6.11a for a small subsection of a 10,000 particle WCA2D-1 system. The panel depicts a "space-time" representation, where time is depicted as a spatial dimension [108]. Fig. 6.11a shows the mobility field in the xy plane, with time perpendicular to the page. Facilitated particle motions

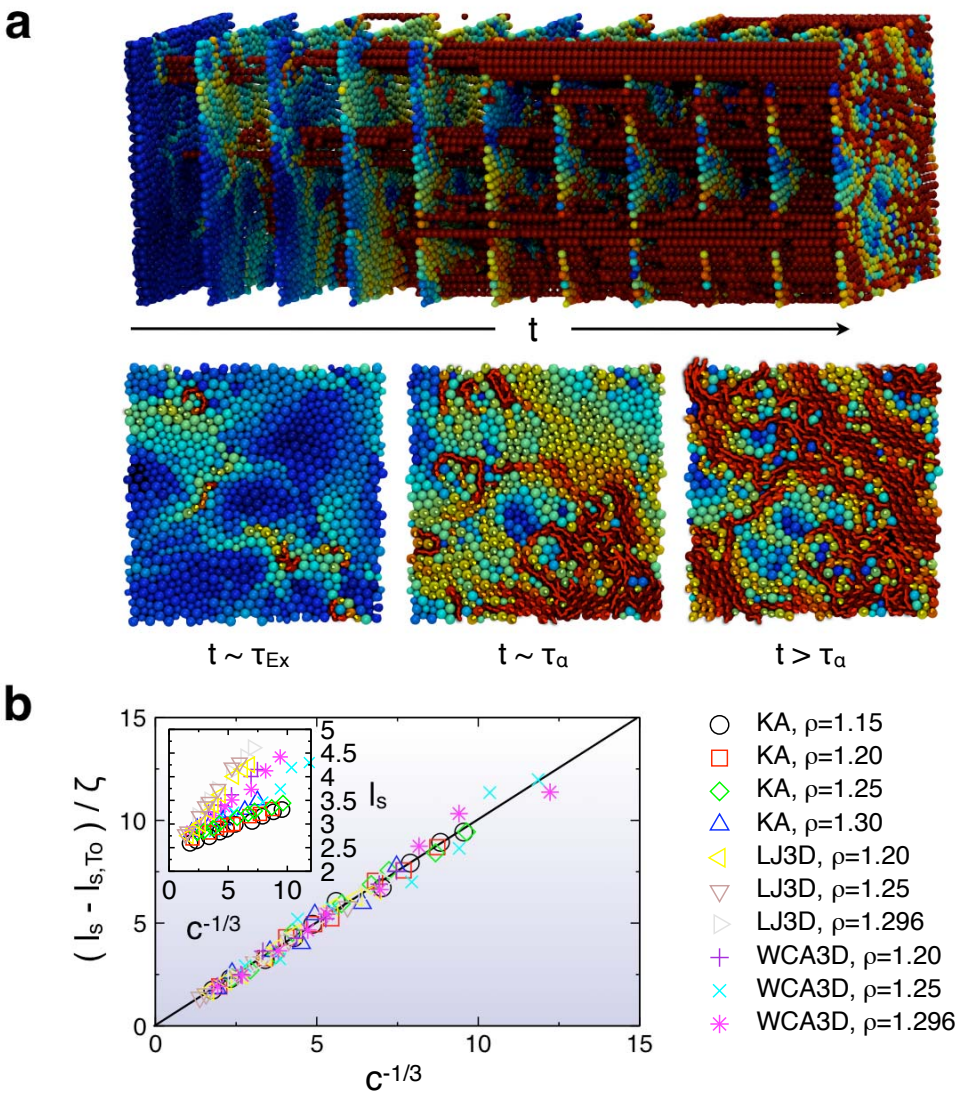


Figure 6.11 Excitations and string-like motion. (a) Space-time plot depicting the formation of strings. Notice that the strings grow piecewise from an initial excitation. (b) Average string length l_s versus the mean distance between excitations $c^{-1/3}$. Excitations are measured based on a lengthscale $a = \sigma$.

on a lengthscale $a = \sigma$ trace out "excitations lines" in space-time [108], highlighting the connectedness of the localized excitations through dynamical facilitation. This phenomenon can also be observed in Fig. 6.10 and from conventional movies of particle motion provided in the SI.

As the initial excitations become increasingly sparse at lower temperatures, the mean string length tends to grow. Since strings are defined by particles that strictly replace the initial position of a neighboring particle (see SI), they are often broken by pairs of particles

Table 6.3 Fitting parameters for data collapse in Fig. 6.11b.

Model	ρ	ζ	l_{s,T_o}	1-R2
KA	1.15	0.047	2.23	4e-3
	1.2	0.052	2.19	6e-3
	1.25	0.046	2.32	5e-3
	1.3	0.054	2.28	2e-3
LJ3D	1.2	0.15	2.13	2e-3
	1.25	0.15	2.19	4e-3
	1.296	0.10	2.51	5e-3
WCA3D	1.2	0.08	2.26	9e-3
	1.25	0.09	2.26	5e-3
	1.296	0.09	2.30	1e-2

that do not precisely satisfy the overlap criterion. As a result, the mean string length is always less than the mean distance between excitations. Nonetheless, as shown in Fig. 6.11b there is a direct relationship between the average string length l_s and the mean distance between initial excitations, $l_s = \zeta c_{a=\sigma}^{-1/3}$. The constant ζ depends on several factors, including the tendency for mobile particles to engage in broken strings, the size of groups of simultaneously exciting particles (Fig. 6.6), and the system density, all of which are independent of temperature. As shown in Fig. 6.11b, the direct relationship holds for the full range of $T < T_o$ for the 10 different 3d systems studied. The relationship between the l_s and c lends a possible explanation to the previously unexplained observation that the strings obey an exponential probability distribution [244, 31]. Because the excitations are distributed randomly in space, the separation l is described by a Poisson random variable, which may provide the physical basis for the observed exponential distribution.

6.2.7 Conclusions

Our collective results support a picture of glassy systems in which particle dynamics are governed by facilitating localized excitations [230]. Our results imply that the mechanism of this facilitation is hierarchical in nature, as originally proposed by Palmer *et al*. [227]. We have demonstrated, for several atomistic systems, a direct connection between the concentration of excitations, bulk transport properties, and a growing dynamical length scale for dynamical heterogeneity. In the future, the microscopic procedure that we have introduced for detecting local excitations in atomistic liquids can be extended to study excitations in more complex simulated systems, such as molecular liquids, or experimental systems, such as colloidal suspensions or granular matter. Our results imply a possible connection

Table 6.4 Model LJ and WCA systems used in this study.

abbreviation	d	f_A	σ_{AA}	σ_{AB}	σ_{BB}	ε_{AA}	ε_{AB}	ε_{BB}	$r_{cut,AA}$	$r_{cut,AB}$	$r_{cut,BB}$	m_A	m_B
KA [133]	3	0.8	1.0	0.8	0.88	1.0	1.5	0.5	$2.5\sigma_{AA}$	$2.5\sigma_{AA}$	$2.5\sigma_{AA}$	1.0	1.0
LJ3D [134]	3	0.5	1.0	11/6	5/6	1.0	1.0	1.0	$2.5\sigma_{AA}$	$2.5\sigma_{AA}$	$2.5\sigma_{AA}$	2.0	1.0
WCA3D [251]	3	0.5	1.0	11/6	5/6	1.0	1.0	1.0	$2^{1/6}\sigma_{AA}$	$2^{1/6}\sigma_{AB}$	$2^{1/6}\sigma_{BB}$	2.0	1.0
WCA2D-1 [252]	2	0.5	1.0	1.2	1.4	1.0	1.0	1.0	$2^{1/6}\sigma_{AA}$	$2^{1/6}\sigma_{AB}$	$2^{1/6}\sigma_{BB}$	1.0	1.0
WCA2D-2 [236]	2	0.3167	1.0	1.1	1.4	1.0	1.0	1.0	$2^{1/6}\sigma_{AA}$	$2^{1/6}\sigma_{AB}$	$2^{1/6}\sigma_{BB}$	1.0	1.0

between excitations and local structure; this relationship can be explored using the existing framework for relating structure to dynamics that has already been applied to particle motions spanning much longer timescales than the excitations considered here [247, 248]. The picture of facilitating excitations does not necessarily preclude alternate descriptions of glassy supercooled liquids, and connections between our results and other theories and scenarios of the glass transition should be explored.

6.2.8 Materials And Methods

The results that we present are based on five different models of atomistic glass formers, including two 2d systems and three 3d systems. All of the systems consist of binary mixtures of 10,000 particles with interactions given by either the Lennard-Jones (LJ), or Weeks-Chandler-Andersen (WCA) potential [249]. The parameters for the models are listed in the table below. Molecular dynamics simulations are carried out in HOOMD-blue [250] on the graphics processing unite (GPU). Further details regarding our simulation methods and transition path sampling scheme are given in the SI.

6.2.9 Supplementary Information: Transition Path Sampling Scheme

The basic goal of our transition path sampling (TPS) scheme is to measure the timescales and cooperativity lengthscales associated with irreversible particle displacements in our systems. To do so, we tag a particle in the system, and sample trajectories over which the particle moves a particular distance a. We are interested in trajectories for which the tagged particle undergoes a discrete displacement of length a over a short period of time rather than a series of smaller intermittent displacements adding up to a over long periods of time. To restrict our sampling to these types of trajectories, we choose our tagged particles from a random distribution of particles that jump at the earliest point along a trajectory. Specifically, we run a trajectory with increments δt, where δt is about two orders of magnitude shorter than the commitment time Δt. For each trajectory $x(t)$, we determine whether the tagged

particle "1" satisfies the path functional:

$$h_{\mathrm{Ex}}[x(t)] = \begin{cases} 1 & \text{if} \quad |\bar{r}_1(t) - \bar{r}_1(0)| \approx 0, \quad 0 < t < t_s \\ & \text{and} \quad |\bar{r}_1(t) - \bar{r}_1(0)| \geq a, \quad \Delta t + t_s < t < \Delta t + 2t_s \\ 0 & \text{otherwise.} \end{cases} \tag{6.12}$$

This functional is impractical to implement directly in computer simulations, since (1) it requires a somewhat arbitrary definition of what constitutes a displacement of approximately zero, and (2) it requires that we check every snapshot along a trajectory, which may consist of hundreds of points. In practice, the functional can be implemented by checking four points along the trajectory:

$$h_{\mathrm{Ex,simple}}[x(t)] = \begin{cases} 1 & \text{if} \quad |\bar{r}_1(\Delta t + 2t_s) - \bar{r}_1(0)| \geq a \\ & \text{and} \quad |\bar{r}_1(\Delta t + t_s) - \bar{r}_1(t_s)| \geq a \\ 0 & \text{otherwise.} \end{cases} \tag{6.13}$$

In words, the functional determines whether the tagged particle has undergone a displacement of length a over a trajectory of length Δt, which is allowed to float according to the variable path length transition path shooting algorithm of reference [233]. Additionally, the functional checks whether the tagged particle has still undergone a displacement of length a when the trajectory is extended by the sojourn time t_s forward and backward in time. This ensures that the particle has not moved back to its original position during the sojourn time. To reduce the possibility for errors, the functional can be checked at intermediate points by extending the trajectory incrementally until t_s is reached. For the results that we report in this article, we use one intermediate interval $t_s/2$.

Each initial transition path is created by choosing a short trajectory from the equilibrium system that satisfies the criterion described above. For each initial path, we perform 100 equilibration trajectories and then 100 production trajectories for which we collect excitation statistics. Typically, 100-500 initial paths are chosen in total.

6.2.10 Supplementary Information: Excitation Characteristics

Previously, we reported representative results for excitations of size $a = \sigma$ where we made no distinction between the different particle types (A or B) when tagging particles. In this section, we explore the effect of the type of the tagged particle and the size of the displacement a on our results. We show that, although the excitations are affected the particle type

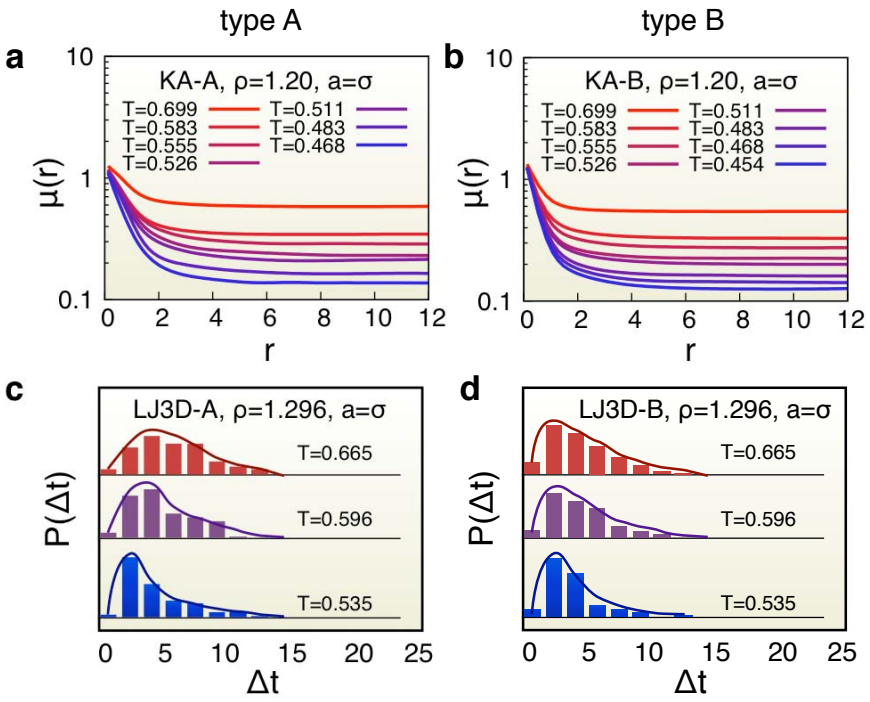


Figure 6.12 Effect of tagged particle type on excitation timescale and cooperativity lengthscale. (a) Mobility field $\mu(r)$ for type A particles in the KA system. (b) Mobility field $\mu(r)$ for type B particles in the KA system. (c) Probability distribution of commitment times Δt for type A particles in the KA system. (d) Probability distribution of commitment times Δt for type B particles in the KA system.

and displacement length, our fundamental results, that the lengthscales and timescales associated with the excitations do not change noticeably with decreasing temperature, remain unchanged.

Supplementary Information: Effect of Tagged Particle Type

Earlier, we reported that the timescales and cooperativity lengthscales associated with the excitations do not change noticeably with temperature, below on onset temperature T_o. Our results were based on randomly choosing exciting particles from the system without considering their type. Here, we show that we obtain the same results when restricting our sampling to tagged particles of a particular type. Fig. 6.12a shows the mobility field $\mu(r)$ for tagged particles of type A and B as a function of temperature for the KA system. We observe that the cooperativity lengthscale remains unchanged for exciting particles of either type, despite the relatively large disparity in particle sizes (see Materials and Methods) and excitation energy scales (see main text) for this system.

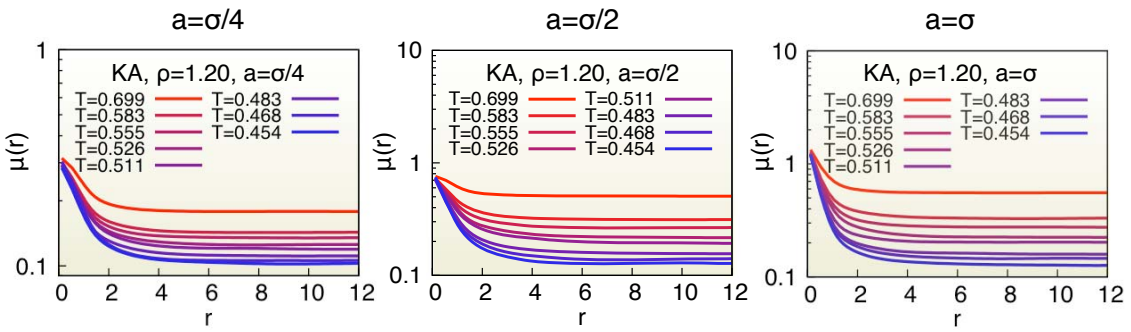


Figure 6.13 Effect of displacement size a on excitation timescale and cooperativity lengthscale. The displacement size is doubled for each panel, left to right. Notice that the cooperativity lengthscale does not change with temperature.

Effect of Tagged Particle Displacement Length a

In the main text, we reported a correlation distance of about 2-5 particle diameters for excitations on a lengthscale $a = \sigma$. Here, we show that this result does not change for smaller displacement distances on the order of a cage size. Fig. 6.13 shows the mobility field $\mu(r)$ for trajectories over which a tagged exciting particle at the origin moves a distance $a = \sigma/4$, $a = \sigma/2$, and $a = \sigma$, respectively from left to right. The cooperativity lengthscale does not change appreciably with the displacement size, and does not appear to grow with decreasing temperature.

6.2.11 Supplementary Information: Measuring the Excitation Energy Scale and Onset Temperature From Excitations

In this section, we describe the detailed method that we use to obtain the values of the excitation energy scales and the onset temperatures reported in this study. We demonstrate that these quantities are fairly robust, and different methods for measuring them yield statistically equivalent results. In the main text, we presented a formula for measuring J_a and T_o from the concentration of excitations c_a. In this section, we expound upon our methodology, and the sensitivity of our results to particular parameters. As outlined in the main text, we determine J_a and T_o by first computing the concentration of exciting particles according to

$$c_a = \left\langle \frac{1}{V} \sum_i^N h_{i,a}[x(t)] \right\rangle . \tag{6.14}$$

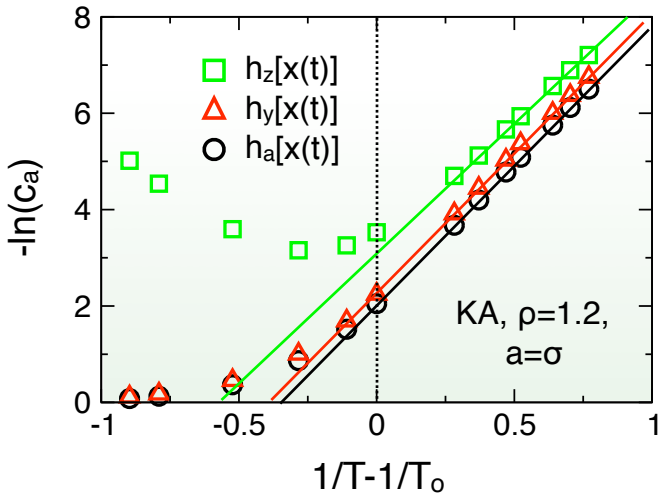


Figure 6.14 Effect of path functional on quantities derived from excitations. The panel depicts plots of the concentration of excitations c versus the temperature T for three different path functionals defined in the text. In words, the functional $h_a[x(t)]$ is the standard functional used for this study, where we check whether a particles has undergone a displacement of a or greater over a time window t_a using energy-minimized coordinates. The second functional $h_y[x(t)]$ is similar to $h_a[x(t)]$, except that we require that the displacement is irreversible (i.e., the particle does not move back towards its original position over at least a sojourn time t_s). The third functional $h_z[x(t)]$ is similar to $h_y[x(t)]$, except that we require that the exciting particle is quiescent for a set period t_s before undergoing a displacement.

where $h_{i,\mathrm{a}}[x(t)]$ is given by

$$h_{i,\mathrm{a}}[x(t)] = \begin{cases} 1 & \text{and} \quad |\bar{r}_i(t) - \bar{r}_i(0)| \geq a, \quad t = t_a \\ 0 & \text{otherwise.} \end{cases} \tag{6.15}$$

The timescale t_a is chosen to be greater than the commitment time Δt to allow exciting particles sufficient time to move, but not so long that particles can excite multiple times. We typically choose $\max(\Delta t) \leq t_a \leq 3\max(\Delta t)$, where $\max(\Delta t)$ is the maximum of the $P(\Delta t)$ distribution obtained from path sampling. Values of c_a are obtained by running 1000 trajectories of length t_a at each temperature, and computing the ensemble average in equation 6.14. The values of J_a and T_o are obtained by constructing plots of $-\ln(c_a)$ versus $1/T - 1/T_\mathrm{o}$ (see main text). The value of $T_{\mathrm{o},a}$ is determined by temperature at which the linear relationship in $-\ln(c_a)$ versus $1/T - 1/T_\mathrm{o}$ breaks down. The value of J_a is determined by the slope of the line for $T < T_\mathrm{o}$.

Effect of Strictly Enforcing Irreversibility

The energy scale J_a simply quantifies the relative rate at which the concentration of excitations c_a changes with temperature. Thus, it is not surprising that we can obtain the same value of J_a for several different methods of estimating c_a; even if a given method misestimates c_a, so long as it does so consistently as a function of temperature, J_a will remain unchanged. Fig. 6.14 shows plots of $-\ln(c_a)$ versus $1/T - T/T_o$ for three different path functionals h[x(t)]. For the first path functional, we apply the usual criterion $h_a[x(t)]$, described in equation 6.15. For the second functional $h_y[x(t)]$, we apply a similar criterion as for h_a, except that we require that the displacement is irreversible. That is, we require that an exciting particle does not move back towards its original position over at least a sojourn time t_s. This is enforced numerically by:

$$h_y[x(t)] = \begin{cases} 1 & \text{and} \quad |\bar{r}_1(t_a + t_s) - \bar{r}_1(t_s)| \geq a \\ & \text{and} \quad |\bar{r}_1(t_a + 2t_s) - \bar{r}_1(0)| \geq a \\ 0 & \text{otherwise.} \end{cases} \tag{6.16}$$

Notice that this functional is similar to the definition of $h_{\text{Ex, simple}}[x(t)]$ defined above for TPS, except that the commitment time Δt, which is allowed to float in path sampling, is set to a constant value t_a here. As shown in Fig. 6.14, $h_a[x(t)]$ and $h_y[x(t)]$ give very similar results, except that $-\ln(c_a)$ is slightly larger for $h_y[x(t)]$, indicating that slightly fewer excitations are detected than for $h_a[x(t)]$ Nevertheless, both To (the point at which the curves begin to deviate from Boltzmann behavior) and J_a (the slope of the lines for $T < T$o) are indistinguishable for the two measurements.

While $h_y[x(t)]$ ensures that the exciting particle does not quickly return to its initial position, it does not enforce that the particle is "quiescent" for a given period before exciting. For the third functional $h_z[x(t)]$, we strictly require that the tagged particle does not move for a sojourn time t_s before undergoing a displacement of length $a = \sigma$. In practice, this is enforced by requiring that the particle move no farther than a (somewhat arbitrary) threshold $\sigma/4$ during the initial sojourn time t_s, beyond which the particle is no longer considered stationary:

$$h_z[x(t)] = \begin{cases} 1 & \text{and} \quad |\bar{r}_1(t_s) - \bar{r}_1(0)| \leq a/4 \\ & \text{and} \quad |\bar{r}_1(t_a + t_s) - \bar{r}_1(t_s)| \geq a \\ & \text{and} \quad |\bar{r}_1(t_a + 2t_s) - \bar{r}_1(0)| \geq a \\ 0 & \text{otherwise.} \end{cases} \tag{6.17}$$

As shown in Fig. 6.14, $h_z[x(t)]$ exhibits a pseudo-parabolic behavior for $-\ln(c_a)$ versus

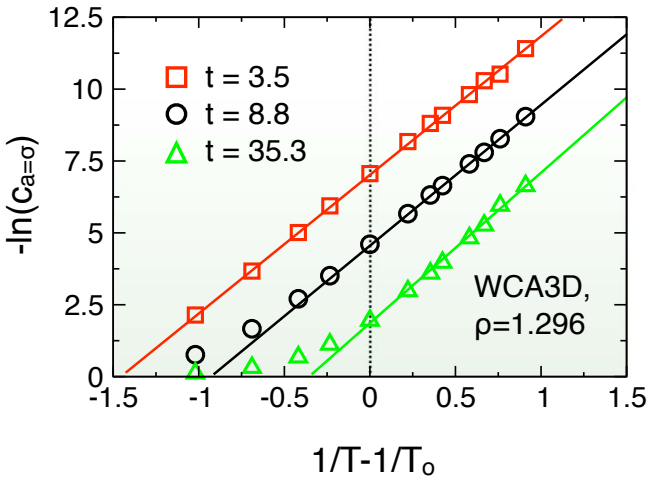


Figure 6.15 Effect of the observation time window on quantities derived from excitations. The panel shows plots of $-\ln(c_a)$ versus $1/T - T/T_o$ for three different values of the time window t_a. For this system, the max commitment time Δt is approximately 6, $\max[\Delta t] \approx 6$. We typically obtain the best results for J_a and T_o for $\max[\Delta t] \leq t_a \leq 3\max[\Delta t]$.

$1/T - 1/T_o$. This reflects the fact that for higher temperatures, the probability of observing a quiescent particle is relatively small, and thus c_a above T_o is not indicative of the true concentration of excitations. Below T_o, the curve based on $h_z[x(t)]$ is similar to that obtained from the other path functionals, except that the parabolic effect mentioned previously persists somewhat, and results in a slightly low measurement of the slope J_a. In practice, any of these measurements give similar results for J_a and T_o.

Effect of Time Window

The Boltzmann-like temperature dependence of c_a on T can only be obtained for timescales on the order of the commitment time, Δt. Fig. 6.15 shows plots of $-\ln(c_a)$ versus $1/T - T/T_o$ for three different values of the time window t_a. For this system, the max commitment time Δt is approximately 6. As shown in Fig. 6.15, for t_a slightly greater than $\max[\Delta t]$, we obtain the expected behavior for plots of $-\ln(c_a)$ versus $1/T - 1/T_o$ (i.e., the curve exhibits Boltzmann behavior, which breaks down for $T > T_o$). For t_a shorter than $\max[\Delta t]$, the Boltzmann behavior never clearly breaks down, since over such short time windows, the breakdown of the separation of timescales between commitment times and waiting times is never observed, even well above T_o. Thus, the measurement of T_o is obscured for very short t_a, although the value of J is largely unchanged. For t_a much longer than $\max[\Delta t]$, multiple excitations can occur for the same particle over the same interval

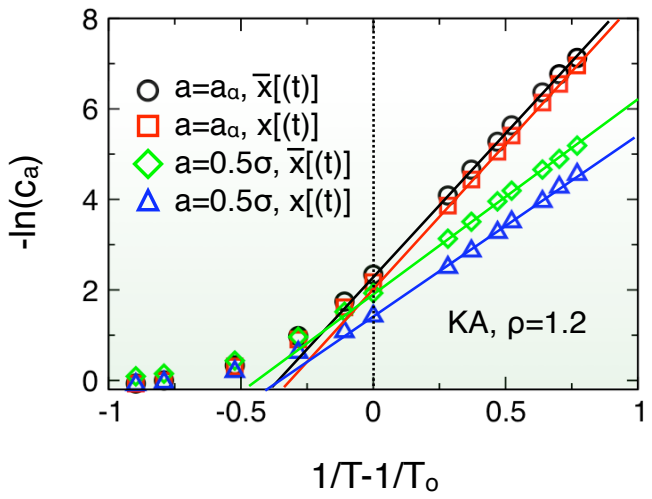


Figure 6.16 Effect of energy minimization on quantities derived from excitations. The panel shows plots of $-\ln(c_a)$ versus $1/T - T/T_\mathrm{o}$, measured using both energy-minimized and unprocessed coordinates, for two different values of the displacements length a. Notice that, although the values of c_a are affected by energy minimization, the values of J_a and T_o (the slopes of the lines and the temperatures at which the Boltzmann fit breaks down) are largely unchanged.

even below T_o, causing a breakdown in Boltzmann behavior, which obscures the value of J_a. The fact that the mean waiting times between excitations are shorter for lower temperatures results in the observed upward concavity of the curve. In summary, although particularly bad choices for t_a can give slightly incorrect results, even in the extreme cases t_a too large or too small described here, we still obtain reasonable estimates of J_a.

Effect of Energy-Minimization / Coarse-Graining

In our study, we used energy-minimized coordinates to rule out the possibility that the observed particle displacements arise due to transient vibrations. Over the short timescales on which excitations occur, time coarse-graining can be applied to obtain identical results. In practice, we find that neither coarse-graining nor energy-minimization is strictly necessary for measuring J_a and T_o from the concentration of excitations. Fig. 6.16 shows plots of $-\ln(c_a)$ versus $1/T - T/T_\mathrm{o}$, where c_a is measured using both energy-minimized coordinates (denoted by an overbar) and unprocessed coordinates. The values of c_a are affected by the different methods, with fewer excitations being recognized for energy-minimized coordinates. Unsurprisingly, this effect is more pronounced for smaller values of the displacement length a. However, the relative rate at which c_a changes with temperature is not dependent on whether or not the coordinates are energy-minimized, and the values of J_a and T_o are largely unchanged. This implies that, provided a is large enough, a very simple procedure

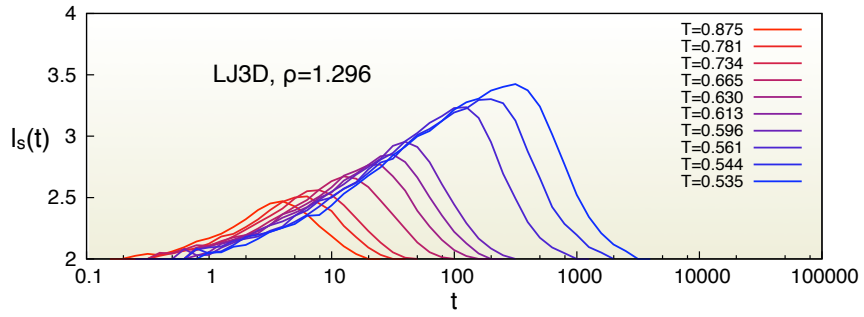


Figure 6.17 Average string length for energy-minimized coordinates as a function of temperature for a binary Lennard-Jones mixture. The dynamical lengthscale l_s is taken as the maximum for $l_s(t)$.

can be applied to measure J_a and T_o in simulations and experiments based on pairs of snapshots, separated by t_a, requiring no time coarse-graining or energy minimization.

6.2.12 Supplementary Information: Measuring the excitation energy scale and onset temperature from transport quantities

Computing J and T_o from transport data requires that we compute the structural relaxation time τ_α as a function of temperature for all of the systems studied. This involves first determining 15-20 statepoints that span the range of accessible temperatures for a given system and density. For each statepoint, the system is quenched from the previous (higher) temperature and equilibrated for a time period $25\tau_\sigma$, where τ_σ is the time required for the mean squared displacement to exceed 1 particle diameter σ. One hundred long trajectories of length $100\tau_\sigma$ are then used to measure the self-intermediate scattering function $F_s(q_0,t)$, where q_0 is the wavevector corresponding to the first peak in the isotropic structure factor $S(q)$ [133]. The alpha relaxation time τ_α is obtained by fitting $F_s(q_0,t)$ in the α relaxation regime to an exponential function. The parameters J and T_o are obtained by optimizing the universal quadratic fit for these τ_α, as outlined in reference [107]. All of the systems under consideration contain 10,000 atoms total (see materials and methods). To sample these systems efficiently, we use HOOMD-blue [250] to run MD simulations on the graphics processing unit (GPU).

6.2.13 Supplementary Information: Measuring Strings

For our study, we have used the average string length l_s to obtain a measure of the characteristic dynamical lengthscale for our systems. Strings are defined by clusters of particles that, within a time interval t, replace the initial position of a neighboring particle. In practice,

quantifying replacement requires that we introduce a threshold value for replacement r_{cut} such that $\left|\bar{r}_i(t) - \bar{r}_j(0)\right| < r_{cut}$, where the overbar denotes energy-minimized coordinates. For our systems, we take r_{cut} in the range 0.5σ for all systems. Mobile particles i and j belong to the same string if i replaces j over a time window t and i and j are neighbors at $t = 0$ (i.e., $\left|\bar{r}_i(0) - \bar{r}_j(0)\right| < r_{nbr}$, where r_{nbr} encompasses the first peak of the radial distribution function $g(r)$). Since particles can only replace one another by moving in a pseudo-one-dimensional line, the string length goes as the number of particles in the string: $n_s \propto l_s$. As shown in Fig. 6.17, the average string length $l_s(t) \equiv \frac{\langle l^2(t)\rangle}{\langle l(t)\rangle}$ grows as a function of decreasing temperature. We define l_s as $l_s(t^*)$, the maximum average string length for a given temperature. The timescale for the maximum string length is identical to $t^*_{\alpha_2}$, the peak time of the non-Gaussian parameter [73, 71]. At the peak time $t^*_{\alpha_2}$, the string length distribution is exponential while the cluster size distribution is a power-law [73, 71, 31].

Chapter 7
Characterizing Transitions

Structural quantities such as order parameters and correlation functions are often employed to gain insight into structural transitions in traditional condensed matter systems, such as systems of rod-like or spherical particles. Since these methods were designed with specific, simple systemes in mind, they can not be directly applied to systems of complex particles encountered in the field of nano and microscale self-assembly. In this chapter, we develop methods for creating robust structural metrics for complex particle structures. We present three articles [34, 38, 37] authored by A.S. Keys, C.R. Iacovella, and S.C. Glotzer that deal with the problem of applying "shape matching" methods from computer science to characterizing particle structures in complex assembled systems. In addition to studying structural transitions, these methods can be used for quantifying structural perfection, identifying unknown structures, and other more abstract applications. In our first article (section 7.1), we provide a broad overview of shape matching methods, and their potential applications in characterizing assembled systems. In our second article (section 7.2), we focus on a particular class of structural metrics known as "harmonic descriptors," and describe how their unique properties can be applied to solve particular problems involving rotational symmetries. In our third article (section 7.3), we provide a more detailed overview of the specific shape matching schemes that can be implemented to solve different types of problems. For all three studies, we provide proof-of-concept examples to demonstrate our methods. Although our examples focus primarily on the field of computational self-assembly, where we predict these methods will be most applicable, we also provide examples from the fields of experimental self-assembly, computational biology, and traditional condensed matter.

7.1 Characterizing Structure Through Shape Matching and Applications to Self Assembly

Structural quantities such as order parameters and correlation functions are often employed to gain insight into the physical behavior and properties of condensed matter systems. While standard quantities for characterizing structure exist, often they are insufficient for treating problems in the emerging field of nano and microscale self-assembly, where the structures encountered may be complex and unusual. The computer science field of "shape matching" offers a robust solution to this problem by defining diverse methods for quantifying the similarity between arbitrarily complex shapes. Most order parameters and correlation functions used in condensed matter apply a specific measure of structural similarity within the context of a broader scheme. By substituting shape matching quantities for traditional quantities, we retain the essence of the broader scheme, but extend its applicability to more complex structures. Here we review some standard shape matching techniques and discuss how they might be used to create highly flexible structural metrics for diverse systems such as self-assembled matter. We provide three proof-of-concept example problems applying shape matching methods to identifying local and global structures, and tracking structural transitions in complex assembled systems. The shape matching methods reviewed here are applicable to a wide range of condensed matter systems, both simulated and experimental, provided particle positions are known or can be accurately imaged.

The preponderance of new nanometer- and micron-sized colloidal particles of nearly arbitrary shape, composition and interaction has made possible the self-assembly of exquisitely complex structures with potential uses in a variety of technologies [12, 13, 11, 253]. Because material properties and behavior are determined by both the global and local shapes, or patterns, within the self-assembled structure [12, 254, 255, 256, 257, 258], methods and tools are needed to characterize the salient structural features of the assemblies. The field of condensed matter physics has traditionally led the way in developing algorithms for characterizing crystal structures and constructing theories to connect these structures to thermodynamics and to overall system properties [259, 260, 261]. These approaches typically involve constructing structural order parameters and/or correlation functions that can discriminate between different building block arrangements and are well developed for systems of point-like, rod-like and spherical particles [262, 53, 57, 263, 264]. Examples include nematic and smectic order parameters for systems of rods citenematic, smectic, liquidcrystals and bond order parameters [53, 265, 132, 57] for 2d and 3d systems of spheres.

However, these functions fail, in many cases, to fully describe the structural complexity of assemblies of more unusual nanocolloids, including those formed from spherical

particles [256, 266], rod-like particles [267, 257], polyhedral particles [33, 268, 269, 210, 270, 271, 272, 273], colloidal molecules [274, 275, 276, 277, 278, 11], patchy spheres [279, 280, 29, 281, 282, 283], arbitrarily-shaped objects [12, 11], polymer-tethered nanoparticles [253, 210, 284, 285, 286, 287, 288], and terminal assemblies resembling biological structures [289, 290]. For example, it is easy to envision that order parameters defined for spherical or rod shaped particles may fail when applied to more complex shaped particles, such as "Y" particles or triangular plates [11]. As a result of the increased complexity of nano building blocks, there are few "model problems" in nano and microscale self-assembly for which generally applicable order parameters can be defined. The dearth of structural metrics has lead many recent experimental and computational studies of assembled systems to rely heavily on visual inspection or *ad hoc* analysis for characterizing structures, rather than well established schemes. This approach is not optimal, since visual inspection can be time consuming and typically less accurate than mathematical analysis, and *ad hoc* analysis can be idiosyncratic, making it difficult to compare structures across independent studies. The impetus for new structural metrics is also driven by advances in microscopy techniques that allow for the direct imaging of nano and microscale systems, which have greatly extended the range of systems for which detailed structural analysis can potentially be performed. For example, the tracking of micron-sized colloidal particles in 2d and 3d is now routine [291, 64, 292, 293, 294], and high-fidelity imaging of nanoparticles [295] and their assemblies [268, 296, 297] is steadily improving. Combined with proper image processing techniques, one can extract much information about structure, such as the particle positions [294, 291] and other key features, providing detailed structural information on par with simulations. Assuming one can construct order parameters sensitive to these unique building blocks and their assemblies, similar routines can be applied to both experimental and simulated systems, allowing for direct comparison [277, 268].

Analysis techniques from the computer science field of "shape matching" offer a potentially powerful solution to the problem of creating general structural metrics for these systems. Shape matching involves defining general structural metrics that can be used to measure the degree of similarity between diverse shapes. Such similarity measures can be applied within the context of traditional condensed matter order parameter and correlation function schemes to obtain analogous quantities for more complex structures. This is possible because, in practice, most standard structural characterization schemes include an implicit concept of matching or shape similarity; that is, the schemes typically measure the degree to which a structure of interest matches another (often ideal) structure. As a familiar example, consider the standard nematic order parameter which gives an optimal value of 1 when the rod-like particles within the system are perfectly aligned, and 0 if the

rods have random orientations. In this case, the order parameter measures the degree to which the local arrangement of rods in the system, described mathematically by the angles between neighboring rods, matches with an ideal reference system with perfect alignment (see Fig. 7.1). Other structural characterization schemes and spatial or temporal correlation functions involve similar underlying concepts of matching. As we will discuss, by modifying these schemes to use shape matching methods, we retain their overall physical insight, but gain the ability to apply them to complex structures. Although we focus exclusively on simulated assembled systems here, these types of methods are general enough that they can be applied to particle systems in general, provided that the particle positions and or orientations can be determined or imaged. Examples of systems, both experimental and simulated, to which shape matching methods can potentially be applied include but are not limited to nanoparticle superlattices created from mixtures of spherical and/or non-spherical nanoparticles [297, 296], microphase separated systems, such as tethered nanoparticles and block copolymers that form crystalline and quasicrystalline domains [298, 194], colloidal ionic crystals [299], dense colloids [64] and granular matter [216, 218].

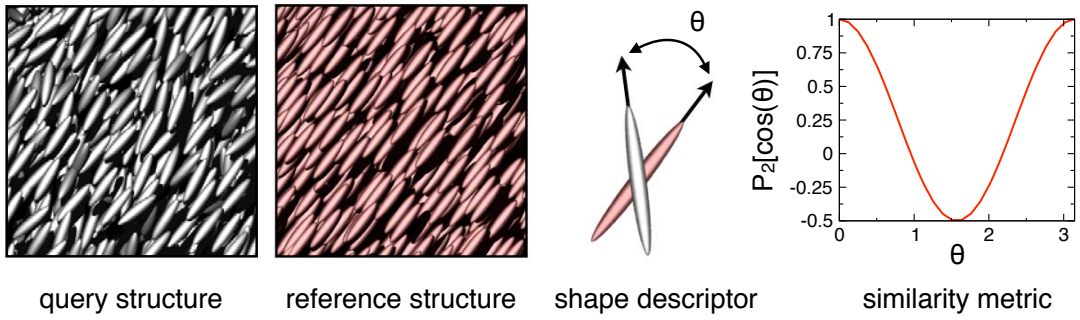


Figure 7.1 Example of an implicit shape matching scheme within the context of a standard order parameter. The panel depicts the process of computing the nematic order parameter $\bar{P}_2$ for a system of rod-like colloidal ellipsoids that assemble into an aligned ordered phase [300]. In the language of a shape matching scheme (see section 7.1.1), the colloidal system acts as a "query structure" that we wish to characterize. An ideal system for which the rods are all oriented along the average global director acts as an implicit "reference structure." The local values of the angles θ between rods in the query structure and reference structure act as "shape descriptors." The Legendre polynomial P_2 acts as a "similarity metric." The global nematic order parameter $\bar{P}_2$ is computed by averaging over local values of $P_2[\cos(\theta)]$.

This review is organized as follows. In section 7.1.1, we review shape matching methods from the literature, restricting our scope to methods that we believe are most immediately applicable to assembled systems. We describe how representative shapes can be extracted from particle systems, review the shape descriptors that are best suited to describe these shapes numerically, and show how they can be compared quantitatively. In section 7.3.5,

we apply a prototype shape matching scheme to three representative example problems from simulations of self-assembly. Our examples include identifying global structures in a microphase-separating system of polymer-tethered nanospheres [298], detecting local icosahedral clusters in a fluid of hard tetrahedral particles [33], and tracking the twisting of a helical sheet formed from polymer-tethered nanorods [301]. In section 7.1.3, we suggest new applications for shape matching methods, including constructing correlation functions, measuring local crystal grains and crystal defects, devising guided computer algorithms to map parameter spaces and search for target structures, and grouping and classifying structures based on particular structural features. To aid in the development and dissemination of new structural analysis methods based on shape matching techniques, we provide accompanying software and examples via the web [302].

7.1.1 Shape Matching

Quantifying how well structures match has been generalized in the context of shape matching [303] (see Fig. 7.2). Familiar applications include matching fingerprints and signatures [303], facial recognition [304] and medical imaging [305]. Shape matching defines the concept of the *shape descriptor*, a numerical "fingerprint" that describes a pattern or shape. Shape descriptors are associated with *query* structures and compared with *reference* structures. The degree of matching between query and reference structures is quantified by a *similarity metric*.

Matching information can be used to create order parameters and correlation functions, identify structures, and perform many other types of structural analysis. Since we can choose virtually any structure as a reference for comparison, shape matching facilitates the creation of highly specific structural metrics. In the following sections, we review the process of constructing a customized structural metric which involves choosing interesting structures to characterize, computing shape descriptors, and using similarity metrics to compare them.

Representative Structural Patterns

Before we can compute a shape descriptor, we must extract a representative structural pattern from the system. This step relies largely on physical intuition; often redundant or unimportant structural information can be discarded out-of-hand to ensure that the matching scheme is only sensitive to important structural features. One standard type of coarse-graining that is often employed, particularly to the case of small clusters of roughly spherical particles, is to consider particle positions exclusively, discarding information regarding particle sizes and

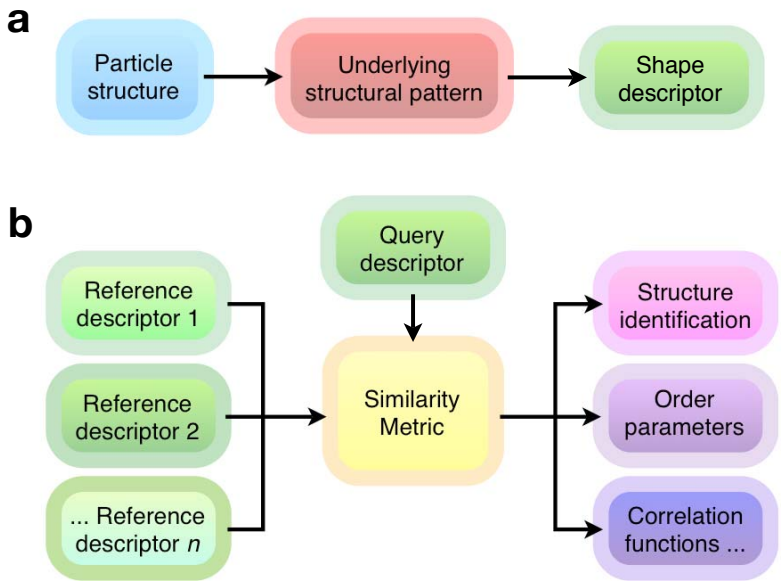


Figure 7.2 Data flow diagram for shape matching. (*a*) A structural pattern is extracted for a given query structure and then indexed into a shape descriptor, which represents a numerical fingerprint for the structure. (*b*) The shape descriptor is then compared with shape descriptors for reference structures to give a measure of similarity between shapes. Depending on how we choose the query and reference structures, the similarity value obtained may be applied to constructing order parameters, correlation functions, or other applications.

shapes which may be nearly identical (see Fig. 7.3). This type of coarse-graining can also be applied to more complex morphologies, such as structures assembled from polyhedral building-blocks [33], or hierarchical assemblies such as micellar systems [187, 306, 307] or virus capsids [289,290], wherein the building blocks assemble into larger structural sub-units that arrange into superstructures. In such cases, the representative structural pattern is given by the positions of assembled sub-units, rather than the individual building blocks (detailed in Example 1 in section 7.3.5, below).

Many complex structures cannot be described by positions alone, and require information regarding building block sizes, shapes and orientations. Such structures can be described by "volumetric data," or "voxel data" (i.e., d-dimensional pixel data), which is represented numerically by a collection of weights or pixel intensities for cells in a grid that spans space. This representation is particularly apt for describing the microphase-separated morphologies assembled from systems of tethered nanoparticles and block copolymers, where spatial density maps for the aggregating species may resemble sheet-like or network domains [307, 308, 306, 309] (see Fig. 7.4). Voxel data captures the essential structural features of these systems, whereas a pattern based on the positions of individual particles

within the superstructure does not. The same rule applies to many other types of structures for which the bulk shape is more important than the underlying particle positions, including all types of phase-separated structures, many complex biological structures such as proteins and macromolecules [310,311], and large but finite (aka "terminal") nanoparticle assemblies. Shape descriptors are typically sufficiently flexible to use either voxel data or point cloud data as an input.

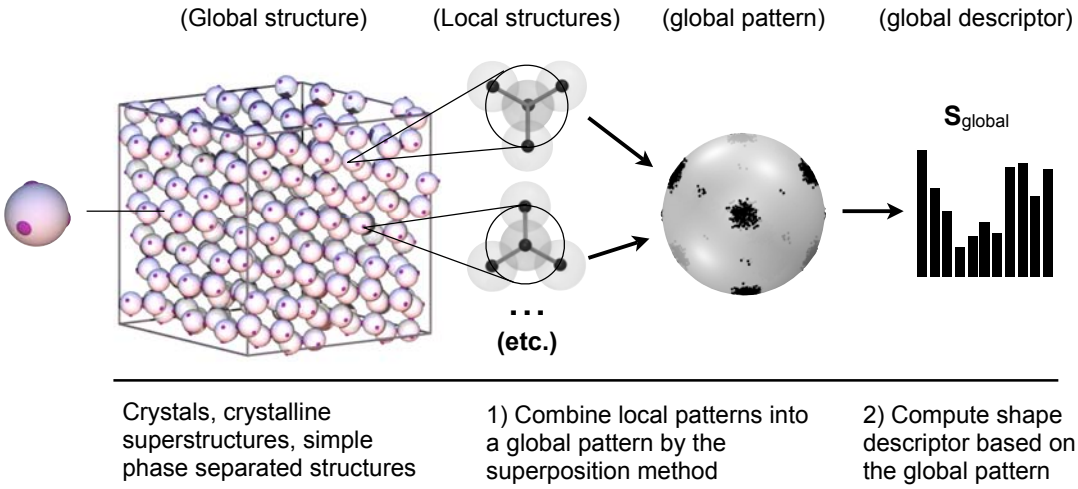


Figure 7.3 Extracting global patterns using the superposition method. For structures with long-range orientational ordering, such as the diamond structure formed by tetragonally patterned patchy spheres depicted in the panel [29], a global pattern is extracted by translating all local clusters [312] or density maps to a common origin. Here, the local structures are represented by particle positions, but more complex representations are possible. A global shape descriptor is then computed for the resulting finite structure.

Shape descriptors are typically constructed to describe finite objects. Thus, when describing global structures such as crystals or bulk disordered systems, local shapes must first be extracted from the infinite system and then combined into finite local patterns that reflect the "global pattern" for indexing. The types of global patterns that we create depends on the structural properties of the system. For structures with long-range orientational ordering, such as crystals and quasicrystals [313], the shape and spatial orientation of local clusters within the system are highly correlated. Thus, a global structural pattern can be obtained by translating all local shapes to a common origin [132], a scheme that we denote as the "superposition method." The visual depiction of the superimposed structures is known as a "bond order diagram [312]," an example of which is depicted for the diamond structure formed by patchy particles [29] in Fig. 7.3. For crystals with multiple particle types, independent global descriptors can be created for each type independently, and a combined

descriptor can be created. Global descriptors based on orientational ordering are applicable to crystalline structures in general, including phase-separated systems arranged in crystalline superstructures [298, 314], where the neighbor directions are computed for the centers of the micelles, cylinders, etc. rather than the individual particles. Some non-crystalline globally-ordered microphase-separated structures, such as layered or network structures, can be described by superposition as well, where global patterns are built up from local density maps, rather than from local point clusters. This reflects the fact that the probability density of observing particles in particular spatial directions within these morphologies is often non-uniform.

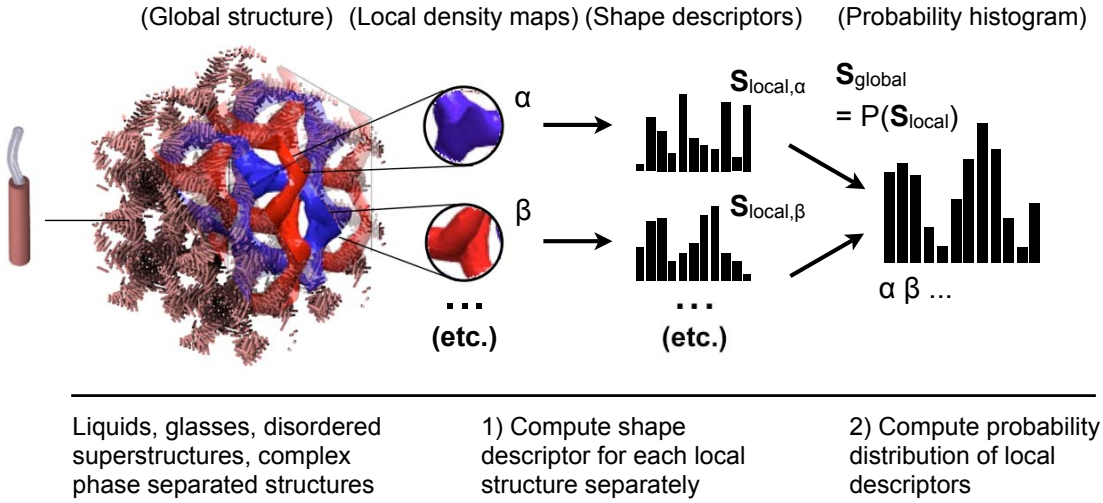


Figure 7.4 Extracting global patterns using the probability distributions method. For structures without long-range orientational ordering, or complex global structures with many different characteristic directions, a global pattern can be built up from the probability distribution of local patterns. The double gyroid formed from tethered nanorods [309], which falls into the latter category, is characterized by computing the distribution of local nanoparticle density maps sampled throughout the structure. The red/blue color scheme emphasizes the bicontinuous nature of the interpenetrating network.

For systems without long-range orientational ordering such as liquids, glasses and amorphous solids, a different strategy must be employed, since, in such cases, the superposition of local structures inherently yields a uniform pattern. In such cases, rather than combining neighbor directions or density maps by superposition, we compute a probability distribution of local patterns. The probability histograms for different structures can then be compared to obtain a measure of similarity between global structures (Fig. 7.4). Computing probability distributions is also useful for certain complex orientationally-ordered structures, for which the superposition of local density maps becomes non-distinguishing due to the presence of

many different characteristic directions within the structure. An example of such a structure is given by the double gyroid structure composed of tethered nanorods [309] shown in Fig. 7.4.

Shape Descriptors

Once we have extracted a representative structural pattern from our particle system, we can compute a shape descriptor to represent the pattern numerically. Depending on the intended application, different shape descriptors may be best suited to describe a particular structural pattern, and this information should be considered when deciding which shape descriptor to compute. Below is a short list of desirable shape descriptor properties within the context of assembled systems:

- **Robustness:** the degree of sensitivity to structural defects or random thermal noise. Some shape descriptors have an inherent data-smoothing mechanism, whereas others require preprocessing to effectively process thermal data.
- **Invariance:** the ability for shape descriptors to remain invariant (i.e., unchanged) under certain mathematical transformations. Invariance under scaling and translations is typically desirable. Additionally, descriptors may be invariant under rotations, mirroring operations, or similarity transformations. Rotation invariance is the most important of these properties for particle systems. For descriptors without rotation invariance, we often must align or "register [315, 316]" objects prior to matching.
- **Efficiency:** the computational effort required to calculate the descriptor. For certain applications, CPU time and memory costs may be a limiting factor for choosing a shape descriptor. For example, efficiency may be an important factor for on-the-fly order parameter calculations that occur during a molecular simulation, whereas for offline data analysis it may be irrelevant. Often, there is a direct tradeoff between computational cost and accuracy.
- **Comparability:** the ease of matching. Shape descriptors should yield similar results for similar structures and different results for different structures. Shape descriptors should be constructed such that similarity is easy to quantify. The numerical similarity should directly reflect the physical similarity between the shapes used to construct the descriptors.

Below, we review some shape descriptors from the computer science shape matching literature. Since shape matching is a broad field, we focus on the subset of methods that are best suited for assembled systems. For a general review of some relevant shape matching methods, see references [317, 318, 319].

Point-Matching Descriptor: For relatively simple structures such as small clusters of atoms, molecules, or nanoparticle/colloidal building-blocks, we can use the particle positions themselves (or a corresponding density map) as a shape descriptor (Fig. 7.5a). Matching

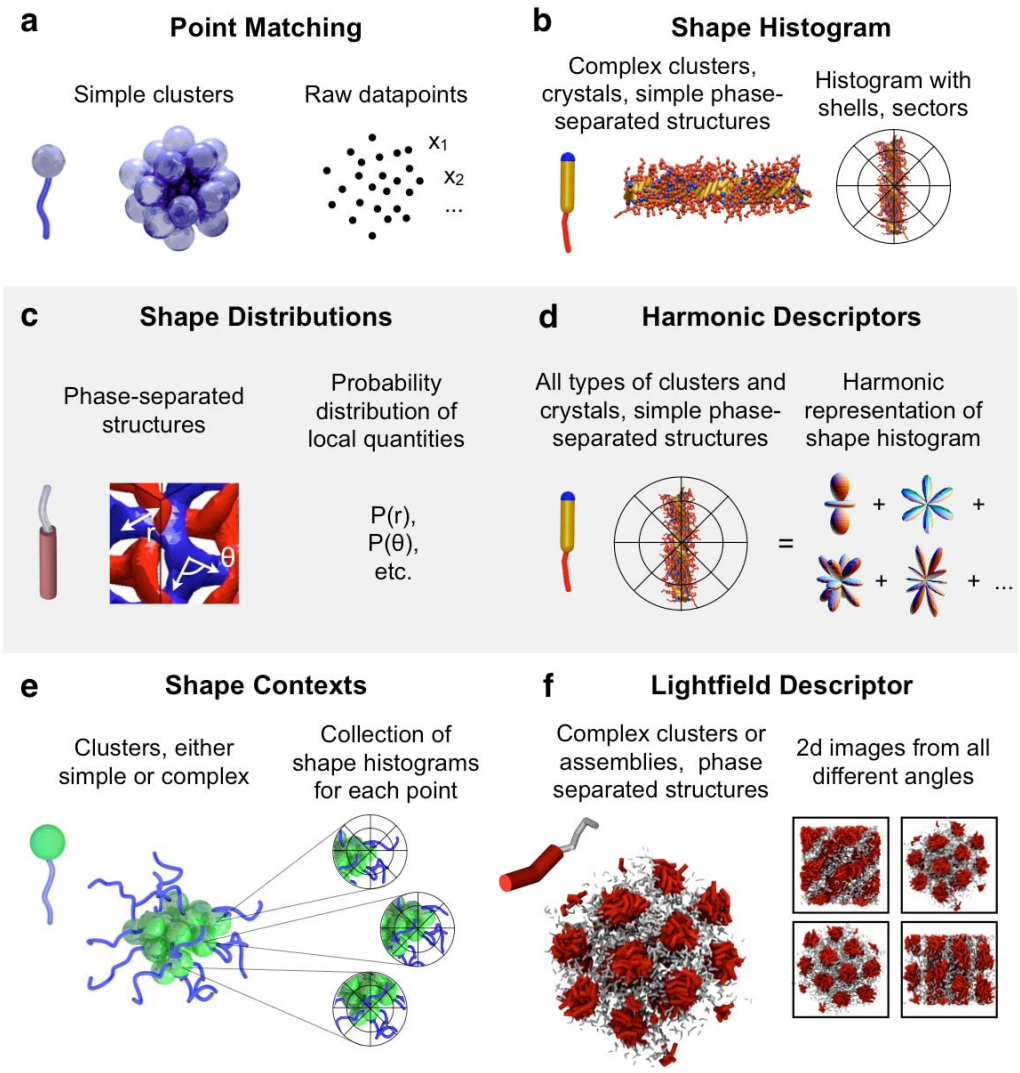


Figure 7.5 Depiction of six different shape descriptors applied to self-assembled systems. (*a*) The point matching descriptor [320, 315]. Descriptor components are given trivially by particle positions or density maps. (*b*) The shape histogram descriptor [321]. The structure is indexed into a spatial histogram consisting of shells and sectors. (*c*) Shape distribution descriptors [322]. The probability distribution is computed for various local measurements, such as the distance or angle between surface points. (*d*) Harmonic descriptors [323, 324, 325, 326]. The shape histogram is decomposed into a convenient harmonic representation, which can be used for rotation-invariant matching. (*e*) The shape contexts descriptor [327]. A coarse histogram is created for each point on the structure. The descriptor is given by the collection of sub-descriptors for each point. (*f*) The lightfield descriptor [328]. Images or projections are constructed from several different vantage points and indexed into individual shape descriptors. The overall descriptor is given by the collection of sub-descriptors for each image.

for this scheme is often based on the root-mean-square (RMS) difference between points, and thus the scheme itself is sometimes referred to as "RMS matching." Point matching

schemes were applied to early attempts at shape matching for macromolecules [329], and more complex variations have since been implemented for proteins [330]. Point matching schemes have the advantage of being conceptually trivial; however, there are many subtleties associated with these schemes that should be considered. First, the descriptor requires an assignment step to determine the optimal correspondence between points in compared structures, which is used to re-order the coordinates in the shape descriptors accordingly. Also, since the descriptors are sensitive to scale, position, and orientation, structures must first be normalized and registered unless the orientations are known beforehand, or rotation-dependent matching is desired. Depending on the application, shapes may be registered based on rigid alignment, or other constraints. Since both assignment and registration are computationally expensive (i.e. they scale poorly with the number of points, n) point matching descriptors should be avoided unless (1) the number of atoms, molecules, or building blocks that make up the structure is small, (2) matching is required for only a few structures, or (3) registration is not required.

Shape Histogram: Another conceptually simple shape descriptor that has been applied to molecular database searches is known as the "shape histogram" [321] (Fig. 7.5b). This descriptor is based on a density map of the structure on a polar or spherical grid. Shape histograms are best suited for describing structural patterns that can be broken down into concentric shells, such as nanoparticle clusters, proteins and macromolecules. Shape histograms are also well suited for indexing global patterns created by the superposition method, as outline above, and can index structures with orientational ordering such as crystals or quasicrystals, and simple microphase separated structures such as layered phases or network structures. The shape histogram has the advantage over the point matching method that no assignment step is required, since the ordering of points is lost during binning. Additionally, the grid resolution can be adjusted to provide a variable degree of coarse-graining. Like the point matching method, the shape histogram requires registration to match non-aligned objects, unless only radial bins are used (i.e., the angular grid resolution is set to zero). However, shape histograms may lose their discerning capabilities without an angular component. If n is large, the cost of registration can be significantly reduced by aligning the histograms themselves rather than the underlying structures.

Shape Distributions: For many applications, registration is too costly and we require rotation-invariant descriptors. A simple, yet powerful method for creating invariants, known as the "shape distributions" scheme [322](Fig. 7.5c), involves computing distribution functions for simple rotationally-invariant local metrics. Such local metrics are defined based on object surfaces; thus this method is best applied to structures with clearly defined, yet

distinguishable, surfaces, such as microphase-separated structures formed by block copolymers [194, 285] or tethered nanoparticles [253, 210, 331, 306] (see, for example, Fig. 7.4). The shape distribution "D2" is defined as the probability distribution of the distance between pairs of surface points. Another similar distribution "A3" is defined by the probability distribution of angles formed by triples of surface points. Similar distributions are defined for higher numbers of points. The distributions D2 and A3 are similar to the radial distribution function $g(r)$ and angular distribution function $a(\theta)$, respectively, although usually only surface particles are considered. Like $g(r)$ and $a(\theta)$, shape distributions are too coarse to distinguish between similar shapes, such as small polyhedral clusters.

Harmonic / Invariant Moment Descriptors: A more complex, but more powerful method for computing invariant descriptors is to compute the harmonic transform of the shape histogram. By disregarding the phase information, we obtain descriptors that are invariant under rotations (Fig. 7.5d). Like the shape histogram, harmonic descriptors are versatile and can be applied to a wide range of structures including complex nanoparticle clusters, proteins and macromolecules, and crystalline or microphase-separated structures. The method by which we compute the harmonic transform of the shape histogram depends on the underlying basis. Invariants can be obtained for shapes on the circle [323](θ-dependence), sphere [324](θ, ϕ-dependence), disk [325] (r, θ-dependence) and ball [326](r, θ, ϕ-dependence). On the unit circle or sphere, the harmonic descriptors are called "Fourier descriptors," whereas on the disk or ball, the descriptors are known as "Zernike descriptors." The implementation of these methods for complex assembled systems is described in detail elsewhere [332]. Harmonic descriptors exhibit an inherent data smoothing mechanism that leaves them better-suited for describing small polygonal or polyhedral clusters than the shape histogram, which is prone to error without sufficient averaging. This property, combined with the property of rotational-invariance, makes harmonic descriptors ideal for describing orientationally-disordered global structures, such as liquids, glasses and certain microphase-separated structures, via the probability distributions method. Harmonic descriptors also contain additional frequency-dependent information regarding the symmetries of the structure. These unique properties of harmonic descriptors have already been successfully applied to constructing orientational order parameters for small clusters and simple crystals [265, 132].

Shape Contexts: It is fairly common in the context of self-assembly experiments and simulations to encounter nearly-ideal assembled structures with localized defects. Thus, it is often desirable to distinguish between local structural dissimilarities that arise due to defects, and "overall" differences in the structure. A brute-force solution to this problem is to explicitly include defective structures in the library of reference structures such that

they may be identified directly [187]; however, his requires *a priori* knowledge of the entire space of potential defective structures. Obtaining such knowledge may be intractable for complex assemblies with many degrees of structural freedom, or unmapped systems whose local motifs have not yet been thoroughly studied. A more general solution is to apply a "partial matching" scheme, such as the "shape contexts" method [327, 333], which is capable of matching structures independently of local defects, as well as identifying such defects (Fig. 7.5e). The shape contexts method combines elements of the point matching scheme with the shape histogram descriptor. Here, a separate shape histogram is computed for each sample point in the structure, where the coordinate system is centered at that point. The points in the query structure are then assigned to their corresponding points in the reference structure by optimizing the match between shape histograms. Outlier points that don't correspond well (i.e., local defects) can be excluded to obtain a partial match, or used to identify the defects. Shape contexts can be applied to any system where local defects might arise, such as atomic or molecular clusters, micro or nanoscale assemblies, or biological structures. Since shape contexts are based on the shape histogram, they have the same limitations when indexing structures with a small number of sample points locally.

Lightfield Descriptor: The shape contexts descriptor is just one example of the more general method for creating new powerful descriptors by combining simpler sub-descriptors . A similar method based on combining sub-descriptors is given by the light-field descriptor [328], which involves projecting 3D structures onto 2D images from 20 vantage points at the vertices of a dodecahedron. This process effectively simulates the act of viewing a structure from different angles by eye, giving the lightfield descriptor its name (Fig. 7.5f). The lightfield descriptor can thus be applied to microphase-separated structures, nano/colloidal scale assemblies, or other structures that can be effectively identified by the trained eye. Each of the 20 2d images is indexed by a 2d descriptor, and assignment is performed for pairs of these descriptors for compared structures to optimize correspondence. In practice, many initial rotations of the dodecahedron are attempted to minimize error due to small offsets in the spatial orientation.

Other Possible Descriptors: In addition to the shape descriptors outlined above, the shape matching literature defines numerous potentially useful descriptors that we have not mentioned here. Some intriguing possibilities include graph based descriptors [334, 335, 336], descriptors based on reflective symmetries [337], and methods based on the similarity of slices of objects [338]. Several structural metrics from the condensed matter literature might also serve as useful shape descriptors for some applications. For example, in the realm of global structures, diffraction patterns, radial distribution functions, or orientation tensors

(e.g. the radius of gyration tensor or the nematic order tensor [339]) could be indexed into shape descriptors. For local structures, analysis schemes such as the common neighbor analysis scheme of reference [263] could be easily incorporated. Although many of the structural metrics from the literature may not be independently distinguishing for a wide range of problems, they may still yield useful information as part of a more general scheme through a combination of descriptors.

Similarity Metrics

The degree to which two shape descriptors match [340] is quantified by a similarity metric. Computing a similarity metric involves reducing the complex information contained in shape descriptors into a single scalar value that indicates the degree of matching. The similarity metric that best suits a particular application depends on both the shape descriptor and the intended physical application. Some desirable properties of similarity metrics are listed below:

- **Metric Behavior:** the ability for a similarity metric to give a value that is proportional to the physical match between the structures. Some similarity metrics satisfy the triangle inequality [340] (i.e., $M(\mathbf{S}_A,\mathbf{S}_B)+M(\mathbf{S}_A,\mathbf{S}_C)\geq M(\mathbf{S}_B,\mathbf{S}_C)$, where M is a similarity metric, and $\mathbf{S}_A$, $\mathbf{S}_B$, $\mathbf{S}_C$ are shape descriptors) and are thus truly metrics, whereas others do not and can be considered pseudo-metrics. It is typically desirable for similarity metrics to range smoothly with the difference between structures.
- **Normalization:** the range of possible matching values for a given matching scheme. For many condensed matter physics applications, we desire similarity metrics that range from 0 to 1 for use as pseudo-order parameters. While many similarity metrics do not vary naturally from 0 to 1, they can often be changed by simply shifting and scaling the interval that defines an ideal and worst-case match. In practice, there is little difference between this type of pseudo order parameter and a standard order parameter in terms of the underlying physics.
- **Specificity:** the degree to which a similarity metric highlights specific differences between shape descriptors. For some applications it is desirable to give more weight to specific important differences between the descriptors.

Often, similarity metrics are based on simple geometric functions, such as the Euclidean distance or vector projection between shape descriptors, which are typically represented as long vectors. Whereas similarity metrics based on the Euclidean distance are particularly common in the shape matching literature [303], schemes based on the vector projection are more commonly (implicitly) applied throughout the condensed matter literature [265, 132, 57]. In practice, the mathematical form of the similarity metric is typically of little consequence; virtually any function can be chosen, provided it ranges smoothly as the

shapes become physically different. In some specific cases, specialized similarity metrics are designed to be used in conjunction with particular shape descriptors. The shape histogram scheme described in section 7.3.3 above utilizes a specialized quadratic form distance function for matching [321], which accounts for mismatches arising from near-misses that occur due to the discrete nature of the histogram bins. The P_2 Legendre polynomial shown in Fig 7.1 is an implicit example of a specialized similarity metric, specifically designed to match the angles of rod-like particles with the ideal angle given by the global director [341].

7.1.2 Example Applications

In this section, we demonstrate the application of shape matching techniques to a few representative problems from our studies of self-assembly. For simplicity, we use the same shape descriptor and similarity metric for all of the examples. Since our goal here is to demonstrate the basic usage of shape matching techniques, our examples should be considered proofs-of-concept rather than optimal solutions to the problems. Additional examples of applications of other shape descriptors to self-assembly may be found in References [342, 332].

Prototype Shape Matching Scheme

For our example problems, we use the 3d Fourier shape descriptor [324], which is the harmonic descriptor defined for patterns on the sphere, $[\theta, \phi]$. We choose this descriptor because it is closely related to the spherical harmonics bond order parameters introduced by Nelson and coworkers [265, 132], and thus many readers will already be partially familiar with them. The basic idea behind the 3d Fourier descriptor is to decompose a 3d structure into one or more patterns on the 2d surface of a sphere, and represent these patterns mathematically by computing the discrete spherical hamonics transform (DSHT). This method of representing a pattern as its harmonic transform is analogous to the way that 1d signals along the perimeter of the circle can be described by their discrete Fourier transform (DFT).

How we extract the patterns on the sphere depends on how data is represented. For simplicity, we use a minimal data representation based solely on particle positions (i.e., point cloud data) for all of our examples; however other types of data, such as volumetric data, can also be easily treated by Fourier descriptors. For our examples, we describe particle structures as patterns on the sphere by (1) translating the structure to the origin, (2) grouping all positions within a radial shell r_s and (3) converting each position $\mathbf{x}$ into its angular direction relative to the origin $[\theta(\mathbf{x}), \phi(\mathbf{x})]$. This is repeated for all n_s radial shells required to describe the full 3d structure, giving n_s patterns on the sphere for each structure.

For each pattern on the sphere, the Fourier coefficients of the DSHT are given by:

$$\mathbf{q}_\ell = \frac{1}{n}\sum_{i=1}^{n} Y_\ell^{m*}[\theta(\mathbf{x}_i),\phi(\mathbf{x}_i)] \quad m = -\ell, -\ell+1, ...\ell. \tag{7.1}$$

The term Y_ℓ^m is a set of spherical harmonics with angular frequency ℓ. The coefficients $\mathbf{q}_\ell$ are vectors with $2\ell+1$ complex components. Although the Fourier coefficients in their complex number form are rotationally-dependent (i.e., their value depends on the spatial orientation of the underlying pattern), we can convert them to their rotationally-invariant form by computing the magnitude of each coefficient. The invariant circular coefficients are given by:

$$|\mathbf{q}_\ell| = \sqrt{\frac{4\pi}{2\ell+1}\sum_{m=-\ell}^{\ell} |q_\ell^m|^2}. \tag{7.2}$$

The Fourier invariants are positive real numbers. Although the coefficient magnitudes themselves can be used directly as order parameters [132], incorporating them into a shape descriptor is often more powerful, since we can compare shapes based on a variety of frequencies and lengthscales. To create a descriptor from the Fourier coefficients, we simply combine the desired $\mathbf{q}_\ell$ or $|\mathbf{q}_\ell|$ into a long vector. For example, a general rotation-invariant shape descriptor that is applicable to patterns on the sphere over a range of symmetries is given by:

$$\mathbf{S}^{F3}_{shell_i} = < |\mathbf{q}_{\ell_{min}}|, |\mathbf{q}_{\ell_{min}+1}|, ...|\mathbf{q}_{\ell_{max}}| > . \tag{7.3}$$

The range of frequencies can be adjusted to obtain a desired level of resolution. For our examples below, we use $\ell_{min} = 4$ and $\ell_{max} = 12$. Since each Fourier descriptor describes the pattern for a given shell, we must combine the Fourier descriptors for each shell to describe the overall shape:

$$\mathbf{S}^{F3} = < \mathbf{S}^{F3}_{shell_1}, \mathbf{S}^{F3}_{shell_2}, \ldots, \mathbf{S}^{F3}_{shell_{n_s}} > . \tag{7.4}$$

For our example applications, we will use a simple similarity metric based on the Euclidean distance $|\mathbf{S}_i - \mathbf{S}_j|$ between harmonic shape descriptors:

$$M(\mathbf{S}_i, \mathbf{S}_j) = 1 - 2\left(|\mathbf{S}_i - \mathbf{S}_j| / |\mathbf{S}_i| + |\mathbf{S}_j|\right). \tag{7.5}$$

This similarity metric is proportional to the Euclidean distance between shape descriptor vectors, but is normalized such that vectors that match perfectly give a value of 1, while vectors that are perfectly anticorrelated give a value of -1. Vectors with no directional correlation (i.e., that are orthogonal) give a value of 0. This normalization allows us to make a clearer analogy between our matching scheme with a typical order parameter; however,

only the relative value of the similarity metric is relevant and the normalization is merely a matter of convenience.

Example 1: Micellar Crystal Structures

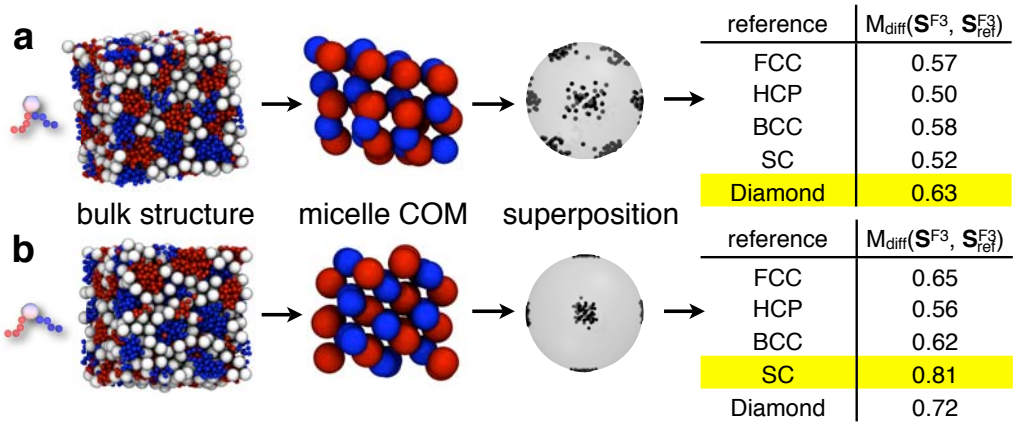


Figure 7.6 Identification of global crystalline structures for a system of ditethered spheres [298, 314]. (*a*) A crystal formed in the ditethered nanosphere system where the planar angle between tether attachment is 30 degrees. Ignoring chemical specificity of the tether micelles, the structure best matches the ideal diamond lattice. (*b*) A crystal formed by the ditethered nanosphere system with planar angle between tether attachment of 60 degrees. Ignoring chemical specificity of the micelles, the structure best matches an ideal simple cubic structure. In both cases, the micelle centers are extracted using a Gaussian filter, and matching is based on the global superposition of local patterns (section 7.3.3).

A straightforward application of shape matching techniques to particle systems is to identify unknown structures by searching a database of known reference structures. Structures are identified by the known structure that gives the best match. Structure identification can be performed for either local structures or for global samples. As a simple example of structure identification for a global sample, consider the ditethered nanosphere system of references [298, 314], which microphase separates into spherical micelles. The micelles themselves pack into an ordered binary crystalline superstructure. Depending on the state point, the system forms different crystals, as shown in Fig. 7.6. The structural pattern that represents the different crystals is obtained by identifying the micelle centers of mass, which comprise the set of positions that describe the system. The micelle centers of mass are determined by creating a density map (i.e., a voxel representation) for the aggregating polymer tethers and then applying a Gaussian filtering algorithm adapted from the colloidal science literature [294, 291] to identify the spheroid centers. Since the superstructure has long-range orientational ordering, a global pattern is given by the superposition of local patterns (see

Fig. 7.3). The pattern for the unknown crystal is compared to those for several standard candidate crystals. For each pattern we compute the 3d Fourier descriptor $\mathbf{S}^{F3}$ described above, with rotationally-invariant coefficients for a single shell, $n_s = 1$. Using this method, the patterns are compared independently of spatial orientation over a single length scale used to construct the local clusters. The unknown crystal is identified by the reference structure that gives the best match. The structures in Fig. 7.6a,b are identified as diamond and simple cubic, respectively, where we do not consider the chemical specificity of the two types of micelles. Notice that the best match does not necessarily give a value that approaches 1; such deviations are common when comparing thermal systems to mathematically perfect reference structures, as we have done here. The micellar system under investigation exhibits thermal disorder as well as polydispersity in the shape and size of the micelles, and thus particle positions deviate from the ideal lattice points. Oftentimes, comparing to reference systems that exhibit similar levels of noise may provide clearer results.

This type of database search has already been applied to particle systems in the context of proteins and macromolecules [329, 343, 330, 344, 345, 310, 311]. Although database searches have only been applied in limited cases to assembled systems [187,309], many standard local structure identification schemes in the condensed matter literature bear a strong resemblance to shape matching identification schemes. For example, the common neighbor analysis (CNA) scheme of reference [263] involves constructing numerical fingerprints for pairs of atoms based on their local neighbor configurations, and identifying local clusters by matching the distribution of fingerprints with those for ideal structures. In the language of shape matching, the collection of CNA fingerprints can be considered a shape descriptor, and the catalogue of ideal fingerprints can be considered a database of reference structures. A similar identification scheme is given by the bond order parameters of reference [132]. Here, particular local structures with strong symmetries, such as small ordered clusters of spherical particles, can be identified by finding structures with bond order parameters that exceed a particular threshold [346]. In this case, the bond order parameters represent shape descriptors, and the threshold values act implicitly as similarity metrics, since the ideal structures are known to have high values of the bond order parameters.

Example 2: Icosahedral Clusters of Tetrahedral Particles

As mentioned in the previous example, a common application of structural characterization schemes is to identify local motifs within a global system. Examples include finding locally stable clusters in liquids [263, 33], colloids and gels [347] and nanoparticle superstructures [187, 309], and identifying structural defects in, or grain boundaries between,

crystalline domains, such as in dense colloids [293]. Often, these local structural characteristics can be directly related to the thermodynamic, mechanical, or other properties of the system.

When detecting local structures in systems without long-range orientational order (i.e. "disordered" systems), we often encounter structures that are not present in our reference library. A structure that does not match with those in the reference library within a certain threshold is considered "disordered," or unimportant [187, 309]. The threshold must be chosen carefully; in thermal systems, an overly-stringent cutoff value might cause a matching scheme to miss highly-ordered structures perturbed slightly from their ideal configurations, whereas an overly-permissive cutoff can misidentify highly disordered structures. In most cases, a sufficiently rigorous cutoff can be defined such that its value does not affect the qualitative results.

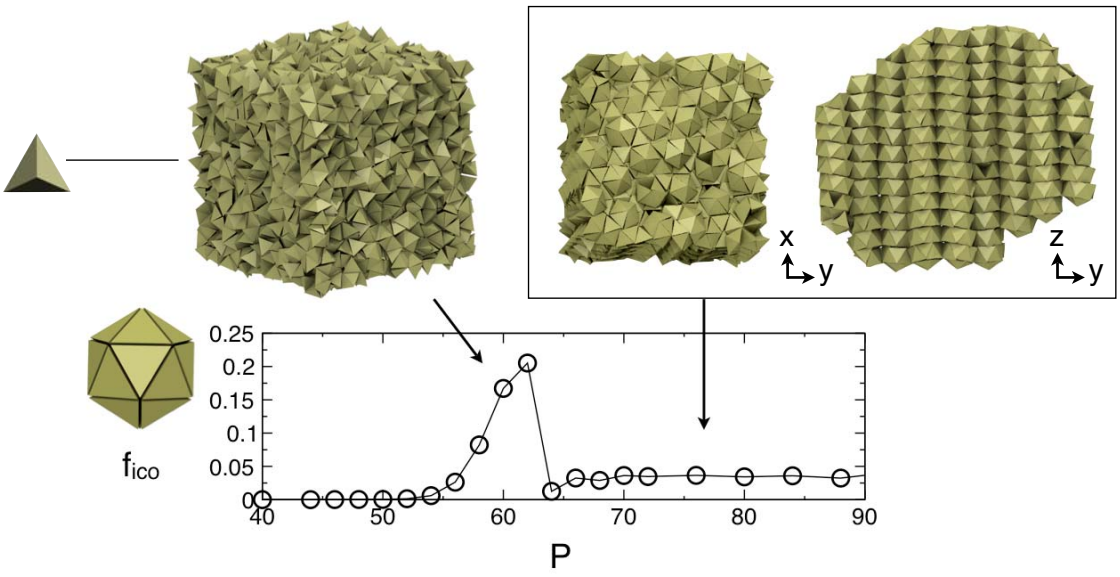


Figure 7.7 Icosahedral clusters in the hard tetrahedron system [33]. As the pressure and the corresponding density increase, icosahedra grow more prevalent until the system transforms into a dodecagonal quasicrystal at $P \approx 62$, at which point the number of icosahedra vanishes.

As an example of identifying ordered local structures in an otherwise disordered system, consider the hard tetrahedron fluid studied in reference [33] (Fig. 7.7a). In this system, an important local motif to both the fluid and the glass, originally identified by visual inspection, is the icosahedron formed by 20 tetrahedra sharing a common vertex. To identify icosahedra in the system, we first cluster all sets of 20 tetrahedra in the system that share a common vertex. The structural pattern for each cluster is defined by the directions of vectors drawn from the center of the cluster through the face of each of the 20 tetrahedra, which for an ideal

icosahedral cluster results in a dodecahedron. Any local cluster i that matches the shape of a dodecahedron with a value of $M_{cut}(\mathbf{S}_i^{F3}, \mathbf{S}_{dodecahedron}^{F3}) > 0.9$ is considered to be in an icosahedral motif. Fig. 7.7 shows the fraction of tetrahedra that participate in at least one icosahedron as a function of pressure. Icosahedra are relatively common in the tetrahedral fluid (below $P = 62$) and become more prevalent with increasing density, persisting into the glass if the fluid is compressed too quickly. As the fluid transforms into a quasicrystal at $P \approx 62$, the fraction of tetrahedra in icosahedra decreases drastically, and vanishes for the ideal quasicrystal without thermal fluctuations. Although the value of M_{cut} may affect the absolute number of icosahedra, the same underlying physical transition is captured for any reasonable value.

Example 3: Assembly of a Helical Ribbon

Another standard application of structural metrics is to track structural transitions, either as a function of time or a changing reaction coordinate. This is typically accomplished by monitoring either an order parameter or correlation function as the system goes through a transition. Tracking structural transitions is important for a wide variety of applications, including elucidating thermodynamic transitions [57, 348, 349, 350, 351] and assembly pathways [29, 352, 353, 268]. Many of the advanced molecular simulation techniques used to study transitions [156, 157, 60, 354, 158] rely on structural metrics in the context of pseudo-reaction coordinates [157], biasing parameters [156], and collective variables [158] to guide the statistical sampling algorithm. Standard order parameters have been devised for various types of ordering, including bond orientational ordering [48, 54, 132, 355], liquid crystalline ordering [341, 262] such as nematic [356] and smectic [349] phases, chiral ordering [264], and helical ordering [357]. Time correlation functions based on these types of order parameters have been applied to creating structural "memory" functions for glassy liquids [138, 139] and ordered motifs attaching to a growing quasicrystal nucleus [30].

As a simple example of using shape descriptors to create an order parameter, consider the ribbon-like bilayer composed of laterally tethered nanorods studied in reference [301], and shown in Fig. 7.8. The initial sheet or ribbon is unstable and eventually relaxes into a stable helical structure. We can track this structural transition by matching the shape of the sheet at a given time t with the final, fully-equilibrated helical structure: $M(\mathbf{S}_t^{F3}, \mathbf{S}_{helix}^{F3})$. Since the structure is 3-dimensional and has radial dependence, we use a Fourier descriptor with $n_s = 6$ radial shells: $r_s = 10\sigma, 30\sigma ... 110\sigma$, where σ is the distance unit corresponding to a Lennard-Jones particle diameter. Since the sheet only changes in terms of its twist in space, we save computational effort by only considering points along the backbone of the

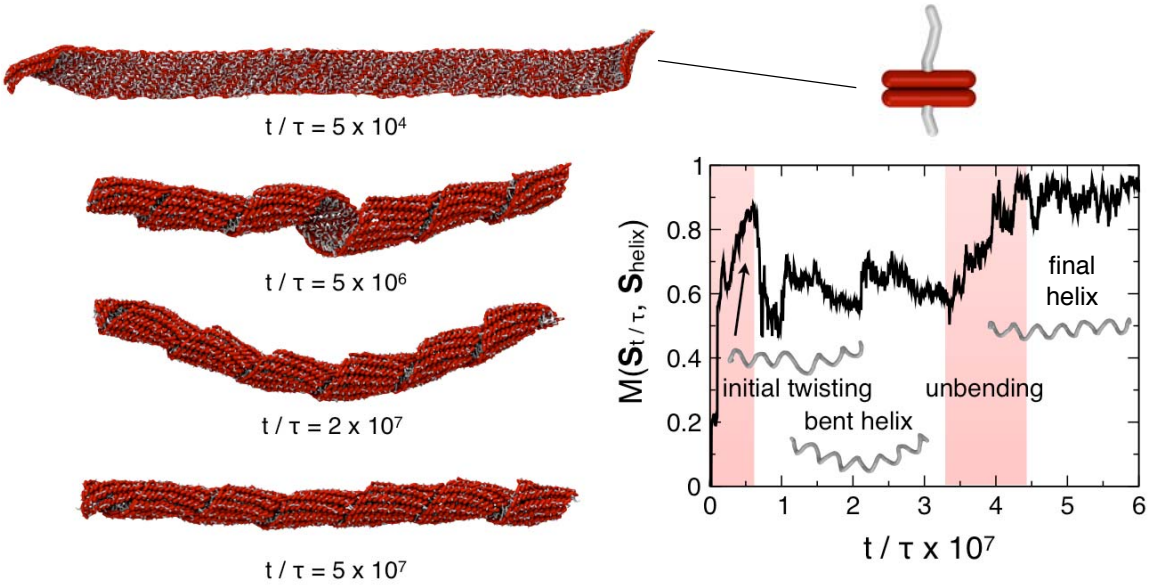


Figure 7.8 Assembly of a helical sheet composed of laterally tethered nanorods [301]. The rods form a bilayer with long attractive tethers on one side, and shorter attractive tethers on the other. As time progresses, the sheet folds into a helix to maximize the favorable energetic interactions between the longer tethers. The matching order parameter $M_{dist}(\mathbf{S}_{t/\tau}, \mathbf{S}_{helix})$ compares the structure at time t with the shape of the final ideal helical structure.

sheet. Fig. 7.8a shows the helical order parameter as a function of time for a long molecular dynamics run. We observe that the sheet begins to twist from both ends simultaneously, which gives rise to a defect at the center of the helix, where a mismatch in the periodicity between the two ends occurs. This results in a tendency for the structure to bend to close the defect. The bend persists for many millions of time steps before annealing into a defect free helix at around $t/\tau = 4.5 \times 10^7$. This behavior is well captured by matching the overall shape of the structure, but is not captured by the more standard H_4 descriptor, applied in the original reference, which only measures the degree of helical ordering and gives an essentially constant value for all times after the completion of twisting at $t/\tau \approx 7 \times 10^6$ [301]. Using H_4 alone, it would appear that the structure is fully formed at this early time, which does not capture the important defect removal behavior, which can also be observed by visual inspection.

7.1.3 Future Outlook

Beyond identifying local and global structures and tracking structural transitions, there are many more applications of shape matching. In this section, we briefly review some areas

in which we are currently applying shape matching for studying self assembly. Additional details may be found in Ref. [342] and in the individual references cited below.

On-The-Fly Structure Identification: For many assembly applications, such as Bottom-Up-Building-Block-Assembly (BUBBA) [358], we are interested in cataloguing unique structures. When enumerating unique structures it is not typically necessary (or feasible) to define a library of reference structures *a priori*, as we did for examples 1 and 2 above. Rather, the reference library can be compiled on-the-fly as new structures are encountered (see Fig. 7.9a). Each new structure is given a unique identifier, and structures that are duplicates are labeled with the same identifier. In addition to cluster enumeration schemes, this type of algorithm can potentially be applied to automatically detect regions of unique ordering in structural phase diagrams.

Space/Time Correlation Functions: In example 3 above, we demonstrated how shape matching could be used to track a structural transition as a function of time, or a reaction coordinate. Another common application of structural metrics is to characterize how structures change in space. In the context of shape matching, this involves choosing structures from different points in the system, rather than ideal structures, as reference structures. Spatial correlation functions are often used to measure structural "correlation lengths." In the condensed matter literature, structural correlation functions have been defined for crystal-like ordering in 2d [48, 54] and 3d [132, 57], nematic ordering [359], and many other more specialized types of ordering. More specialized types of spatial correlation functions have been widely applied as well. One example is the $q_6 \cdot q_6$ scheme of references [57, 62], which detects ordered crystal nuclei based on spatial correlations between local bond order parameters. This scheme can be adapted to identify crystal nuclei in general by replacing q_6, which is only sensitive to particular crystal structures, with other shape descriptors that are applicable to a particular crystal under investigation. Fig. 7.9b depicts the formation of a diamond-structured crystal nucleus (yellow) in a system of patchy particles, identified by replacing q_6 with the $\ell = 3$ Fourier coefficient, q_3 [29].

Structure Grouping and Classification: The field of self-assembly involves a wealth of particle building blocks and the assemblies they form; thus it is sometimes useful to categorize or classify structures based on particular structural features. For example, reference [11] ranks different building blocks for self-assembly based on their shape anisotropy. Shape matching methods can provide numerical metrics by which to classify structures. Structures can be ranked based on the degree to which they exhibit a particular structural feature of interest, or by how well they match ideal structures exhibiting a particular feature. For

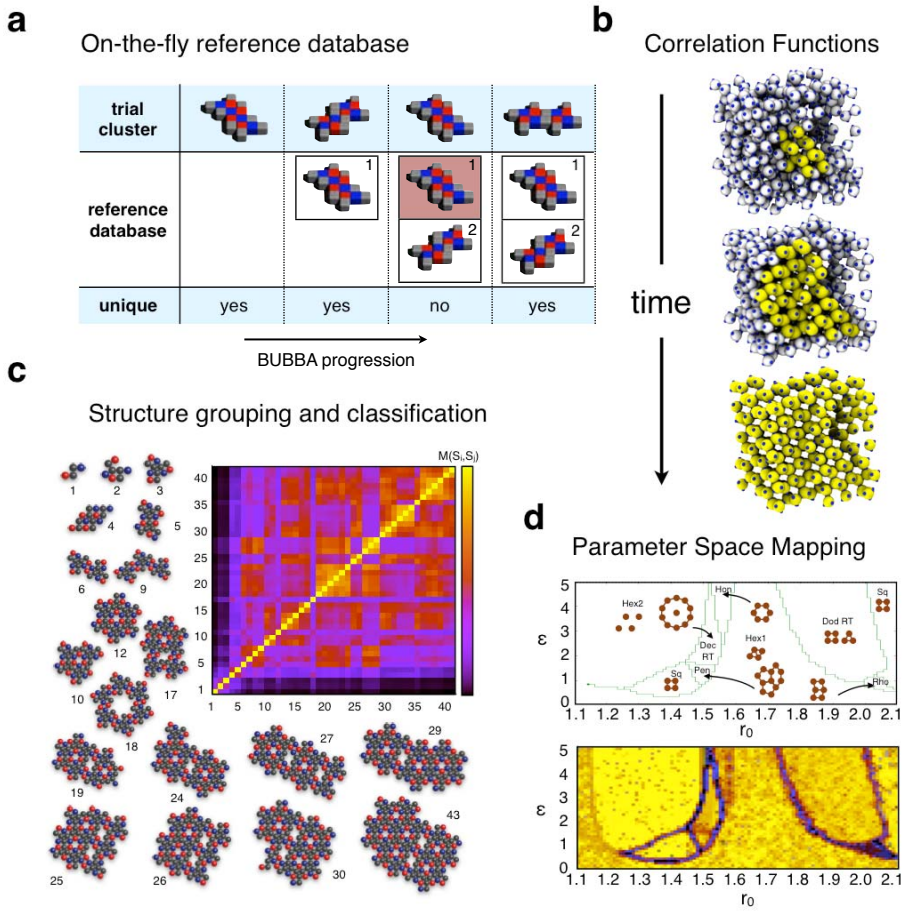


Figure 7.9 Potential uses for shape matching in assembly applications. (*a*) Searching parameter spaces for unique structures. The panel depicts the Bottom-Up-Building-Block-Assembly (BUBBA) algorithm [358]. (*b*) Computing spatial correlation functions. The panel depicts detecting a growing diamond crystal nucleus in a system of patchy particles [29]. (*c*) Structure grouping and classification. The panel depicts a similarity matrix (i.e., all of the pairwise similarity values) for 2d clusters of different sizes. Groups of similar structures are identified by bright boxes about the line y=x. (*d*) Abstract correlation functions. The panel depicts a structural phase diagram for the 2d Lennard-Jones Gauss system (top), created by visual inspection [120], compared with a phase diagram for the same system generated automatically using a shape matching algorithm (bottom).

example, structures can be ranked based on their 6-fold symmetry by computing the value of their $\ell = 6$ Fourier descriptor, which is proportional to the degree of 6-fold symmetry. Similarly, we can create groups of structures that exhibit a particular structural feature by comparing shape descriptors. One example of a technique used to visually group similar structures is given by plotting a matrix of pairwise similarity values known as a "similarity matrix" or "heat map" [360], as depicted in Fig. 7.9c for 2d colloidal clusters [361]. Groups of clusters with similar structural features produce bright blocks, indicating that clusters

within this region of parameter space match well. Grouping objects based on shape similarity has also been applied recently to macromolecules and proteins [310, 344].

Abstract Correlation Functions: Thus far, we have either extended the applicability of standard condensed matter order parameters and correlation functions by incorporating shape matching, or applied standard shape matching applications directly in the context of assembled systems. However, in addition to extending existing applications for use with assembled systems, shape matching allows us to invent new methods that have not yet been explored. For example, rather than creating correlation functions in space and time as we typically do for condensed matter systems, we can create abstract correlation functions in parameter space. Fig. 7.9d depicts a parameter space correlation function computed for the 2d Lennard-Jones Gauss system [120], which identifies structural phase boundaries (purple) by finding points in parameter space that do not match well with their neighboring points. This correlation function is able to reproduce the structural phase diagram produced in reference [120] by visual inspection of over 5000 independent configurations. This scheme is just one example of how shape matching algorithms can replace the human element in searching for target structures, and rapidly mapping parameter spaces. The ability to expedite self-assembly research by automating the study of unique systems may represent one of the most important uses for shape matching moving forward.

Summary: The example applications and shape descriptors that we have provided here represent only a small subset of the vast range of possibilities yet to be explored. In the future, the wealth of shape descriptors from the shape matching literature should be tested for different classes of particle systems to expand the scope of order parameters available to the fields of experimental and computational assembly. New abstract order parameters and correlation functions, such as the phase space correlation function of Fig. 7.9d, can be constructed to expand the algorithms used to explore new systems. More immediately, the relatively simple algorithms outlined here can be applied to existing assembled systems to enhance our ability to gain insight into the underlying physics of these complex systems.

7.2 Harmonic Order Parameters for Characterizing Complex Particle Morphologies

Order parameters based on spherical harmonics and Fourier coefficients already play a significant role in condensed matter research in the context of systems of spherical particles. Here, we extend these types of order parameter to more complex shapes, such as those

encountered in nanostructure self-assembly applications. To do so, we build on a powerful set of techniques that originate in the computer science field of "shape matching." We demonstrate how shape matching techniques can be applied to identify unknown structures and create highly-specialized *ad hoc* order parameters. Additionally, we investigate the special symmetry properties of harmonic descriptors, and demonstrate how they can be exploited to provide optimal solutions to certain classes of problems. Our techniques can be applied to particle system in general, both simulated and experimental, provided the particle positions can detected, or the particles can be accurately imaged.

7.2.1 Introduction

Quantitative measures of symmetry and order, such as order parameters and correlation functions, are often applied within the chemical sciences to study the structural properties of particle systems. Structural quantities are particularly important in condensed systems, including nano and colloidal scale self-assembly applications, where subtle differences in particle ordering can greatly effect the thermodynamic, physical, chemical, electrical, and optical properties of a material or device [362, 363, 364, 351, 297, 12, 11]. Some particularly useful order parameters, known as "bond order parameters," were introduced by Nelson and co-workers in the context of 2d and 3d simulations of point-particles [48, 132]. These order parameters have since been widely applied to both simulated and experimental systems for quantifying crystal-like ordering amongst spherical particles. Some common applications of bond order parameters include, but are not limited to, identifying small ordered clusters [132, 346, 187], constructing static [365, 138, 366, 57, 62, 367] and temporal [139, 30] correlation functions, identifying structural defects [293], and studying nucleation and growth [57, 62]. Bond order parameters have the advantage that they can give a representative value for a given structure regardless of spatial orientation. Additionally, they are robust under random perturbations due to thermal noise and highlight important rotational symmetries.

Since bond order parameters were originally designed to quantify order in small point clusters, they cannot be applied directly to complex structures or particle shapes. Thus, they fail, in many cases, to fully describe the complex structures that arise in contemporary disciplines, such as nanoscale and colloidal assembly, soft matter physics, and the biological sciences. The field of nanoparticle assembly in particular encompasses a vast range of the structural complexity that is possible for particle systems [11, 253, 12, 13]. Here, nanometer and micron-sized colloidal particles with a wide range of shapes, compositions and interparticle forces self-assemble into unique structures, such as complex crystals reminiscent of

atomistic condensed matter [33, 268, 269, 210], phase-separating structures similar to those observed for block copolymer and surfactant systems [253, 210, 284, 285, 286, 287, 288], and hierarchical assemblies that resemble certain biological structures [289, 368]. Because there are no general structural metrics or order parameters than can be applied to these systems, *ad hoc* analyses or visual inspection are often employed instead. This often yields relatively inaccurate or incomplete results when compared to statistical analysis.

Here, we generalize and extend bond order parameters for applications involving complex structures. Our primary focus is on the field of self-assembly, where we predict that these new structural characterization schemes are most immediately required and most readily applicable. However, the structural metrics that we introduce, known as "harmonic descriptors," can be applied across diverse fields, including soft matter, macromolecular and biological sciences, and other applications where complex particle structures are encountered. To derive general structural metrics for these systems, we draw strongly from the computer science field of shape matching [34, 37], which aids us both in contriving mathematical representations for describing complex structures, and in quantifying structural similarity based on these representations. Our approach involves constructing "shape descriptors" that are roughly based on a collection of many bond order parameters computed over a range of lengthscales. By quantifying the similarity between pairs of these shape descriptors, we can derive order parameters and correlation functions that are applicable to complex structures. In this respect, the order parameters that we introduce actually represent the degree to which two collections of "sub-order parameters" match. This pairwise comparative approach to constructing order parameters is necessary for applications involving complex structures, which, unlike simple clusters of spherical particles, require a range of structural metrics to describe them completely. Shape matching techniques based on the harmonic descriptors described here have already been applied to complex particle systems in the context of fast database searches for retrieving macromolecules and proteins [344, 310, 345]. We have applied similar database searches in the context of characterizing local structure in nanoparticle assemblies [187, 33]. In addition to this type of application, we demonstrate how harmonic descriptors can be applied to the a broad range of structural characterization problems to which bond order parameters have traditionally been applied for simpler systems.

This article is organized as follows. In section 7.2.2, we describe how to extract patterns from complex particle systems that can be described by harmonic shape descriptors. In section 7.2.5 we describe, in detail, how the descriptors can be computed mathematically, and explore their unique properties, such as rotational invariance and sensitivity to rotational symmetries. In section 7.2.10, we describe how the harmonic descriptors can be used to solve representative problems from the field of computational and experimental

self-assembly and computational biology. In section 7.2.13, we explore how the special symmetry properties of harmonic descriptors can be applied to solve unique problems that are not easily solved by other types of shape descriptors. The methods that we describe here are applicable to all types of particle systems, both simulated and experimental, for which particle positions are known or can be accurately imaged. To aid with the dissemination of these techniques and new algorithms using harmonic shape descriptors as a basis, we provide a software library via the web [302].

7.2.2 Pattern Extraction

The first step in the structure characterization schemes that we will employ is to extract a representative structural pattern from the particle system that can be "indexed" (i.e., described mathematically) by a harmonic shape descriptor. While this is relatively trivial for small clusters of spherical point particles [132], for complex structures, some physical and mathematical intuition is often required. As we will demonstrate in detail in section 7.2.5, the harmonic shape descriptors that we introduce are best suited to index patterns on the unit circle, sphere, disk or ball. Such representations are often sufficient to distinguish between even very similar structures with a high degree of precision. The patterns may represent the particles themselves or some interesting pattern formed by their positions or density profile. In some cases, the raw data may be preprocessed to better extract important structural features, for example, by spatial coarse-graining, time averaging, or potential energy minimization.

Structural patterns in general can be described by a set of positions $\{\mathbf{x}\} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ and corresponding weights $\{f\} = \{f_1, f_2, \ldots, f_n\}$. For point cloud data, or raw particle coordinates, $\{\mathbf{x}\}$ represents the particle positions and the weights $\{f\}$ are equivalent and usually taken to be 1. For voxel data (i.e., volumetric data, often used to describe density maps), $\{\mathbf{x}\}$ represents the positions of bins on a grid with weights $\{f\}$. The same representation can be used to describe experimental images; in this case, $\{\mathbf{x}\}$ represents the positions of the pixels and $\{f\}$ represents their intensity. This notation allows us to write general equations for shape descriptors in section 7.2.5.

7.2.3 Local Structures

As originally shown in the context of bond order parameters [48, 132], a cluster of point particles is one simple type of structure that can be trivially represented by the projection of the points onto the surface of a circle or sphere. The patterns for two different point clusters

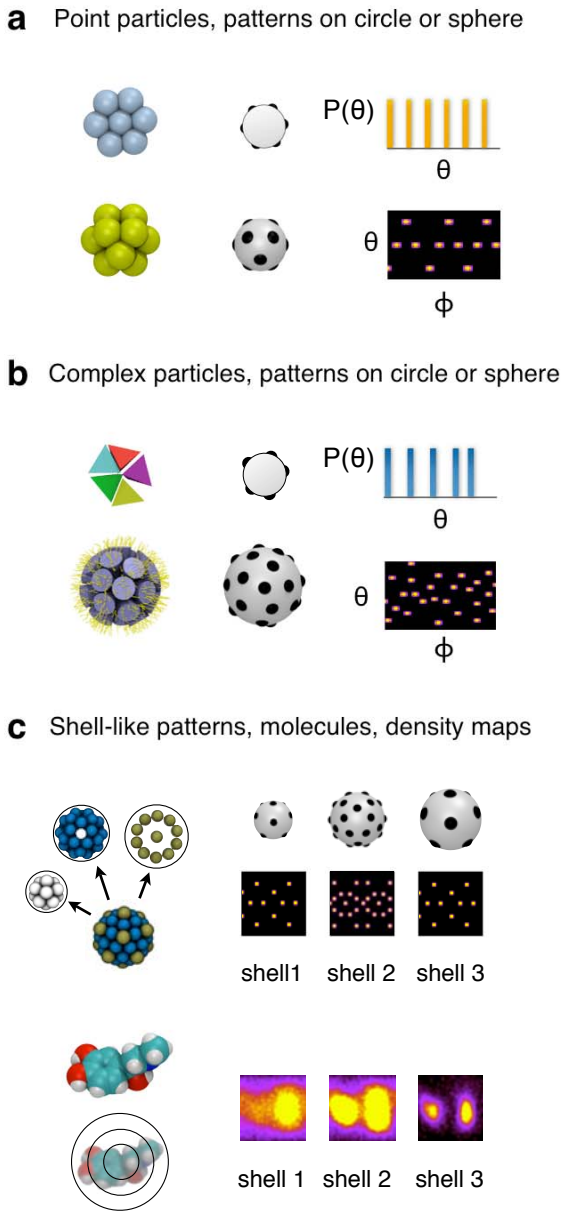


Figure 7.10 Extracting local patterns. (*a*) Clusters of point particles are expressed as patterns on the perimeter of the circle or on the surface of the sphere. (*b*) More complex particles are often idealized as point particles to the same effect. (*b*) Point clusters with *r*-dependence or spatial density maps can be decomposed into independent patterns for multiple radial shells.

are shown in Fig. 7.10a. Notice that we exclude the center particle from the pattern, since it has no specific direction. The clusters need not strictly consist of point particles; so long as particle shape is not important, the structural pattern can be described by placing points at the particle centroids, as depicted in Fig. 7.10b. Often, local structures are isolated from the bulk system by applying a clustering algorithm. One standard scheme is to cluster all particles within a cutoff range [132, 187]. More specialized schemes can be applied for specific applications, such as particles with complex shape [33].

Projecting patterns onto the circle or sphere neglects radial information. Therefore, such projections can lead to non-distinguishing patterns for structures with radial dependence. One solution is to decompose structures into concentric shells, projecting each shell onto the circle or sphere independently, as depicted in Fig. 7.10c. Structures can be compared by matching corresponding shells. Alternatively, complex patterns can be represented on the unit disk or unit ball. These representations account for radial dependence, as well as angular dependence. As described in section 7.2.7, representation on the disk or ball is optimal when decomposition into shells is inaccurate, or gives degenerate representations.

7.2.4 Global Structures

Structures with long-range ordering, such as crystals and phase-separated structures, cannot be directly indexed on the circle, sphere, disk or ball. Rather global patterns are constructed by combining different pieces of local information. One such method, originally introduced in the context of bond order parameters [132], is based on computing the superposition of all local patterns in a sample [312]. Fig. 7.11a shows the superposition of local patterns in an face centered cubic crystal, where local patterns are defined by the projection of neighbor directions on the unit sphere. Since crystals have long range orientational ordering, the neighbor directions are coincident throughout the sample. This type of structural pattern is independent of the shape of the underlying particles, and thus can be applied to many assembled structures, such as crystals of patchy particles [29], polyhedral particles [33], or phase-separated structures that form micelles or cylinders arranged in crystalline superlattices, such as dendrons [196], block copolymers [364], or tethered nanoparticles [307, 306, 331, 298, 187]. This superposition scheme is also applicable to many non-crystalline global phase-separated structures, such as layered and network structures, often formed by block copolymers or tethered nanoparticle systems [307, 306, 331]. Fig. 7.11b shows the superposition of local density maps represented on the unit ball for a phase-separated sheet structure formed by tethered nano-spheres [331].

For structures without long range orientational ordering, the superposition of local pat-

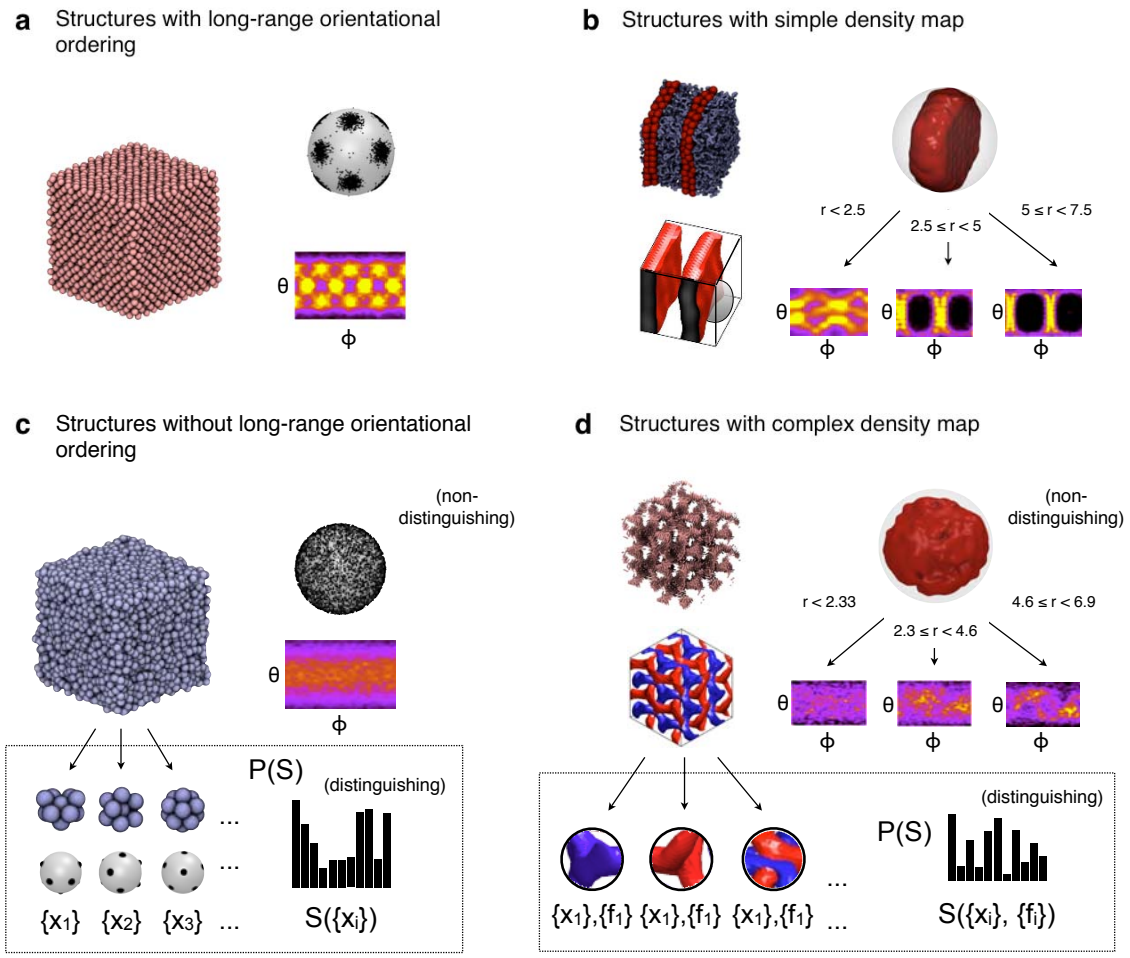


Figure 7.11 Extracting global patterns. (*a*) Superposition of local patterns for an fcc crystal. (*b*) Superposition of local patterns for a phase separated lamellar structure formed by tethered nanospheres [187]. (*c*) Superposition and probability distribution of local patterns for a liquid. Since the structure has no long range orientational ordering, superposition is non-distinguishing. (*d*) Superposition and probability distribution of local patterns for a double-gyroid structure formed by tethered nano rods [307, 309]. For this complex structure, the superposition of local patterns is non distinguishing, even though the structure has long-range ordering.

terns results in a random pattern over long ranges. Fig. 7.11c shows the superposition of local patterns for an atomic liquid, which results in a uniform distribution on the sphere. Since the same pattern is inherent to all liquids, gases, gels, etc., regardless of the underlying particle shape, superposition gives no structural information, other than indicating an absence of long-range orientational order. Orientationally-disordered structures of all types can be differentiated by considering the probability distributions function of local patterns, rather than the superposition. The probability distribution scheme is also useful when comparing structures with complex local patterns where superposition becomes degenerate or non-distinguishing, such as complex network structures [309]. This is depicted for the double-gyroid structure formed by tethered nano-rods [307, 309] in Fig. 7.11d.

7.2.5 Harmonic Descriptors

Given a pattern on the unit circle, sphere, disk or ball, the harmonic shape descriptors reviewed in this section can be used to index the pattern into a compact vector representation. This vector, or "shape descriptor," can be compared with other shape descriptors to obtain a quantitative measure of similarity between structures. In the derivation that follows, we introduce harmonic descriptors from a perspective that draws on elements of shape matching and signal processing. This contrasts with the physics-based perspective presented in the original derivation of bond order parameters [132]. The alternate perspective is meant to highlight the fundamentally different way in which the structural metrics are used; whereas bond order parameters are often applied directly as order parameters, harmonic descriptors represent structural fingerprints that must be matched to obtain structural information, which can subsequently be used for constructing order parameters.

Before we introduce harmonic descriptors, it is important to understand why they are more useful than simpler shape descriptor methods. Consider, for example, the problem of mathematically comparing two simple structures, such as the clusters shown in Fig. 7.10a. Perhaps the most obvious way to describe the different structural patterns is to simply use the coordinates themselves as the shape descriptor. While this greatly simplifies the initial step of creating a shape descriptor, it complicates matching significantly, since we do not typically know *a priori* the optimal correspondence between the coordinates in different lists. Reordering the lists is an optimization problem that can be solved, for example, by applying the Hungarian method [369]. This type of problem scales as $O(N^3)$ and thus quickly becomes inefficient for large N. A simple solution to the assignment problem is to create a probability histogram on the unit circle, sphere, disk or ball, as depicted in Fig. 7.10. Since the histogram bins are independent of the order of the particles in the list

(and number of particles), no assignment is required. Although this representation, known as the "shape histogram [321]" is very useful for some applications, one drawback is that to compare patterns in a way that is rotation-invariant, the patterns (or the histograms) must be aligned prior to matching. This "registration [315, 316]" step is computationally expensive and potentially inaccurate if applied naively. One elegant solution to these problems is to compute the discrete Fourier transform (DFT) for each shell in the shape histogram. The DFT transforms the pattern into its frequency-domain representation, which can be used to obtain rotation-independent harmonic descriptors through a simple mathematical operation as described in the following section. As an additional advantage, harmonic descriptors have adjustable frequency parameters that can be tuned to highlight important rotational symmetries or give a variable degree of coarse-graining. The manner in which we compute the harmonic descriptors depends on the coordinate system best used to describe the structure. In the following sections, we first introduce Fourier descriptors, which are suited for indexing shapes on the unit circle (θ dependence) or sphere (θ, ϕ dependence) and then introduce Zernike moments which are suited for indexing shapes on the unit disk (r, θ dependence) or ball (r, θ, ϕ dependence).

7.2.6 Fourier Descriptors

Fourier descriptors are designed to efficiently index structural patterns on the circle or sphere. The basic idea behind the Fourier transform is to represent a function as the sum of harmonic components. For a pattern along the 1d perimeter of the circle, we can do this by summing complex exponential terms:

$$f(\theta_j) = \sum_{\ell=0}^{\ell_{max}} \psi_\ell \exp\left[i\ell\theta_j\right] \quad j = 1, 2, ...n_{bin}. \tag{7.6}$$

Here, $f(\theta_j)$ is the intensity of the pattern at a particular point along the perimeter of the circle θ_j. The terms ψ_ℓ, known as "Fourier coefficients," indicate the strength of the pattern for a particular frequency ℓ. Typically, we cut off the frequency ℓ at some finite value ℓ_{max}, since the information for high-frequency ℓ becomes increasingly dominated by noise in the structure. The $\ell = 0$ and $\ell = 1$ terms only contain information regarding the position and center of mass of the pattern and are sometimes excluded. If the pattern consists of more than one radial shell, we compute the Fourier transform for each shell independently.

The Fourier coefficients contain structural information that can be used to create shape

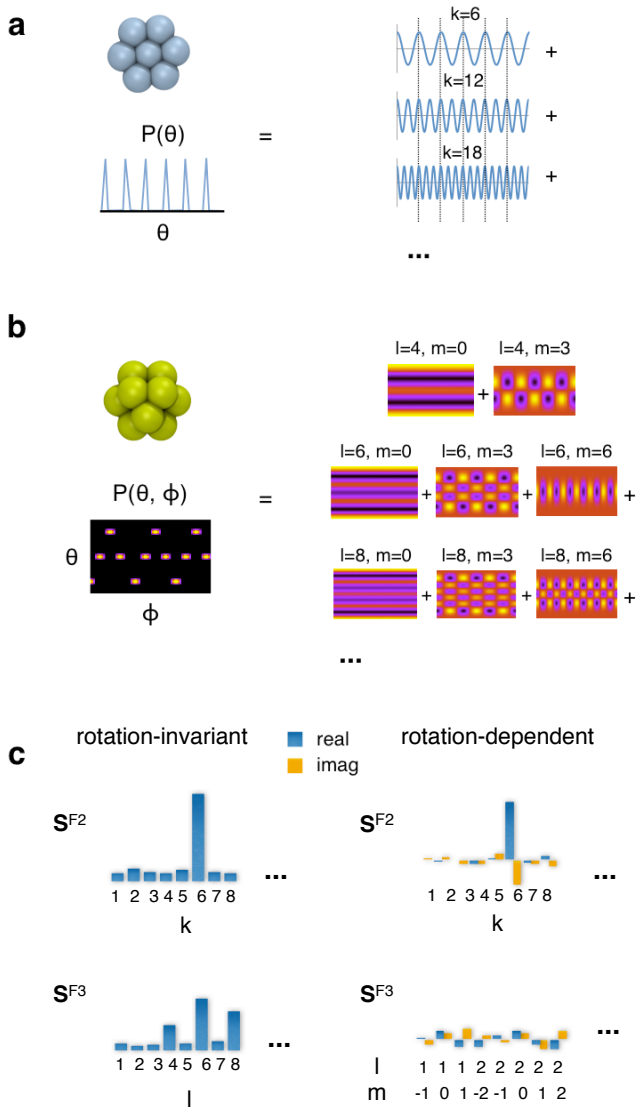


Figure 7.12 Fourier descriptors. (*a*) Decomposition of the angular pattern obtained for a 2d hexagonal cluster into a sum of sines and cosines. Since the cluster has 6-fold symmetry, the leading coefficients are given by multiples of $\ell = 6$. (*b*) Decomposition of the angular pattern obtained for a 3d fcc cluster into a sum of spherical harmonics. Since the cluster has 4 and 6-fold symmetry, the leading coefficients are given by multiples of $\ell = 4$ and $\ell = 6$. (*c*) Schematic of the resulting Fourier descriptors. The value of the each component in the descriptor is proportional to the contribution of the corresponding harmonic to the overall sum. Notice that rotation-invariant Fourier descriptors contain real, positive components, while rotation-dependent descriptors contain complex components.

descriptors with various properties. The coefficients are given by:

$$\psi_\ell = \frac{1}{n_{pts}} \sum_{j=1}^{n_{pts}} f_j \exp\left[i\ell\theta_j\right]^*. \tag{7.7}$$

The term θ_j is the angle of an input point $\mathbf{x}_j$ with intensity f_j from our pattern $\{\mathbf{x}\},\{f\}$. The coefficients ψ_ℓ are complex numbers. The * denotes the complex conjugate. As outlined in the previous section, the input points from our pattern can represent either the positions of bins in the circular shape histogram or the raw data if no binning is performed. The representations become equivalent as the bin size approaches zero and only one point can occupy a given bin.

Since the Fourier transform is a frequency-domain representation of the pattern, it gives the strongest signal for frequencies that reflect periodicities in the pattern around the circle. That is, patterns with n-fold rotational symmetry yield a high values of ψ_ℓ (and, as we will discuss later, $\mathbf{q}_\ell$). Since rotational symmetries are relatively insensitive to small changes to the pattern, Fourier descriptors are relatively insensitive to thermal noise, particularly for low-frequency coefficients. Although the Fourier coefficients in their complex number form are not rotation-invariant, we can convert them to their invariant form by computing the magnitude of each coefficient. The invariant circular coefficients are given by:

$$|\psi_\ell| = \psi_\ell \psi_\ell^* = \sqrt{\Re(\psi_\ell)^2 + \Im(\psi_\ell)^2}. \tag{7.8}$$

The Fourier invariants are positive real numbers.

To obtain some physical understanding of the properties of Fourier coefficients, consider the small cluster in Fig. 7.13a. If we normalize the cluster to the unit circle, the centroid (unweighted center of mass) is given in complex coordinates by conjugate of the $\ell = 1$ coefficient ψ_1^*. Using a frequency term other than $\ell = 1$ multiplies each angle θ by a factor, effectively stretching or compressing the pattern along the circle. We see that choosing $\ell = \psi_n$ where ψ_n is a rotational symmetry of the cluster causes the different angles θ_j to coincide, resulting in a non-vanising centroid for the transformed cluster. Thus, the Fourier coefficient with $\ell \neq 1$ represents the centroid of the stretched or compressed pattern. Although the position of the centroid is dependent on the cluster orientation, this distance from the origin to the centroid is invariant under rotations. Thus, we obtain a rotation-invariant descriptor by computing the magnitude of the coefficient.

These properties of Fourier coefficients have been exploited in the context of bond order parameters [48, 132]. For example, particular coefficient magnitudes, such as ψ_6 (or, analogously $\mathbf{q}_4$ and $\mathbf{q}_6$, see below) have been used directly as scalar order parame-

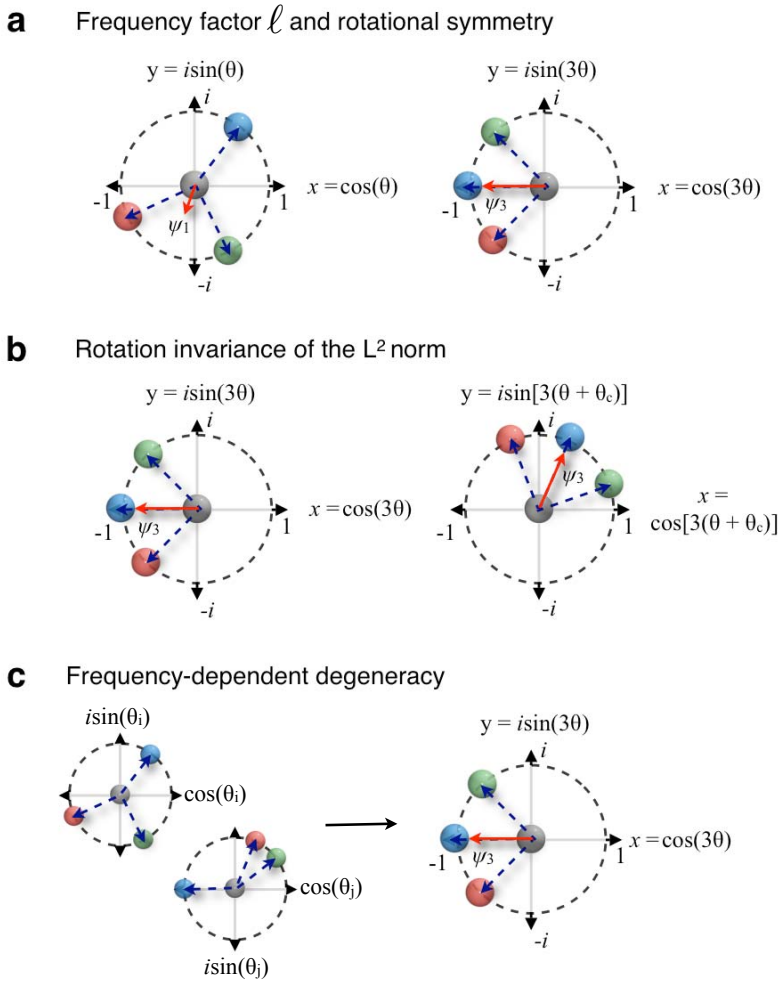


Figure 7.13 Properties of harmonic descriptors. (*a*) The depicted cluster is somewhat 3-fold symmetric. When each angle θ is multiplied by a frequency factor $\ell = 3$, the centroid position moves farther from the origin ($|\psi_3| > |\psi_1|$). (*b*) When the cluster is rotated (i.e., a constant factor θ_c is added to each angle θ), $|\psi_3|$ remains constant, although ψ_3 changes. (*c*) When the clusters on the left are scaled by a frequency factor $\ell = 3$, the transformed centroids ψ_3 become equivalent, even though the underlying shapes are different.

ters [132, 350]. This method is often sufficient for the simple structures that we encounter in systems of spherical particles. However, more complex structures often require a full range of coefficients, and order parameters must be constructed by comparing sets of coefficients, rather than evaluating particular coefficients. To create a descriptor from the Fourier coefficients, we simply combine the desired ψ_ℓ or $|\psi_\ell|$ into a long vector. For example, a general rotation-invariant shape descriptor that is applicable to patterns on the circle over a range of symmetries is given by:

$$\mathbf{S}^{\mathrm{F2}} =< |\psi_0|, |\psi_1|, ...|\psi_{\ell_{max}}| > . \tag{7.9}$$

It is easy to imagine how different combinations of the Fourier descriptors can be used to create shape descriptors with different levels of robustness, sensitivity to particular symmetries. For many applications it is common to include only coefficients with specific symmetries, or to use rotation-dependent coefficients.

As outlined in the previous section, many 3d structures are well represented by patterns on the surface of the sphere. The analogy to the 1d DFT on the 2d surface of the sphere is known as the discrete spherical harmonics transform (DSHT), given by:

$$f_j(\theta_j, \phi_j) = \sum_{\ell=1}^{\ell_{max}} \sum_{m=-\ell}^{\ell} q_\ell^m Y_\ell^m(\theta_j, \phi_j). \\ j = 1, 2, ...n_{bin} \tag{7.10}$$

The terms $Y_\ell^m[\theta_j, \phi_j]$ are spherical harmonics, defined by $Y_\ell^m(\theta, \phi) = N_\ell^m P_\ell^m(\cos\theta)\exp(im\phi)$. The term N_ℓ^m is a normalization factor $\sqrt{(2\ell+1)(\ell-m)!/(\ell+m)!}$ and P_ℓ^m is a Legendre polynomial. We see that the DSHT is similar to the DFT except for an additional term depending on the polar angle θ.

The Fourier coefficients for the DSHT are given by:

$$\mathbf{q}_\ell = \frac{1}{n_{pts}} \sum_{j=1}^{n_{pts}} f_j N_\ell^m Y_\ell^{m*}(\theta_j, \phi_j). \\ m = -\ell, -\ell+1, ...\ell \tag{7.11}$$

Unlike the circular coefficients ψ_ℓ, which are complex numbers, the spherical coefficients $\mathbf{q}_\ell$ are $2\ell+1$ dimensional complex vectors. Like the circular coefficients, the spherical Fourier coefficients are robust under noise, sensitive to rotational symmetries corresponding to ℓ, and can be used to construct invariants. Although the geometrical interpretation of these properties is more complex than for the 1d case, the same principles apply.

The rotation-invariant version of the spherical coefficients is given by:

$$|\mathbf{q}_\ell| = \sqrt{\frac{4\pi}{2\ell+1}\sum_{m=-\ell}^{\ell}|q_\ell^m|^2}. \tag{7.12}$$

Like the circular invariants $|\psi_\ell|$, the spherical invariants $|\mathbf{q}_\ell|$ are positive real numbers.

In analogy with our example above, we can create a rotation-invariant shape descriptor for a pattern on the sphere by:

$$\mathbf{S}^{\mathrm{F3}} =< |\mathbf{q}_0|, |\mathbf{q}_1|, ...|\mathbf{q}_{\ell_{max}}| > . \tag{7.13}$$

Again, the optimal descriptor for a particular application depends on the desired properties such as robustness, sensitivity to specific symmetries, and rotation invariance.

Notice that although we use the notation "ψ_ℓ" and "$\mathbf{q}_\ell$" to highlight the connection with bond order parameters, we have redefined the order parameters slightly, by changing the sign of the complex exponential (i.e., the conjugates in equations 7.7 and 7.11). This sign change is inconsequential to the properties of the coefficients; our redefinition simply allows us to highlight the important relationship between ψ_ℓ and $\mathbf{q}_\ell$ and the DFT, which is standard and extensively studied. An overview of the Fourier descriptors method is given in Fig. 7.13. As a general rule of thumb, "ψ_ℓ" is used when we only wish to describe the 2d ordering of a system (e.g., the spatial ordering within a single plane in a confined fluid [350, 351] or the crystalline arrangement of cylindrical domain [298, 314]) and "$\mathbf{q}_\ell$" is used when we wish to describe the 3d ordering (e.g., the spatial ordering in a 3d crystal [57] or structure of a compact 3d cluster [187]).

7.2.7 Zernike Descriptors

As mentioned in the previous section, when patterns cannot be properly represented by the surface of a single circle or sphere, one solution is to break the pattern up into independent radial shells, and compute the Fourier descriptor for each shell independently. We then construct a shape descriptor by combining the Fourier descriptors for each shell into a long vector:

$$\mathbf{S}^{\text{F, multishell}} =< \mathbf{S}^F_{shell_1}, \mathbf{S}^F_{shell_2}, ...\mathbf{S}^F_{shell_n} > . \tag{7.14}$$

Here, $\mathbf{S}^{\mathrm{F}}$ represents a Fourier descriptor, either $\mathbf{S}^{\mathrm{F2}}$ or $\mathbf{S}^{\mathrm{F3}}$, as defined in the previous section.

While this scheme is sufficient for many problems, it has the drawback that small perturbations to the particle positions can cause maxima in the pattern to shift between shells,

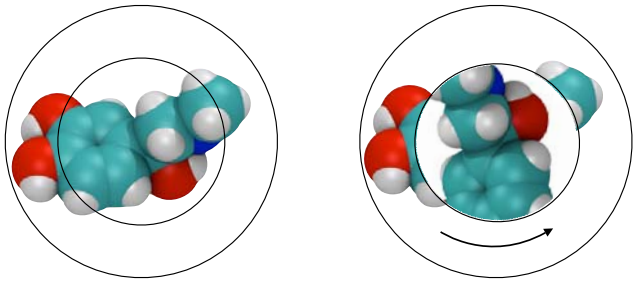


Figure 7.14 Drawbacks of Fourier descriptors. Rotation-invariant Fourier descriptors for structures with multiple shells are insensitive to the relative orientation of inner and outer shells. Using invariant Fourier descriptors to describe the two structures shown in the schematic shown erroneously produces identical results.

causing errors in matching, particularly when n_{pts} is small. Another drawback is that, since the shells are treated independently, the rotation-invariant Fourier descriptors are insensitive to relative orientations between the different shells; this is depicted in Fig. 7.14 for two structures that would erroneously have matching descriptors.

We can solve these problems by representing our pattern in a coordinate system with radial dependence, which allows us to properly index patterns defined on the disk or ball rather than the unit circle or sphere. To do so, we use Zernike radial polynomials in our expansion [325, 326].

We can express the intensity of a pattern at a given point on the unit disk as:

$$f(r_j,\theta_j) = \sum_n \sum_\ell a_{n\ell} R_{n\ell}(r_j) \exp\left[i\ell\theta_j\right]. \qquad j = 1,2,\ldots,n_{bin} \tag{7.15}$$

The terms θ_j and r_j represent the position of a point on the unit disk. The value of ℓ is restricted such that $\ell \leq n$ and $(n-\ell)$ is an even number. The expansion coefficients $a_{n\ell}$ are known as "Zernike moments" and can be considered analogous to Fourier coefficients for the r,θ coordinate system. The function $R_{n\ell}(r)$ is a radial polynomial, where r is the radial distance from the center of the disk. Thus, the 2d Zernike expansion is very similar to the 1d Fourier expansion, but with an additional radial term.

The Zernike moments are given by [325]:

$$a_{n\ell} = \frac{n+1}{n_{pts}\pi} \sum_{j=1}^{n_{pts}} f_j(r_j,\theta_j) R_{n\ell}(r_j) \exp\left[-i\ell\theta_j\right]. \tag{7.16}$$

The terms θ_j and r_j represent the position of an input point $\mathbf{x}_j$ in polar coordinates, normal-

ized on the unit disk. Again, we require that $\ell \leq n$ and $(n-\ell)$ is even. Each moment is a complex number. Since the radial polynomial is only dependent on r, the same invariance relations hold for the Zernike moments as for the Fourier coefficients. To define a rotational invariant moment on the disk, we take the complex magnitude of the moment:

$$|a_{n\ell}| = a_{n\ell} a_{n\ell}^*. \tag{7.17}$$

The 2d Zernike invariants are positive real numbers. We can create a Zernike descriptor by combining many moments in a vector. For example, we can create a 2d rotation-invariant Zernike descriptor by:

$$\mathbf{S}^{Z2} = < |a_{00}|, |a_{11}|, |a_{20}|, |a_{22}|, ... |a_{\ell_{max}\ell_{max}}| > . \tag{7.18}$$

Like the Fourier descriptors, the frequency parameter ℓ has a straightforward relationship with the rotational symmetry of the pattern. Thus, we can sometimes choose important moments *a priori*. However, we typically take all moments with ℓ within a limiting frequency ℓ_{max}.

We can express a pattern on the unit ball as the sum of 3d Zernike moments:

$$f(r_j, \theta_j, \phi_j) = \sum_n \sum_\ell \sum_m z_{n\ell}^m R_{n\ell}(r_j) Y_\ell^m(\theta_j, \phi_j). \\ j = 1, 2, ... n_{bin} \tag{7.19}$$

The terms $[r_j, \theta_j, \phi_j]$ give the position of a point on the unit sphere. Again, we require $\ell \leq n$ and $(n-\ell)$ is even. The moments are defined similarly to the Fourier coefficients on the surface of the sphere, but again with an additional radial component. The 3d Zernike moments are given by [326]:

$$\mathbf{z}_{n\ell} = \frac{3(n+1)}{4 n_{pts} \pi} \sum_{j=1}^{n_{pts}} f_j R_{n\ell}(r_j) N_\ell^m Y_\ell^{m*}(\theta_j, \phi_j). \\ m = -\ell, -\ell+1, ... \ell \tag{7.20}$$

The variables $[r_j, \theta_j, \phi_j]$ represent the position of an input point $\mathbf{x}_j$ in spherical coordinates, normalized on the unit sphere. Whereas the 2d Zernike moments $a_{n\ell}$ are complex numbers and the 3d Zernike moments $\mathbf{z}_{n\ell}$ are complex vectors of length $2\ell+1$. Analogously to the spherical Fourier coefficients, we take the magnitude of the complex vector $|\mathbf{z}_{n\ell}|$ to define

invariant moments on the unit ball:

$$|\mathbf{z}_{n\ell}| = \sqrt{\frac{4\pi}{2\ell+1}\sum_{m=-\ell}^{\ell}|z_{n\ell}^{m}|^2}. \tag{7.21}$$

The 3d Zernike invariants are positive real numbers. Like the Fourier coefficients, they are sensitive to the rotational symmetries of the pattern and are robust under small perturbations. We can create a rotation invariant symmetry-independent 3d Zernike descriptor according to:

$$\mathbf{S}^{Z3} =< |\mathbf{z}_{00}|, |\mathbf{z}_{11}|, |\mathbf{z}_{20}|, |\mathbf{z}_{22}|, ...|\mathbf{z}_{\ell_{max}\ell_{max}}| > . \tag{7.22}$$

When computing either multiple-shell Fourier descriptors or Zernike moments it is essential that the patterns being compared are normalized consistently. In the case of Zernike moments, all points in $\{X\}$ must lie on the unit ball or disk. Typically, patterns are normalized by translating the centroid to the origin and rescaling the coordinates such that every point on the pattern has a radial distance less than 1. This scheme is sufficient for the majority of patterns that we encounter in particle systems. An overview of the Zernike scheme is depicted in Fig. 7.15.

7.2.8 Computational Considerations

Fourier and Zernike coefficients can be computed from either point cloud data (i.e., raw particle positions) or voxel data (i.e., volumetric data or pixel data). As mentioned previously, the two representations are essentially equivalent; point cloud data represent the limit of zero bin size, where each f_i is equivalently 1. While this distinction does not affect the properties or definition of the descriptors, it becomes important when considering the computational cost of a matching application. If the input data is point cloud data, we must compute the descriptors for each structure independently. However, for volumetric data, we can compute the contribution to each coefficient for each point on the grid beforehand, and then simply multiply by the intensity of each shape f_i to compute the value of the coefficients. This can greatly reduce the computational cost when n_{pts} is large or many coefficients are used.

As an additional consideration, the time required for computing the transforms themselves can be greatly reduced by computing the fast Fourier transform (FFT) rather than the DFT (or the equivalent for the appropriate coordinate system). Methods for computing the FFT and the discrete spherical harmonics transform, respectively are given in references [164] and [370]. An efficient method for computing Zernike coefficients is given in reference [326].

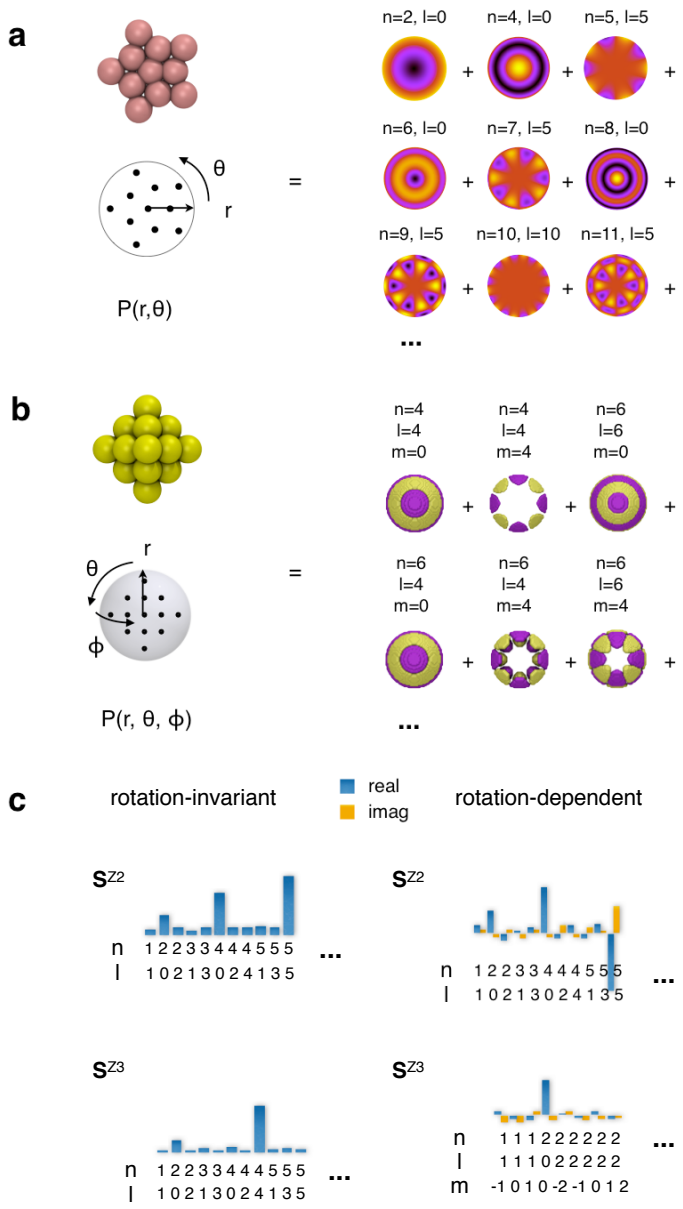


Figure 7.15 Zernike descriptors. (*a*) Decomposition of the pattern obtained for a 2d pentagonal cluster with Zernike moments. Since the cluster has 5-fold symmetry, the leading coefficients are given by multiples of $\ell = 6$. (*b*) Decomposition of the pattern obtained for a 3d fcc cluster with Zernike moments. Since the cluster has 4 and 6-fold symmetry, the leading coefficients are given by multiples of $\ell = 4$ and $\ell = 6$. (*c*) Schematic of the resulting Zernike descriptors. The value of the each component in the descriptor is proportional to the contribution of the corresponding moment to the overall sum. Notice that rotation-invariant Zernike descriptors contain real, positive components, while rotation-dependent descriptors contain complex components.

7.2.9 Quantifying Similarity

The shape descriptors derived in the previous section can be considered compact mathematical representations of the underlying particle structures. The physical similarity between different particle structures can then be quantified by the mathematical similarity between shape descriptors. Shape descriptor similarity is quantified by a similarity metric "M," that gives a scalar value that is proportional to the similarity between descriptor pairs. For convenience, we define M such that, by construction, it lies on the interval $M \in [0,1]$, or sometimes $M \in [-1,1]$. This strengthens the analogy between M and what we normally consider to be an order parameter, since order parameters typically give a value of 1 for perfectly ordered structures and 0 for perfectly disordered structures. Since our harmonic shape descriptors are defined as vectors, we can define similarity metrics based on standard vector operations. For example, one simple similarity metric is given by the Euclidean distance between shape descriptor vectors

$$M_{dist}(\mathbf{S}_1,\mathbf{S}_2) = 1 - |\mathbf{S}_1 - \mathbf{S}_2|/(|\mathbf{S}_1| + |\mathbf{S}_2|). \tag{7.23}$$

Variations of M_{dist} are common throughout the shape matching literature. Another simple similarity metric is proportional to the dot product between shape descriptors:

$$M_{dot}(\mathbf{S}_1,\mathbf{S}_2) = \mathbf{S}_1 \cdot \mathbf{S}_2/(|\mathbf{S}_1||\mathbf{S}_2|). \tag{7.24}$$

Schemes similar to M_{dot} have been used in applications involving standard bond order parameters, such as measuring correlation lengths [132, 365] and identifying crystal grains [57].

Similar information is obtained from M_{dist} and M_{dot}; the only difference is that whereas M_{dist} is more sensitive to the absolute difference between vector components, M_{dot} is more sensitive to the overall direction of the vector and the sign of the components. Thus, M_{dot} is often superior when matching non-ideal structures from a particle system to mathematically perfect reference structures, since thermal noise will tend to damp the frequency domain signal, but the descriptor will retain the same basic character for a given class of structures and hence the same direction. The M_{dot} metric may also be favorable when comparing rotation-dependent harmonic descriptors, which may contain either negative or positive components, whereas M_{dist} may be favorable for invariant descriptors, where all values are inherently positive. Since compared shapes are usually at least grossly similar, matching values are rarely 0 for either metric. As is discussed in greater detail in reference [37], it is often necessary to determine a lower bound on M by comparing to structures that are known to match poorly to obtain a baseline value.

7.2.10 Example Applications

Shape similarity information obtained from evaluating the match, M, between shape descriptors, can be applied to create various types of structural metrics for complex particle systems. In this section, we provide several example applications that can be addressed by the use of shape matching with harmonic descriptors. We provide a more extensive range of example applications based on shape matching methods in general in reference [37]. Several additional examples of shape matching applications based on the $\mathbf{S}^{\mathrm{F3}}$ descriptor are given in reference [34].

7.2.11 Order Parameters and Correlation Functions

Perhaps the most standard application of order parameters is to track structural transitions, either as a function of time or a changing reaction coordinate. As an example, consider the protein "Ubiquitin [371]" shown in Fig. 7.16, which unfolds as it is pulled from both ends. This is a standard example problem from the NAMD and VMD tutorials [372, 373], both of which are available online. Since the structure is 3-dimensional and has radial dependence, we can index it using a Zernike descriptor on the unit ball, $\mathbf{S}^{\mathrm{Z3}}$. To match the shape independent of the orientation of the sheet, we take rotation invariant moments with ℓ in the range $4 \leq \ell \leq 12$. We use both the initial folded state i and the final unfolded state f as reference states. Fig. 7.16 shows the unfolding transition as a function of time t in a NAMD molecular dynamics simulation. The protein unfolds in three steps (blue dashed lines), in agreement with visual inspection. The noise in the data is indicative of the thermal fluctuations in shape observed at this temperature.

As a slightly more complex example of characterizing transitions, consider the cylindrical phase separated structures formed by ditethered nanospheres [314] shown in Fig. 7.17. The structure of the equilibrium system goes through two transitions as a function of the effective inverse temperature, first from a disordered structure to a tetragonal cylinder/tetragonal-mesh (TC/TM) phase and then to a similar tetragonal cylinder (TC/TC) phase. The abbreviations indicate the structure of the tethers (blue, red) and nanoparticles (white), respectively. We can quantify this behavior by matching the global patterns obtained at different temperatures with ideal structures taken from within the three structural regimes: the disordered regime, the TC/TM regime, and the TC/TC regime. The global pattern for each structure is characterized by the probability distribution of local density maps for each particle type, as depicted in Fig. 7.11d. To capture ordering on a range of lengthscales, density maps are computed for 4 radial shells ranging from $r = [3\sigma, 4.5\sigma, \ldots, 9\sigma]$, where σ is the charac-

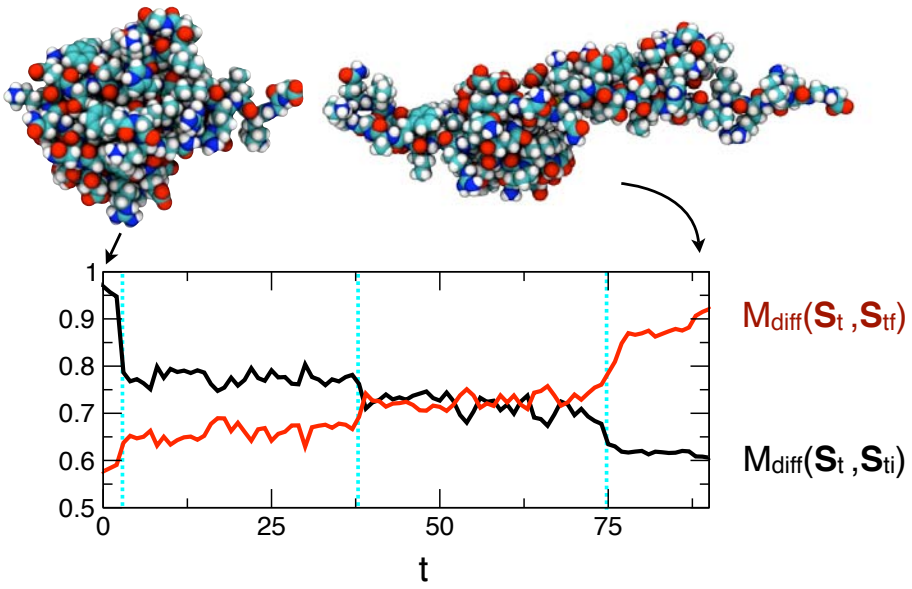


Figure 7.16 Tracking structural transitions in the protein Ubiquitin as it is pulled from both ends [372, 373]. The protein is pulled linearly as a function of time. The black curve shows a matching order parameter using the initial folded configuration i as a reference structure. The red curve shows the order parameter using the final unfolded configuration f as the reference structure. The blue lines highlight the times at which significant structural changes occur.

teristic lengthscale, given by a the diameter of the tether beads in the simulation. For each shell, we compute the rotation-invariant Fourier descriptors $\mathbf{S}^{\mathrm{F3}}$, where we take a range of frequencies $4 \leq \ell \leq 12$. A pseudo-order parameter for each reference structure is then given by $M_{dist}(\mathbf{S}^{\mathrm{F3}}, \mathbf{S}^{\mathrm{F3}}_{ref})$. Fig. 7.16 shows the order parameters for the three reference structures as a function of inverse temperature. We observe that the structural transition between the three phases is smooth and continuous, as verified by visual inspection. In reference [37], we show that, for this particular problem, we can obtain a nearly identical result using a simpler shape descriptor akin to the radial distribution function $g(r)$. However, for more complex phase separated structures, harmonic descriptors typically give a better representation of the underlying shapes than such coarse measurements.

In addition to characterizing how structures change as a function of time or a reaction coordinate, another common application of structural metrics is to characterize how structures change in space, by computing correlation functions. In this case, we choose structures from different points in space, rather than ideal structures, as references. Several examples of spatial correlation functions based on bond order parameters have already been defined in the context of measuring lengthscales for crystal-like ordering [48, 54, 132, 366]. This involves measuring quantities such as $\langle M(\mathbf{S}_i, \mathbf{S}_j) \rangle (r_{i,j})$, that give the average similarity

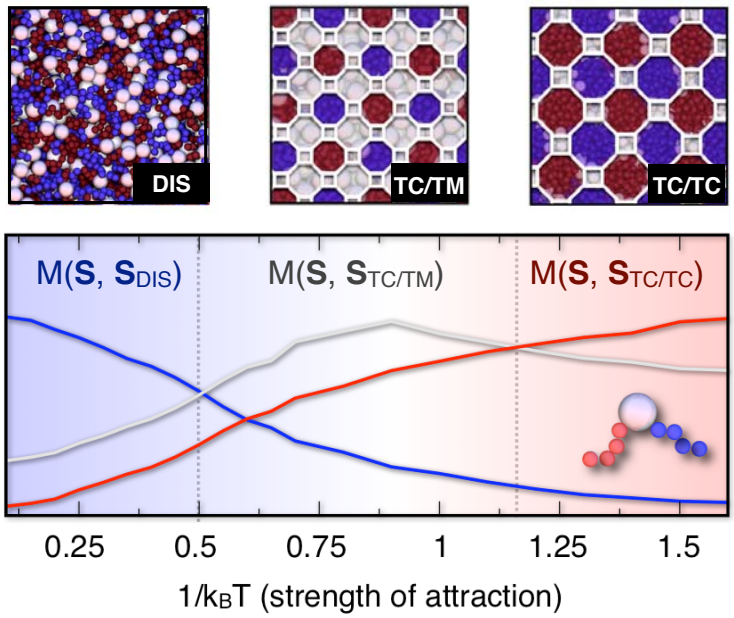


Figure 7.17 Transitions in ditethered nanospheres. (*a*) Depiction of the three structures formed by a ditethered nanosphere system [314] (left to right: disordered, TC/TM, TC/TC). Notice that the cylindrical structures span the simulation box in the z-dimension, into the page. (*b*) Matching order parameter for the three reference structures as a function of inverse temperature.

value as a function of the radial separation $r_{i,j}$. Typically, a particular rotation-dependent Fourier coefficient, for example ψ_6 or $\mathbf{q}_6$, is chosen for $\mathbf{S}$ and M_{dot} is chosen for M. An alternative, and well known, spatial correlation function based on bond order parameters is given by the $q_6 \cdot q_6$ crystal grain detection scheme of reference [57], which has been applied to studying nucleation and growth and characterizing crystalline defects in several simulation and experimental studies involving close-packed and bcc crystals [62, 367, 292, 293]. Here, crystal grains are identified within a bulk liquid by first noticing that, for many crystals, local clusters within crystal grains match with their neighbors in terms of *both* shape and orientation, whereas clusters in the liquid do not. Thus, pairs of particles i, j in grains typically satisfy $M(\mathbf{S}_i, \mathbf{S}_j) > M_{cut}$, where $\mathbf{S}$ is a rotation-*dependent* harmonic descriptor and M_{cut} is a sufficiently high value so as to exclude poor matches. Even in the liquid, some pairs of particles inevitably satisfy $M(\mathbf{S}_i, \mathbf{S}_j) > M_{cut}$ due to random fluctuations. Thus, a local indicator of crystal-like ordering is given by:

$$f_i = \frac{1}{n_{nbr}} \sum_{j}^{n_{nbr}} \Theta[M(\mathbf{S}_i, \mathbf{S}_j) - M_{cut}]. \tag{7.25}$$

Here, Θ is the Heaviside function. Typically, we enforce $f_i >= f_{cut}$, where the value of

f_{cut} is chosen to distinguish liquid and crystal-like configurations [57, 62]. Although the original scheme is based on the $\mathbf{q}_6$ Fourier coefficient as a shape descriptor (for identifying fcc,hcp, and bcc crystals) and the M_{dot} similarity metric, other instances of $\mathbf{S}$ and M can be used depending on the structure under investigation. For example, in references [29] and [30], we used different bond order parameters as shape descriptors to identify particles in the diamond lattice and in dodecagonal quasicrystals, respectively. Arbitrarily complex crystal structures can be treated using this method by harmonic descriptors with a full spectrum of Fourier coefficients or Zernike moments. We apply this scheme in the context of two examples highlighting the special symmetry properties of harmonic descriptors in section 7.2.13 below.

7.2.12 Database Search and Structure Identification

In addition to computing order parameters and correlation functions, another common application of bond order parameters is to identify local structures, such as icosahedral clusters [346, 30, 187, 309] within a bulk system. This typically involves choosing a cutoff value for a particular Fourier coefficient (for example, $|q_6|$) above which a cluster is identified as a the structure of interest. This is, in its essence, a rudimentary shape matching scheme, where the Fourier coefficient provides a coarse description of the cluster shape, and the cutoff acts as a similarity metric. This type of structure identification scheme can be applied within a much broader context by using harmonic descriptors to perform a database search for an unknown structure. The unknown structure is identified as the structure from the database of known structures that gives the best match. Database searches based on harmonic descriptors have already been applied to proteins and macromolecules [344, 310]. In an earlier publication, we applied a database search to identify local structures within a phase separated system of tethered nanoparticles [187]. In the future, they may be applied to more abstract problems, such as data mining for web-accessible particle structures.

Database searches based on harmonic descriptors can be applied to a wide range of complex local and global structures. As an example, consider the tunneling electron microscopy (TEM) image depicted in Fig. 7.18, which shows nanoparticles arranged in a binary crystalline superlattice from reference [297]. Although this structure was identified as the AuCu crystal viewed along the 100 directions by visual inspection [297], assume for the for the purpose of this example that the structure is unknown. The structure of the lattice can be identified by finding a best match from a database containing the images of known reference structures. For our proof-of-concept example, we use a minimal reference database consisting of four different ideal binary crystal structures; however for more re-

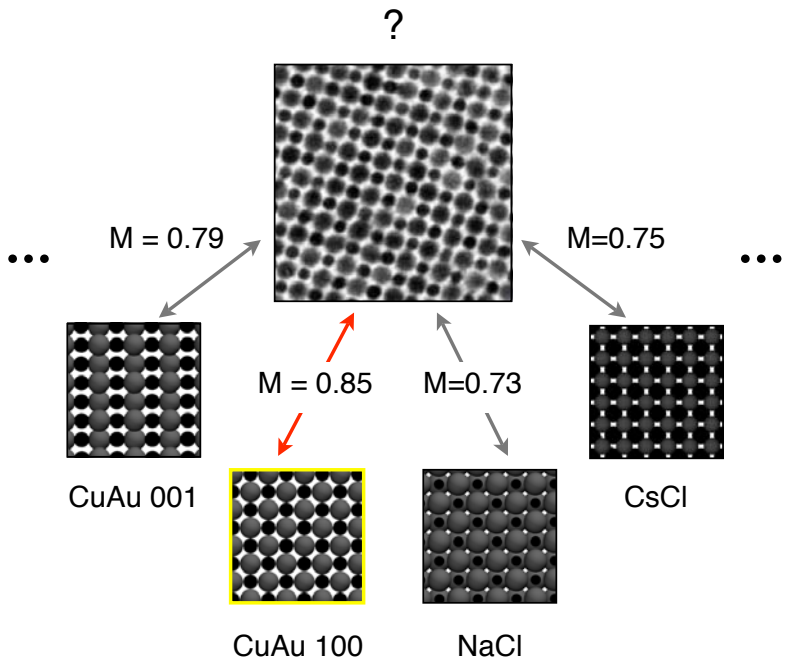


Figure 7.18 Structure identification using a tunneling electron microscopy (TEM) image from reference [297]. The structure depicted in the image is identified by finding the best match from a database of candidate structures. For this particular example, the reference structures are created by mathematical construction. The image is correctly identified as depicting the CuAu structure viewed along the 100 direction.

alistic problems, the database may be much more expansive. The reference structures are created by mathematical construction and rendered by placing spheres at the lattice positions. In practice, matching can be performed using other non-ideal images or other experimental images. The images are indexed for comparison using the 2d Zernike descriptor $\mathbf{S}^{Z2}$. As mentioned previously, harmonic descriptors can be used to describe images [325], where $\{\mathbf{x}\}$ and $\{f\}$ represent the positions and intensities of individual pixels. In practice, the pixel intensities are inverted ($I_i = -I_{i,0}$), since, for the current set of images, the particles are darker than the background. To ensure that the matching algorithm is not affected by the different particle shades, we apply a binary thresholding criterion $I_i = \Theta(I_i - I_{cut})$. To extract a global pattern from the image, we use the probability distributions method as depicted in Fig. 7.10. (Notice that although the structures are crystalline, the superposition method is not applicable, since the particle centers are not known). For each local structure, we compute $\mathbf{S}^{Z2}$ descriptors with rotation-invariant moments and frequencies in the range $0 \leq \ell \leq 8$. Our overall results are not impacted by the inclusion of higher frequencies. For each image, local descriptors are computed for 100 different randomly chosen pixels. For

each pixel, the range of neighboring pixels used to construct the local descriptor corresponds to roughly three particle diameters. The local descriptors are then combined into an overall probability histogram descriptor for matching. As shown in Fig. 7.18, the unknown structure most closely resembles the CuAu lattice along the 100 direction, in agreement with visual inspection [297]. Additional orientations of the crystalline lattices could also be considered to better identify structures with complex ordering.

This same basic identification scheme can be used for all types of structures, either simulated or experimental, local or global. A subtlety arises when disordered local structures are possible, since it is typically infeasible to construct a reference database for the vast space of "disordered" structures. As outlined in references [187, 37] this problem can be avoided by setting a minimum value for the best match, below which structures are considered disordered.

In addition to database searches, experimental images can be used for all of the other applications described here or in references [34] and [37], including quantifying structural perfection, computing order parameters and correlation functions, grouping and classifying structures, etc. The only caveat is that, like simulation data, the images must be properly normalized such that the particle sizes and lattice spacings (if applicable) are the same for all compared structures. In addition to images of particles, information can be extracted from other types of images, such as diffraction patterns. For many applications, image processing can be applied to experimental images to simplify the data [291, 294, 300], for example, by identifying particle centroids. We explore the combination of image processing techniques and shape matching algorithms in a separate publication [374].

7.2.13 Special Properties of Harmonic Descriptors

In the previous section, we applied harmonic descriptors within the context of general particle shape matching applications, similar to those outlined in reference [37]. Thus, although the harmonic descriptors have useful properties, such as rotation invariance, we could just as easily base our examples on other shape descriptors with similar properties. In contrast, in this section we explore applications for which harmonic descriptors and Fourier coefficients are specifically well-suited, due to their unique symmetry properties.

7.2.14 Matching to Within an n-Fold Rotation

One such application is the problem of matching structures that are a unique rotation of one another. As an example, consider the problem of detecting the local crystal grains of

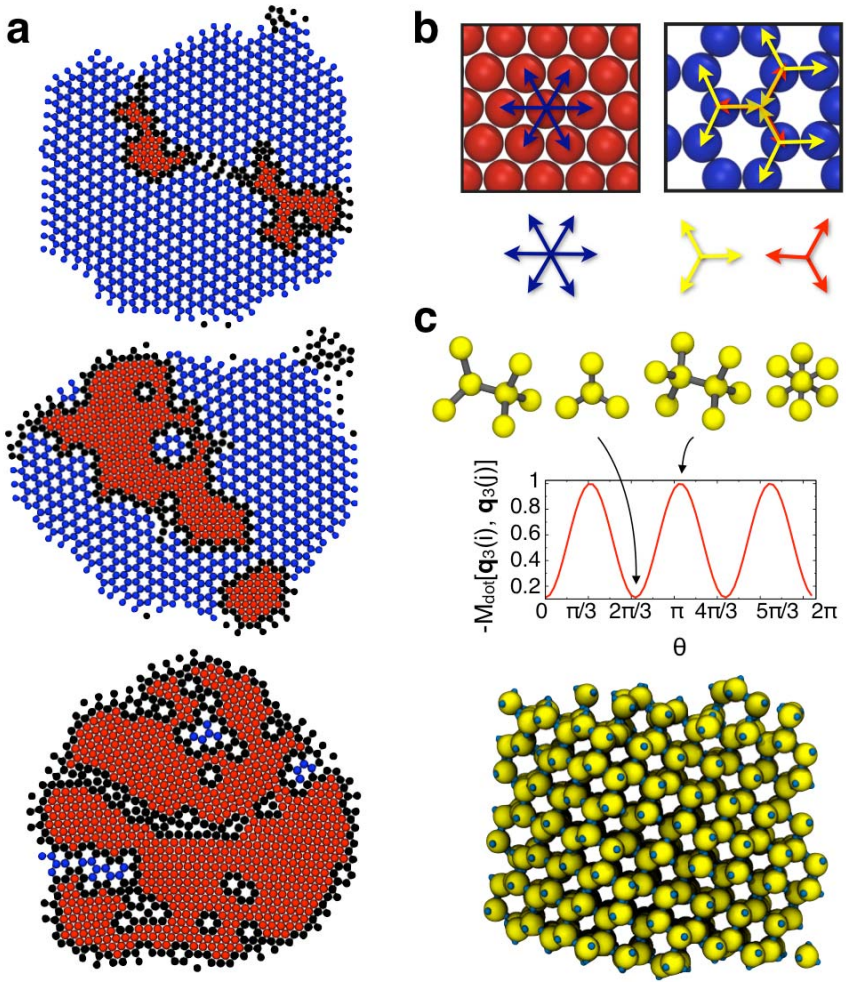


Figure 7.19 Honeycomb (hc) to hexagonal (hex) Transition in the 2d Lennard-Jones Gauss (LJG) system [120]. (*a*) Annealing results obtained for a LJG well-minimum position $r_0 = 1.57$ for three different values of the well depth parameter ε. From top to bottom: $\varepsilon = 2.7, 2.2, 1.4$. Particles in the hc structure are colored blue, while particles in the hex structure are colored red. All other particles are colored black. (*b*) Depiction of the local neighbor configurations in the hex and hc structures. (*c*) Diamond lattice formed by patchy particles [29, 280]. Analogously to the hc lattice, the diamond lattice can be characterized by the negated match of the third-order harmonic [29].

different structures in the 2d system depicted in Fig. 7.19a, where particles interact via the Lennard-Jones Gauss (LJG) potential [120]. For a well-position parameter $r_0 = 1.57$, the system forms two crystal structures: the honeycomb (hc) structure at high values of the well-depth parameter ε and the hexagonal (hex) structure at low ε [120]. As outlined in section 7.2.11, local crystal-like pairs are typically identified by local clusters that match in terms of both shape and orientation. This method is easily applicable to the hex lattice; however, in the case of the hc lattice, the triangular first neighbor shells of neighboring particles are mirrored and rotated by 60° in the plane (see Fig. 7.19b). Thus, they match in terms of shape, but not in terms of orientation. The symmetry properties of the Fourier coefficients pose a unique solution to this problem. In the space of the $\ell = 3$ Fourier coefficient ψ_3, the triangular neigbor shells in the hc lattice are precisely antiparallel (i.e., $M_{dot}[\psi_3(i),\ \psi(j)_3] = -1$, where i and j are neighbors). Thus, a matching criterion can be constructed based on $-M_{dot}(\psi_3,\ \psi_3)$ to determine whether two neighbor shells are in the ideal hc configuration. In Fig. 7.19a, particles in the hc structure are colored blue, particles in the hex structure are colored red, and other particles, that don't belong to a particular crystal grain, are colored black. The 3d analogy of this method, using $\mathbf{q}_3$ in the place of ψ_3, was used to measure the number of particles in local diamond lattice grains in a system of patchy particles in reference [29] (see Fig. 7.11c).

7.2.15 Matching Based on Rotational Symmetries

A similar application for which Fourier coefficients are uniquely suited is the problem of matching structures based on their rotational symmetries rather than their shapes. As an example, consider the problem of matching local neighbor shells in the decagonal (i.e., 10-fold symmetric) quasicrystal formed in the 2d LJG system [120] (Fig. 7.20a). Over the range indicated, the neighbor shells exhibit strong 10-fold rotational symmetry with a common direction, but the neighbor shells have different shapes. Thus, our criterion for detecting local crystal grains outlined in section 7.2.11 fails for most shape descriptors, since the underlying shapes don't match. As a solution, we can describe each local cluster with the $\ell = 10$ Fourier coefficient ψ_{10}. Since the clusters are 10-fold symmetric, and oriented in the same direction, the complex number ψ_{10} is identical regardless of whether the clusters are missing particles. Local quasicrystalline grains can then be detected as outlined in section 7.2.11, using $M_{dot}(\psi_{10},\ \psi_{10})$ to identify local crystal-like pairs.

In reference [30] we use an analogous scheme to detect ordered grains in a 3d dodecagonal (12-fold symmetric) quasicrystal (Fig. 7.20b). In this case, the structure has hundreds of different neighbor shell directions [312], which exhibit strong 12-fold symmetry. This is

depicted by the superposition of local patterns over the range $r < 2.31\sigma$ (i.e., ~ 2 neighbor shells) in Fig. 7.20c. Considering this longer range ensures that each local region contains a sizable fraction of 12-fold directions. Since we are only interested in the rotational symmetry and directionality of these local patterns, we remove the r-dependence from the patterns prior to matching (Fig. 7.20c). To capture the 12-fold symmetry, we use a matching criterion based on $M_{dot}(\mathbf{q}_{12},\ \mathbf{q}_{12})$. Particles with a minimal fraction of solid-like matches f_{cut} are considered to be locally quasicrystalline. The cutoffs are determined by taking the crossover points in the probability distributions $P(M_{dot}(\mathbf{q}_{12},\mathbf{q}_{12}))$ and $P(f_{solid})$ (Fig. 7.20d). Following reference [30], we take $M_{cut} = 0.45$, $f_{cut} = 0.5$. As depicted in Fig. 7.20e, this criterion is sufficient to determine a small quasicrystal nucleus in the bulk liquid.

Notice that although both of our examples here are based on quasicrystalline structures, the method of matching dissimilar structures based on their rotational symmetries is applicable to wide range of structures. As a trivial example, we return to our previous problem of detecting crystal structures in the 2d LJG system. Suppose now that we require an order parameter that detects crystalline grains in general, either hex or hc. In this case, we can use a scheme based on the $\ell = 6$ Fourier descriptor ψ_6, since in the space of $\ell = 6$ harmonics, the triangle neighbor shells of the hc structure and the hexagon neighbor shells of the hex structure are equivalent.

7.2.16 Orientation About a Symmetry Axis

As a final example of an application that exploits the symmetry properties of Fourier coefficients, consider the problem of aligning a crystal structure about a particular symmetry axis. This is useful, for example, for computer algorithms that depend on the orientational direction of a structure, such as automatically computing the diffraction pattern or rendering images of particle data. Suppose, for example that the desired symmetry axis is the $\ell = \ell_0$ plane, and the desired alignment direction is the z-axis. The crystal can be iteratively rotated to maximize the Fourier coefficient $|\psi_{\ell_0}|$ where all particles are projected into the xy plane. If the crystal consists of a single grain, a relatively small test cluster can be used to perform the optimization, greatly enhancing computational efficiency.

As an example, consider the problem of aligning the face-centered cubic (fcc) crystal shown in Fig. 7.21a along an 8-fold planar symmetry axis. Finding the optimal rotation that maximizes $|\psi_8|$ in the plane is solved by applying a simple simulated annealing Metropolis Monte-Carlo (MC) scheme. Our MC scheme involves attempting trial rotations of a small test cluster from the center of the box, which are accepted according to the Hamiltonian $\beta|\psi_8|$, with the fictitious energy function $-|\psi_8|$. The inverse temperature β is increased

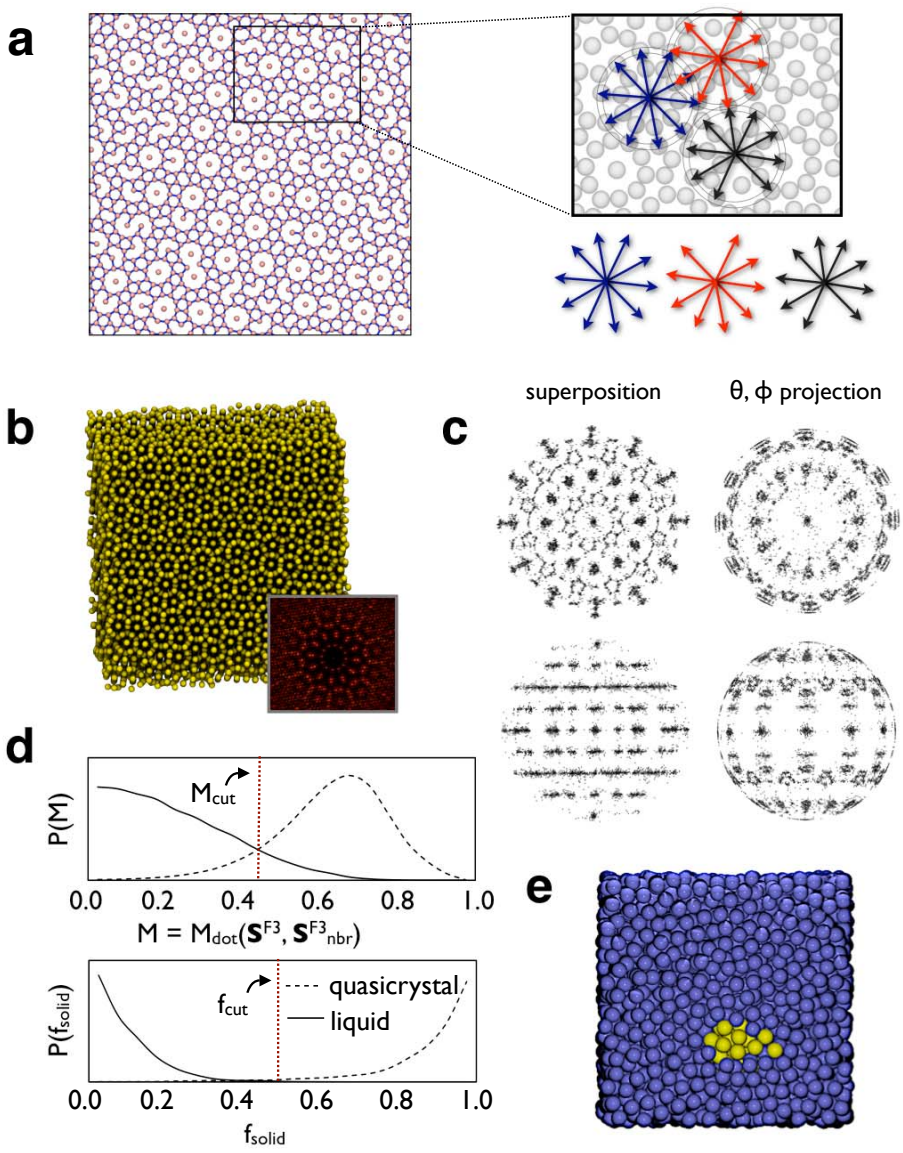


Figure 7.20 Matching based on rotational symmetries. (*a*) Decagonal (10-fold) quasicrystal formed in the 2d LJG system [120]. The call-out depicts local structures within the system, which are 10-fold symmetric but dissimilar in shape. (*a*) Dodecagonal (12-fold) quasicrystal formed in the 3d Dzugutov system by MD simulation [119], with a simulated diffraction image (lower right). (*b*) superposition of all neighbor clusters with a cutoff range 2.31σ. The images on the right disregard r-dependence. (*c*) Probability distributions of bond correlations and solid-like neighbors used to determine the criterion for local quasicrystal grains. (*d*) Dzugutov liquid with a single small quasicrystal nucleus of about 50 atoms.

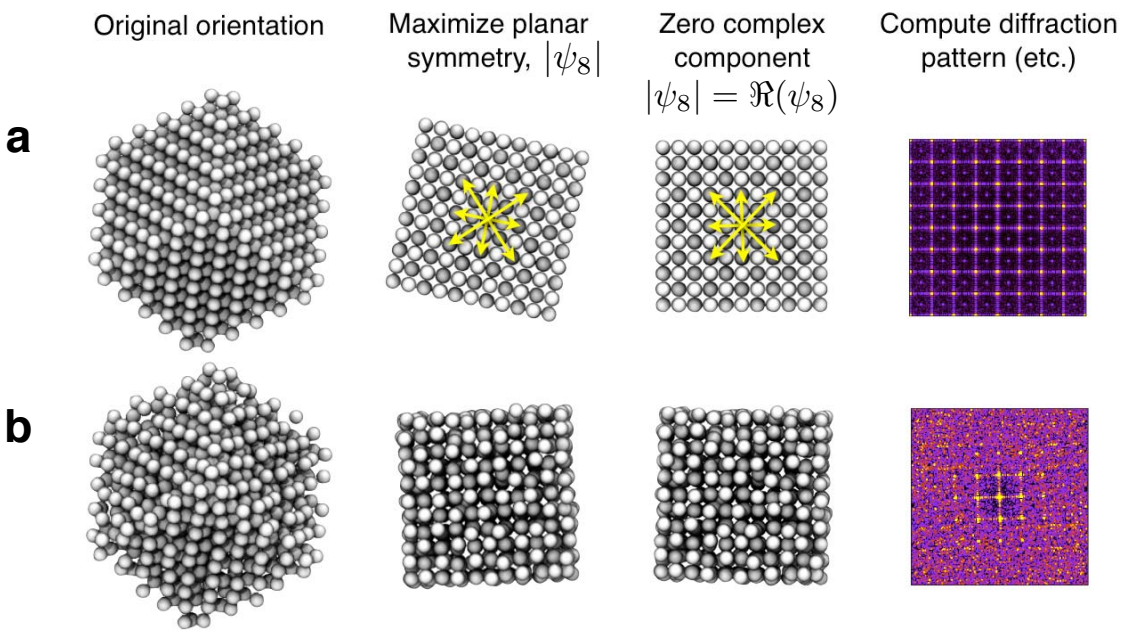


Figure 7.21 Orientation about a symmetry axis. (*a*) Depiction of an algorithm for orienting a face-centered cubic (fcc) crystal about an 8-fold planar symmetry axis, in this case, for the purpose of automatically computing the diffraction image. First, the planar symmetry, measured by $|\psi_8|$, is maximized such that the 8-fold symmetry axis is oriented in the z-direciton (out of the plane). Then the structure is rotated in the xy plane to zero the complex component $\Im(\psi_8)$, such that the symmetry axis is oriented with the xy axis. Finally, the diffraction image is computed. (*b*) The same process, repeated for a structure with thermal disorder.

from 0.5 to 1000 over 100 steps. In practice, the scheme only converges to a local minimum in the energy $-|\psi_8|$, so many different initial orientations are attempted to find a global minimum. As depicted in Fig. 7.21, optimizing $|\psi_8|$ aligns the 8-fold symmetry axis of the structure along the z-axis, but the structure is misaligned with the x and y-axes. To perform this alignment, we zero the complex component $\Im(\psi_8)$, such that $\Re(\psi_8) = |\psi_8|$. As depicted in Fig. 7.21b, this scheme is quite robust, even under a large amount of thermal noise. More complex optimization algorithms may be applied to improve computational efficiency and accuracy over our simple MC scheme. This type of orientation algorithm is potentially useful for matching diffraction data, since a large number of simulated diffraction images for 3d structures can easily be computed automatically about various symmetry axes.

7.2.17 Summary and Future Outlook

In summary, we have demonstrated how bond order parameters, already defined for particle structures on the unit circle and sphere, can be extended to index structures on the unit

disk or ball. We have demonstrated how these bond order parameters can be used to create harmonic shape descriptors, which can in turn be applied to create unique, highly specialized order parameters and automatically identify unknown particle structures. In addition to the minimal proofs of concept reviewed here, more complex matching applications are explored in reference [37].

In the future, the ability to describe structures numerically lends itself to many novel applications. In the short term, matching applications can be used to automate structural analysis for large datasets. Shape descriptors can be used within the context of many of the enhanced computational algorithms used in self-assembly and computational biology, such as path sampling [157, 354] or metadynamics [158] in the context of pseudo order parameters or collective variables to guide the sampling. Shape descriptors can also serve as the basis for new optimization algorithms such as genetic algorithms [375], which often rely on energy rather than structure as a numerical measure of fitness. In addition to abstract computational applications, shape matching algorithms can be applied to experimental images to obtain quantitative insight into experimental data. By combining shape matching algorithms with new image processing schemes, much of the same information that we have obtained for simulation data can be obtained for experimental systems as well.

7.3 Characterizing Complex Particle Morphologies Through Shape Matching: Descriptors, Applications, and Algorithms

Many standard structural quantities, such as order parameters and correlation functions, exist for common condensed matter systems, such as spherical and rod-like particles. However, these structural quantities are often insufficient for characterizing the unique and highly complex structures often encountered the emerging field of nano and microscale self-assembly, or other disciplines involving complex structures such as computational biology. Computer science algorithms known as "shape matching" methods pose a unique solution to this problem by providing robust metrics for quantifying the similarity between pairs of arbitrarily complex structures. This pairwise matching operation, either implicitly or explicitly, lies at the heart of most standard structural characterization schemes for particle systems. By substituting more robust "shape descriptors" into these schemes we extend their applicability to complex particle structures. Here, we describe several structural characterization schemes and shape descriptors that can be used to obtain various types of structural information about particle systems. We demonstrate the application of shape matching algorithms to a

variety of example problems, for topics including local and global structure identification and classification, automated phase diagram mapping, and the construction of spatial and temporal correlation functions. The methods are applicable to a wide range of systems, both simulated and experimental, provided particle positions are known or can be accurately imaged.

7.3.1 Introduction

It has been long recognized that in condensed matter systems there exists a strong connection between thermodynamics and particle packing [259, 260, 261]. The spatial arrangement of particles in a given phase determines the free energy of the system, as well as, in many cases, the physical, chemical, electrical, and optical properties. As a result, it is natural to attempt to gain insight into systems in general by characterizing and monitoring both global and local structure. In standard condensed systems, this is typically achieved by constructing order parameters or correlation functions that are sensitive to the way particles are arranged. Several order parameters and correlation functions have been contrived for standard classes of condensed matter, including, e.g., systems of rod-like and spherical particles [262, 53, 57, 263, 264]. Standard examples include the $\bar{P}_2$ nematic order parameter for rod-like liquid crystalline systems and the bond order parameters of Nelson and coworkers for detecting crystalline ordering in systems of spherical particles [48, 132]. These types of standard metrics have found widespread use in both computational condensed matter physics as well as the colloidal sciences, where standard systems of rod-like or spherical particles are often studied.

In the emerging field of colloidal and nanoscale self-assembly, unique building blocks can form assembled morphologies that often deviate from those expected in traditional condensed matter systems [11, 253, 12, 13]. Beyond the self-assembly of spherical [256, 266] and rod-like [267, 257] particles [12, 11], examples of assembled systems include ordered structures formed from polyhedrally shaped [33, 268, 269, 210, 270, 271, 272, 273] or patterned particles [279, 280, 29, 281, 282], and phase-separated domains reminiscent of those formed by block copolymers and surfactants [253, 210, 284, 285, 286, 287, 288]. In these systems too, system stability and properties are, in many cases, strongly linked to their global structure and local packing [12, 254, 255, 256, 257, 258]. However, constructing general order parameters for assemblies of particles of complex shape and interaction anisotropy [11, 12] is considerably more challenging than for traditional condensed systems, where the particle shapes and morphologies are comparatively much more standardized. As a result of the increased complexity and vast design space, there are few "model problems" in nanoscale

self-assembly for which generally applicable order parameters can be defined. This dearth of structural metrics has lead many recent studies of assembled systems to rely heavily on visual inspection or *ad hoc* analysis for characterizing structures. This approach is not optimal, since visual inspection is often more time consuming and less accurate than mathematical analysis, and *ad hoc* analysis can be idiosyncratic, making it difficult to compare structures across independent studies.

In this article, we address the problem of creating general structural metrics for complex colloidal and nanoscale assemblies, and other systems with a high degree of structural complexity. To do so, we combine the physical insights underlying many standard condensed matter order parameters with the mathematical insights provided by the computer science field of "shape matching." We show that virtually all of the standard structural characterization schemes from the condensed matter literature can be broken down, fundamentally, into the problem of quantifying the degree to which structures match. Structural similarity, in turn, can be quantified by using robust "shape descriptors" from the field of shape matching, which can be applied to arbitrarily complex structures. We decompose several order parameters, correlation functions, and other standard structural characterization schemes into their core elements, such that they can be used with arbitrary shape descriptors to extend their applicability. Additionally, we introduce new, more abstract structural characterization schemes that can also be used with arbitrary shape descriptors, to solve novel problems that arise in computational self-assembly. The shape matching methods that we provide will facilitate the creation of new structural metrics that are standardized, improving accuracy and comparability, but are also still flexible enough to be applied to the new classes of complex structures that arise in assembly problems.

This article is organized as follows. In section 7.3.2, we provide an overview of the shape matching framework and terminology that we will employ and describe how it connects to some standard structural characterization schemes from the condensed matter literature. In section 7.3.3, we review some relevant shape descriptors from the shape matching literature that can be applied to assembled systems. In section 7.3.4, we introduce some simple "similarity metrics" that can be used together with the shape descriptors from section 7.3.3 to measure structural similarity. Finally, in section 7.3.5, we introduce general algorithms based on shape descriptors and similarity metrics that can be used to obtain various types of structural information for complex particle systems. To demonstrate the usage of these algorithms, we apply shape matching to systems in the fields of nanoscience, computational self-assembly and condensed matter. Our examples include identifying local and global structures, quantifying structural changes as a function of time or a control variable, constructing correlation functions, mapping structural phase diagrams, and grouping similar

structures. We cover a wide range of systems including ordered phases formed from spherical and point-like particles, a fluid of tetrahedrally-shaped particles with locally ordered motifs [33], self-assembled systems of tethered nanoparticles with various nanoparticle shapes [298, 314, 306], patchy colloidal tetrominoes [358], a helical ribbon formed from tethered nanorods [301], a model protein [371], gold nanowires [376,377], and small clusters of water molecules [378]. The examples that we provide are applicable to particle systems in general, provided that the particle positions and, in some cases, orientations, can be detected. Although not explicitly treated here, other data representations such as images or diffraction data can also be used to obtain structural metrics within the shape matching framework. To aid in the development and dissemination of shape matching techniques, we provide accompanying software and examples via the web [302].

7.3.2 Shape Matching Overview

The concept of quantifying how well structures match has been generalized in the context of computer science applications known as "shape matching" techniques [303] (see Fig. 7.22). Familiar applications include matching fingerprints and signatures [303], facial recognition [304] and medical imaging [305]. Shape matching schemes have already been applied to systems of particles, particularly in the realm of fast database searches for proteins and macromolecules [329, 343, 330, 344, 345, 310, 311]. In some specific cases, shape matching schemes have been explicitly applied to local structure identification in problems in condensed matter and nanoscale self-assembly [187, 309, 30].

The basic idea of shape matching is to "index" structures into mathematical fingerprints known as "shape descriptors," $\mathbf{S}$, and then compare them using a similarity metric $M(\mathbf{S}_i, \mathbf{S}_j)$ to obtain both a quantitative and qualitative measure of similarity between the structures. For mathematical simplicity, we constrain our shape descriptors here to be vectors containing an arbitrary number of components. Matching can then be performed using straightforward vector operations, based on, e.g., the degree of alignment of or distance between shape vectors. The values of similarity metrics $M(\mathbf{S}_i, \mathbf{S}_j)$ are scalars. Matching information is used to create order parameters and correlation functions, or to identify structures by comparing "query" structures to "reference" structures. Since we can choose virtually any structure as a reference, this scheme facilitates the creation of highly specific structural metrics. The workflow for an application within the shape matching framework is shown in Fig. 7.22.

To apply these ideas to particle systems, we begin by asserting that most standard structural metrics include an implicit concept of "matching." That is, an order parameter or correlation function typically tells us the degree to which a structure of interest matches

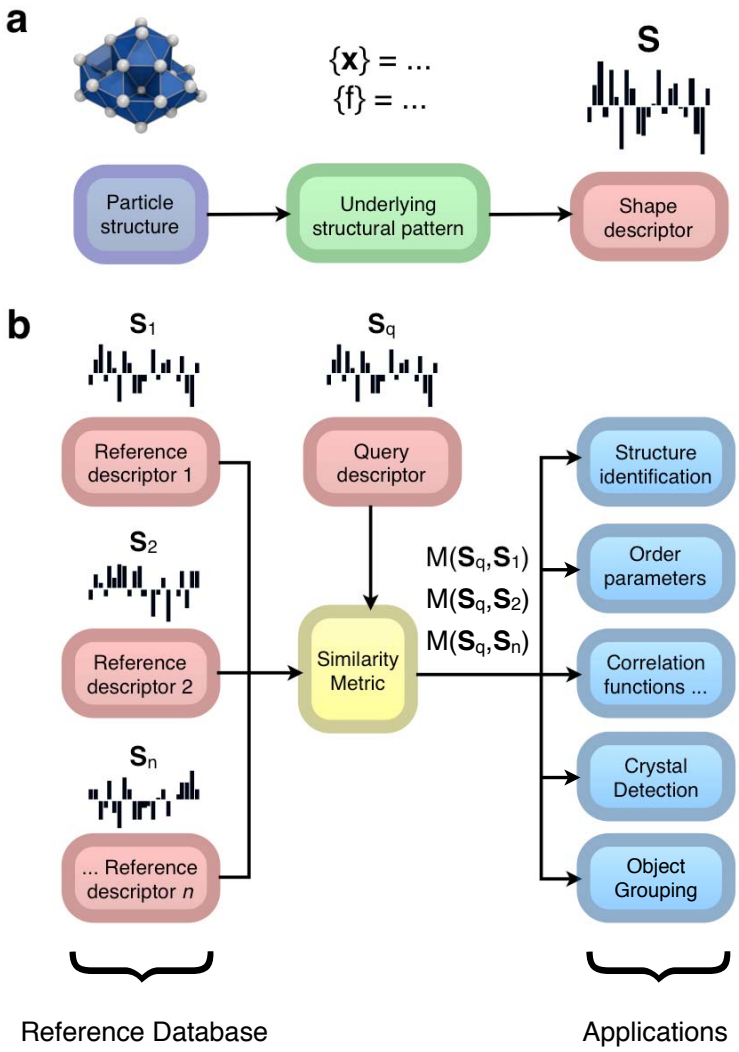


Figure 7.22 Data flow diagram for shape matching. (*a*) A representative pattern is extracted for a given particle structure and then indexed into a structural fingerprint known as a shape descriptor, **S**. The depicted cluster is an energy-minimized quantum Lennard-Jones cluster [379, 380]. (*b*) Shape descriptors are then compared to obtain similarity information M, which can be applied to within the context of various structural characterization schemes.

another (often ideal) structure. Most standard structural characterization schemes implicitly fit within the shape matching framework, and can be decomposed into query structures, reference structures, shape descriptors, and similarity metrics.

For example, consider the well-known order parameter $\bar{P}_2$ which detects nematic (aligned) liquid crystalline ordering:

$$\bar{P}_2 = \langle P_2(\cos\theta)\rangle = \left\langle \frac{3\cos^2\theta - 1}{2} \right\rangle. \tag{7.26}$$

The function P_2 is the second Legendre polynomial [381] and θ is the angle between the axis of the molecule and the "local director" $\mathbf{d}$ that indicates the preferred direction of the overall sample [262]. As the direction deviates from the preferred direction, P_2 decreases proportionally. The global nematic order parameter $\bar{P}_2$ is obtained by computing the ensemble average of P_2, denoted by angle brackets. In this scheme, the query structure is the system of particles and the reference structure is an ideal nematic liquid crystal with director $\mathbf{d}$. The shape descriptor is given by the collection of angles between the molecular axes and $\mathbf{d}$ and the similarity metric is given by the Legendre polynomial P_2. The order parameter $\bar{P}_2$ gives an optimal value of 1 when the structure matches a perfectly aligned liquid crystal with director $\mathbf{d}$, and tends toward zero the more the structure deviates from this ideal case.

As another classical example, consider the hexatic correlation function for 2d systems of spherical particles, or disks [48, 54]:

$$g_6(r) = \left\langle \sum_{i\neq j} \frac{\psi_6(i)\psi_6^*(j)\delta(r - |\mathbf{r}_i - \mathbf{r}_j|)}{\left[\delta(r - |\mathbf{r}_i - \mathbf{r}_j|)\right]} \right\rangle. \tag{7.27}$$

The quantity ψ_6 is a "bond order" parameter, defined as [53, 48]:

$$\psi_6(i) = \frac{1}{n}\sum_j^n \exp(i6\theta_j). \tag{7.28}$$

Here, n is the number of atoms in the first neighbor shell of an atom i, and θ is the direction of neighboring atom j. The value of $\psi_6(i)\psi_6^*(j)$ approaches $\sqrt{2}$ when the particles i and j are both in hexagonal local environments with the same spatial orientation, and varies towards 0 otherwise. Thus, like the standard radial distribution function $g(r)$, $g_6(r)$ measures the degree of spatial ordering; however, whereas $g(r)$ is sensitive to translational ordering generally, $g_6(r)$ is specifically sensitive to aligned hexagonal ordering. In this scheme, the query and reference structures are the pairs of neighbor shell clusters with for atoms i and j, which are a distance r apart, the shape descriptors are the local values of ψ_6, and

the similarity metric is the 'dot' operator, which measures the coherence between two ψ_6 descriptors. Thus, $g_6(r)$ measures how closely the local environment of a particle at the origin matches with that of a particle a distance r away in terms of both hexagonal shape and spatial orientation.

Although $\bar{P}_2$ and $g_6(r)$ are specific schemes, the physical insights underlying them are general. Recasting standard schemes within the shape matching framework allows us to obtain the same types of information, but with different shape descriptors $\mathbf{S}$ and similarity metrics M that are better suited for the unique and complex structures observed in assembled systems. The latter provides the main substance our article, but first, we introduce several shape descriptors and similarity metrics in the following sections.

7.3.3 Shape Descriptors

The shape descriptors that we describe in this section are adapted from the computer science field of shape matching. Since shape matching is a broad field, we focus on the small subset of methods that are best suited for particle systems and are applicable in both two and three dimensions. For a general review of some relevant shape matching methods, see references [317, 318, 319].

The first step towards creating an order parameter within the shape matching framework is to index the shapes representing the structure of interest into one or more shape descriptors $\mathbf{S}$. For simplicity, we consider in our framework shape descriptors to store structural information in a vector, which may contain real or complex components. However, shape descriptors may take other forms.

In addition to containing structural information, shape descriptors may possess other desirable properties and contain additional data, which may determine which descriptor is optimal for a particular application. One important property of shape descriptors is "invariance," defined as the ability for the descriptor to remain unchanged under certain mathematical transformations, such as scaling, translations, or rotations. In the context of particle systems, rotation invariance is a highly desirable property, since many applications involve comparing structures in a way that is independent of their spatial orientation. For descriptors without rotation-invariance, alignment or "registration [315, 316]" algorithms must be employed prior to matching to remove orientational dependence. Since particle systems often exhibit thermal noise, another desirable property of shape descriptors is robustness under small perturbations. However, this property must be balanced with the property of sensitivity, so that descriptors are still capable of detecting subtle structural differences. Another important consideration is the amount of computational time required

to compute and compare the descriptors, which may vary drastically for different schemes. Often, there is a direct tradeoff between computational cost and accuracy and attention to detail.

In the following sections, we provide a brief overview of some shape descriptors with different combinations of these properties that are well suited for self-assembled systems of particles. These descriptors are not representative of the full realm of possibilities, but rather are meant to serve as demonstrative examples. It is important to note that in principle, there are no limits on how the shape descriptor is calculated. Here, we constrain our analysis to descriptors that can be described mathematically as a vector, since this simplifies the process of writing general similarity metrics in section 7.3.4. However, in general, not all descriptors can be represented in this way, and thus require different similarity metrics.

Data Representations

Particle systems are typically represented as either a set of points (point cloud data) or solid objects (volumetric data). Both types of data can be represented by a set of position vectors $\{X\} = \{\mathbf{x}_1, \mathbf{x}_2, ...\mathbf{x}_n\}$ and weights $\{f\} = \{f_1, f_2, ...f_n\}$. Point cloud data $\{X\}$ typically represents particle positions, in which case the weights f_i are all 1. For volumetric data, the position vectors $\mathbf{x}_i$ represent the location of voxels (n-dimensional pixels) with intensities given by f_i. There is no formal rule regarding how to best represent input data for a given system. In general, point cloud data is optimal when particle shapes are not important, such as is the case with point particles. Volumetric data is optimal when particle size, shape, or orientation are important, such as with systems of rods or polyhedra, or the system is coarse-grained in space, such as with phase-separated structures. Image processing algorithms [294, 291, 298] can often be employed to change between the two data representations.

Point-Matching or RMS Descriptor

For relatively simple structures, such as small clusters or macromolecules comprised of particles (atoms, molecules, etc.), we can use the particle positions themselves (or a corresponding density map) as a shape descriptor (Fig. 7.23a). Mathematically, the point matching descriptor $\mathbf{S}^{\mathrm{RMS}}$ is defined trivially by the pointset $\{X\}$:

$$\mathbf{S}^{\mathrm{RMS}} = \{\mathbf{x}_1, \mathbf{x}_2, ...\mathbf{x}_n\}. \tag{7.29}$$

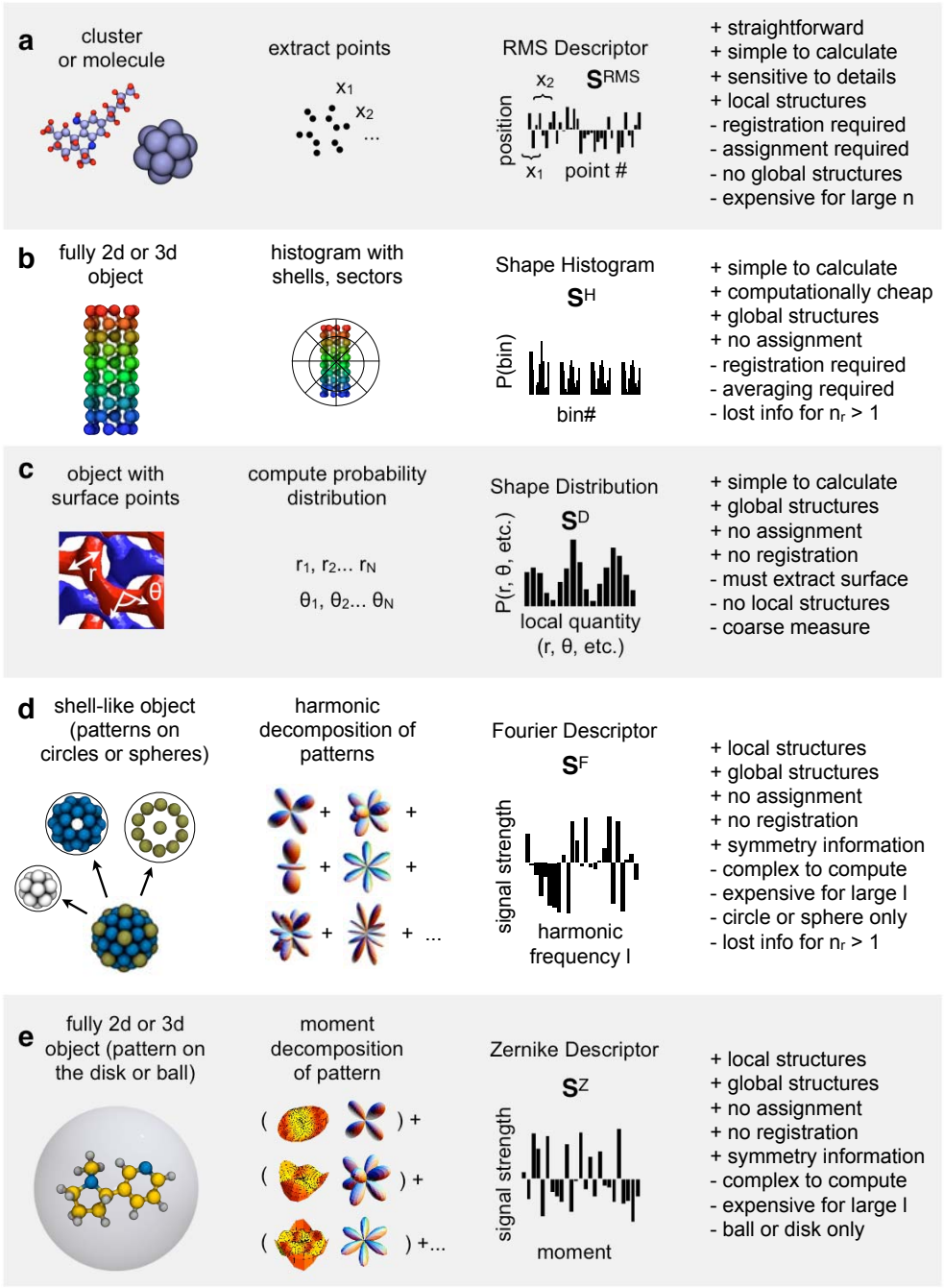


Figure 7.23 Depiction of five different shape descriptors. (*a*) The RMS descriptor [320, 315]. Descriptor components are given trivially by particle positions or density map. (*b*) The shape histogram descriptor [321]. The structure is indexed into a histogram consisting of n_r shells and n_θ sectors. (*c*) The D2 shape distribution descriptor [322]. The probability distribution is computed for various local measurements, such as the distance or angle between surface points. (*d*) The Fourier descriptor [323, 324]. A pattern along the perimeter of the circle or on the surface of a sphere is decomposed into a harmonic representation. (*e*) The Zernike descriptor [325, 326]. A pattern on the unit disk or unit ball is decomposed into a harmonic representation.

Here, each $\mathbf{x}_i$ is a d-dimensional vector representing the position of the ith point in $\{X\}$. The point-matching or $\mathbf{S}^{\mathrm{RMS}}$ shape descriptor [329, 320, 315] is a vector with $n \times d$ components. Typically the centroid is subtracted off and the vectors in $\{X\}$ are normalized, e.g. by dividing by the average distance between points. Matching for this scheme is often based on the root-mean-square (RMS) difference between points, and thus the scheme itself is often referred to as "RMS matching." Point matching schemes were applied in early attempts at shape matching macromolecules [329], and more complex variations have since been implemented [330]. Point matching schemes have the advantage of being conceptually simple; however, there are many subtleties associated with these schemes that can give rise to complications. First, point matching requires an assignment step to determine the optimal correspondence between points in compared structures. The coordinates in the shape descriptors are then re-ordered accordingly. As a coarse approximation, points can be assigned based on the minimum distance or the maximum dot product between individual coordinates. That is, a point i on the query structure, is assigned to the point j on the reference structure that maximizes the fitness $w_{i,j}$, which can be defined as, e.g., $w_{i,j} = f_i f_j |\mathbf{x}_i \cdot \mathbf{x}_j|$. The latter scheme has the disadvantage that it is possible to assign multiple points on the query structure to a single point on the reference structure. A more robust method involves creating a "fitness matrix" that records the degree of correspondence between all pairs of points:

$$F = \begin{pmatrix} w_{1,1} & w_{1,2} & \dots w_{1,n_r} \\ w_{2,1} & w_{2,2} & \dots w_{2,n_r} \\ \vdots & \vdots & \vdots \\ w_{n_q,1} & w_{n_q,2} & w_{n_q,n_r} \end{pmatrix}. \tag{7.30}$$

The variables n_q and n_r represent the number of points in the query and reference structures, respectively. We can then use a numerical technique, such as the Hungarian method [369], to efficiently determine the optimal assignment matrix that maximizes the overall fitness of the match. An additional subtlety arises when $n_q \neq n_r$. In this case, outliers can be excluded to obtain a "partial match" between structures. This is accomplished by sequentially removing points with the lowest total fitness w_i, defined as $w_i = \sum_{j=1}^{n} w_{i,j}$. The number of points excluded depends on the desired application. For partial matching, we might exclude $|n_q - n_r|$ points from whichever structure contains the fewest points. For excluding outliers, we might exclude all points with w_i below a certain threshold.

In addition to requiring assignment, the RMS descriptor also has the drawbacks that it is sensitive to scale, position, and orientation, and structures must first be normalized and registered unless the orientations are known beforehand, or the application utilizes rotation-dependent matching. Depending on the application, objects may be registered based

on rigid alignment, or other constraints [320]. Rigid registration can be achieved using either the iterative closest point (ICP) method [315], which involves minimizing the distance between points on compared objects by iterative rotations and translations, or the principle components analysis (PCA) method [316], which aligns objects with common principle axes. The ICP method has the disadvantage that it is non-trivial to implement, computationally expensive for structures with many points, and must be performed for all pairs of compared shapes. Moreover, it is prone to error if applied naively; the ICP method converges to a local minimum, so many initial orientations need be attempted to ensure the quality of the local minimum. The PCA method is only applicable to objects with distinct principle axes and thus fails for spherical objects. Despite the simplicity of the point-matching shape descriptor, implementation of the RMS method can often be non-trivial. Since both assignment and registration are computationally expensive (i.e. they scale poorly with n) point matching descriptors should be avoided unless (1) n is small, (2) matching is required for only a few structures, or (3) registration is not required.

Shape Histogram Descriptor

Another conceptually simple shape descriptor that has been applied to molecular database searches is known as the "shape histogram" [321] (Fig. 7.23b). This descriptor is based on a density map of the structure on a polar or spherical grid. The shape histogram is constructed in 2d by first generating n_θ equiangular gridlines on the unit circle:

$$\theta_i = 2\pi i/n_\theta \qquad i = 0,\ldots,n_\theta - 1. \tag{7.31}$$

The value of n_θ is chosen so as to capture important structural features while balancing computational efficiency. Typically, we take $2n \leq n_\theta < 4n$, where n is the number of data-points. Structures with radial dependence can be divided into n_r concentric shells. A given component in the 2d shape histogram descriptor is then given by:

$$S^{\mathrm{H2}}_{jn_\theta+k} = \sum_{i=1}^{n} f_i \delta\left(\left\lfloor \frac{n_r|\mathbf{x}_i|}{r_{max}} \right\rfloor - j\right)\delta\left(\left\lfloor \frac{n_\theta \theta(\mathbf{x}_i)}{2\pi} \right\rfloor - k\right). \tag{7.32}$$

The $\mathbf{S}^{\mathrm{H2}}$ descriptor then contains $n_\theta n_r$ real components, one for each bin in the histogram:

$$\mathbf{S}^{\mathrm{H2}} = \left\langle S^{\mathrm{H2}}_1, S^{\mathrm{H2}}_2, \ldots, S^{\mathrm{H2}}_{n_\theta n_r} \right\rangle. \tag{7.33}$$

The 3d version of the shape histogram is constructed in a similar way, except that in this

case, there are many different ways to construct the grid. An equiangular grid with and n_ϕ azimuthal bins and $\frac{1}{2}n_\phi$ polar bins is given by:

$$\theta_i = \pi i/n_\theta, \quad \phi_j = 2\pi k/n_\phi; \tag{7.34}$$

$$i = 0,1,\ldots,n_\theta - 1, \quad j = 0,1,\ldots,n_\phi - 1.$$

The total number of cells defined by the gridlines is $\frac{1}{2}n_\phi^2$. The 3d equiangular grid introduces artifacts near the poles of the sphere where the cells are small compared to the equator. Such artifacts are inherent to 3d grids on the sphere; there is no way to create an evenly-spaced grid on the sphere with equivalent cells. However, there are several alternatives to the equiangular grid, such as the rectilinear grid, icosahedral grid, etc., that give more evenly-sized cells [382]. A given component in the 3d shape histogram descriptor for an equiangular grid is given by:

$$S^{\mathrm{H3}}_{\substack{jn_\theta n_\phi \\ +kn_\phi \\ +l}} = \sum_{i=1}^{n} f_i \delta\left(\left\lfloor \frac{n_r|\mathbf{x}_i|}{r_{max}} \right\rfloor - j\right) \delta\left(\left\lfloor \frac{n_\theta \theta(\mathbf{x}_i)}{\pi} \right\rfloor - k\right) \delta\left(\left\lfloor \frac{n_\phi \phi(\mathbf{x}_i)}{2\pi} \right\rfloor - l\right). \tag{7.35}$$

As for $\mathbf{S}^{\mathrm{H2}}$, shapes with r-dependence are indexed by computing separate angular histograms for each radial shell. The $\mathbf{S}^{\mathrm{H3}}$ descriptor contains $\frac{1}{2}n_\phi^2 n_r$ real components, one for each bin in the histogram:

$$\mathbf{S}^{\mathrm{H3}} = \left\langle S^{\mathrm{H3}}_1, S^{\mathrm{H3}}_2, \ldots, S^{\mathrm{H3}}_{n_r n_\phi^2/2} \right\rangle. \tag{7.36}$$

The shape histogram has several advantages over the point matching method. No assignment step is required, since the ordering of points is lost during binning. Additionally, the grid resolution can be adjusted to provide a variable degree of coarse-graining. Like the point matching method, the shape histogram requires registration to match non-aligned shapes, unless only radial bins are used (i.e., $n_\theta = n_\phi = 1$). However, shape histograms may lose their discerning capabilities without an angular component. If n is large, the cost of registration can be significantly reduced by aligning the histograms themselves rather than the raw data. Shape histograms are best suited for describing structures that can be broken down into concentric circles or spheres. Examples include nanoparticle clusters, proteins and macromolecules. Shape histograms are also well suited for indexing global structures with orientational ordering such as crystals or quasicrystals, wherein the bond or neighbor directions of particles create a global pattern on the circle or sphere, as described in section 7.3.3.

Shape Distributions

For many applications, registration is costly and rotation-invariant descriptors are optimal. A simple yet powerful method for creating rotation-invariant descriptors is given by the "shape distributions" scheme [322] (Fig. 7.23c). This scheme involves creating distribution functions for simple invariant local metrics. The shape distribution "D2" is defined as the probability distribution of observing two surface points i and j a distance r apart. A given component in the D2 descriptor is given by:

$$S_k^{\mathrm{D2}} = \sum_{i \neq j} f_i f_j \delta \left[\left\lfloor n_r \frac{\left(\left| \mathbf{x}_i - \mathbf{x}_j \right| - r_{min} \right)}{r_{max} - r_{min}} \right\rfloor - k \right]. \tag{7.37}$$

The D2 descriptor is the collection of n_r radial components:

$$\mathbf{S}^{\mathrm{D2}} = \left\langle S_1^{\mathrm{D2}}, S_2^{\mathrm{D2}}, \ldots, S_{n_r}^{\mathrm{D2}} \right\rangle. \tag{7.38}$$

Notice that this function is similar to the standard radial distribution function $g(r)$, except that there is no ideal gas normalization and the function is typically computed only for points on the surface.

A similar distribution "A3" is defined by the probability of observing an angle θ between three surface points:

$$S_l^{\mathrm{A3}} = \sum_{i \neq j \neq k} f_i f_j f_k \delta \left(\left\lfloor \frac{n_\theta \mathbf{x}_j \cdot \mathbf{x}_k}{\pi \sqrt{(|\mathbf{x}_j - \mathbf{x}_i||\mathbf{x}_j - \mathbf{x}_i|)}} \right\rfloor - l \right). \tag{7.39}$$

The A3 descriptor is the collection of n_θ components:

$$\mathbf{S}^{\mathrm{A3}} = \left\langle S_1^{\mathrm{A3}}, S_2^{\mathrm{A3}}, \ldots, S_{n_\theta}^{\mathrm{A3}} \right\rangle. \tag{7.40}$$

Notice that this function is similar to the angular distribution function $a(\theta)$.

Similar distributions can be contrived for sets of four, five, etc. points; however, D2 and A3 were shown to have the best discerning capabilities for the structures tested in reference [322]. Shape distributions are best applied to structures with clearly defined, distinguishable surfaces, such as phase-separated structures formed by block copolymers [194, 285] or tethered nanoparticles [253, 210, 331, 306]. Like $g(r)$ and $a(\theta)$, shape distributions are too coarse to distinguish between similar shapes, such as small polyhedral clusters.

Fourier Descriptors

For shapes with more subtle differences, such as localized nanoparticle clusters, macromolecules, or global crystal structures, we can apply a more complex but more powerful technique for creating rotation invariants based on computing the harmonic transform of the shape histogram. By disregarding phase information from the harmonic transform, we obtain descriptors that are invariant under rotations. The formulae for the harmonic transform depend on the underlying basis. Invariants can be obtained for shapes on the unit circle [323] (θ-dependence), sphere [324, 383] (θ, ϕ-dependence), disk [325] (r, θ-dependence) or ball [326] (r, θ, ϕ-dependence). On the unit circle or sphere, the harmonic descriptors are known as Fourier descriptors (Fig. 7.23d). On the unit disk or ball, the descriptors are known as Zernike descriptors (Fig. 7.23e), which we discuss in the following section.

The Fourier descriptors are based on the Fourier transform, which involves decomposing a function into a sum of harmonic components. The Fourier coefficients for a 2d pattern are obtained by computing the discrete Fourier transform for each "shell" s of the 2d shape histogram, $\mathbf{S}^{\mathrm{H2}}$, defined in section 7.3.3:

$$\psi_{\ell,s} = \frac{\sum_{j=0}^{n_\theta - 1} S^{\mathrm{H2}}_{sn_\theta + j} \exp\left(-i\ell \frac{2\pi j}{n_\theta}\right)}{\sum_{j=1}^{n_\theta} S^{\mathrm{H2}}_{sn_\theta + j}}. \tag{7.41}$$

Here, n_θ is the number of sectors in each shell s in the shape histogram. By considering each shell independently, we reduce a 2d problem (a function of r and θ) to n_r 1d problems (functions of θ only). The coefficients ψ_ℓ are complex numbers.

Although the Fourier coefficients in their complex number form are not rotation-invariant (which may be beneficial for some applications), they can be converted to an invariant form by computing the magnitude of each coefficient. The invariant coefficients for a pattern on the circle are given by:

$$|\psi_\ell| = \psi_\ell \psi_\ell^* = \left[\Re(\psi_\ell)^2 + \Im(\psi_\ell)^2\right]^{1/2}. \tag{7.42}$$

Here * denotes the complex conjugate. The Fourier invariants are positive real numbers. To create a Fourier descriptor for a given shell s, we take a collection of desirable coefficients:

$$\mathbf{S}^{\mathrm{F2}}_s = \left\langle |\psi_{\ell_{min},s}|, |\psi_{\ell_{min}+1,s}|, \cdots |\psi_{\ell_{max},s}| \right\rangle. \tag{7.43}$$

Here, we use invariant coefficients; however, rotation-dependent coefficients are useful for many applications [132, 57, 62, 30]. The coefficients are sensitive to patterns with angular

frequencies that match the parameter ℓ. For example, ψ_4 is large for 4-fold patterns, ψ_6 is large for 6-fold patterns, etc. Specific coefficients can be chosen to describe structures with particular angular frequencies. In general, an arbitrary pattern can always be described by a sufficiently large range of ℓ. For the problems that we consider, we typically take ℓ in the range $\ell_{min} \sim 2$, $\ell_{max} \sim 10$. The overall Fourier descriptor is given by concatenating the descriptors for each shell into a vector:

$$\mathbf{S}^{\mathrm{F2}} = \left\langle \mathbf{S}_1^{\mathrm{F2}}, \mathbf{S}_2^{\mathrm{F2}}, \ldots, \mathbf{S}_{n_r}^{\mathrm{F2}} \right\rangle. \tag{7.44}$$

An analogous scheme can be used for 3d objects, where shells in the shape histogram have $[\theta, \phi]$ dependence. The Fourier coefficients are obtained by computing the discrete spherical harmonics transform for each "shell" s of the 3d shape histogram, $\mathbf{S}^{\mathrm{H3}}$:

$$\mathbf{q}_{\ell,s} = \frac{\sum_{j=0}^{n_\theta - 1} \sum_{k=0}^{n_\phi - 1} S^{\mathrm{H3}}_{\substack{sn_\theta n_\phi \\ +jn_\phi \\ +k}} N_\ell^m P_\ell^m \left[\cos\left(\frac{\pi j}{n_\theta}\right)\right] \exp\left(-im\frac{2\pi k}{n_\phi}\right)}{\sum_{j=1}^{n_\theta} \sum_{k=1}^{n_\phi} S^{\mathrm{H3}}_{\substack{sn_\theta n_\phi \\ +jn_\phi \\ +k}}}. \tag{7.45}$$

Here, N_ℓ^m is a normalization factor $\sqrt{(2\ell+1)(\ell-m)!/(\ell+m)!}$, and P_ℓ^m is a Legendre polynomial [381]. The variable m is an integer $m \in [-\ell, \ell]$. Therefore, unlike the circular coefficients ψ_ℓ, which are complex numbers, the spherical coefficients $\mathbf{q}_\ell$ are vectors with $2\ell+1$ complex components. Rotation-invariant versions of the coefficients can be obtained by computing the vector magnitude:

$$|\mathbf{q}_\ell| = \left(\frac{4\pi}{2\ell+1} \sum_{m=-\ell}^{\ell} |q_\ell^m|^2 \right)^{1/2}. \tag{7.46}$$

Like the Fourier invariants on the circle $|\psi_\ell|$, the Fourier invariants on the sphere $|\mathbf{q}_\ell|$ are positive real numbers. To create a Fourier descriptor for a given shell s, we take a collection of desirable coefficients:

$$\mathbf{S}_s^{\mathrm{F3}} = \left\langle |\mathbf{q}_{\ell_{min},s}|, |\mathbf{q}_{\ell_{min}+1,s}|, \cdots |\mathbf{q}_{\ell_{max},s}| \right\rangle. \tag{7.47}$$

Again, we have chosen invariant coefficients, but rotation dependent coefficients may also be used. The overall Fourier descriptor on the unit sphere is given by:

$$\mathbf{S}^{\mathrm{F3}} = \left\langle \mathbf{S}_1^{\mathrm{F3}}, \mathbf{S}_2^{\mathrm{F3}}, \ldots, \mathbf{S}_{n_r}^{\mathrm{F3}} \right\rangle. \tag{7.48}$$

Again, different combinations of coefficients can be used to create shape descriptors with different levels of robustness, and sensitivity to particular symmetries.

By using harmonic descriptors we gain many of the same advantages of the shape histogram, but without the need to register the objects or histograms. Like the shape histogram, harmonic descriptors are well suited for describing a wide variety of shapes including nanoparticle clusters, proteins and macromolecules, crystals composed of arbitrarily shaped particles and, in some cases phase separated structures. Harmonic descriptors exhibit an inherent data smoothing mechanism; thus they are typically better-suited for describing small polygonal or polyhedral clusters than the shape histogram, which is prone to error without sufficient averaging. These properties, along with the unique ability to yield symmetry-specific information, have already been successfully applied to constructing orientational order parameters for small clusters of point particles and simple crystals in the context of bond order parameters [48, 54, 132, 53]. While the bond order parameters scheme focuses primarily on the numerical values of specific coefficients (often ψ_6 and q_6), the shape matching approach more closely resembles a signal processing application, where an entire spectrum of Fourier coefficients are utilized. Additionally, while the bond order parameters were defined for point clusters that form patterns on the circle or sphere, the descriptors introduced here can be applied to volumetric objects and objects with r-dependence. Notice that in the limit of infinitesimal angular bin size and a single radial bin ($n_\theta, n_\phi \to \infty$, $n_r = 1$), our definitions of ψ_ℓ and $\mathbf{q}_\ell$ become nearly equivalent to the bond order parameters, only differing by a sign in the complex exponential. This only makes the direct mathematical connection between harmonic descriptors and the Fourier transform more explicit; the change is otherwise inconsequential. We explore the properties of Fourier descriptors in more detail in a separate paper [38].

Zernike Descriptors

The Fourier descriptors introduced in the previous section have pseudo r-dependence. That is, radial information is incorporated by decomposing the structure into concentric shells and then computing independent descriptors for each shell. This is problematic for structures with a small number of sample points, such as small clusters, since random perturbations move points between nearby shells. A second, more subtle drawback occurs when attempting to distinguish between structures for which the shapes of the shells are similar, but the relative orientation of the shells within the structure are different. Since the descriptors are computed for each shell independently, rotation-invariant descriptors are insensitive to relative orientations within the structure [324]. As a result, in many cases, it

is preferable to compute harmonic descriptors with full r dependence, known as "Zernike descriptors [325, 326]" (Fig. 7.23e).

The coefficients of the Zernike expansion, known as "Zernike moments" are computed by adding a Zernike radial polynomial to the Fourier coefficients:

$$R_n^m(r) = \sum_{k=0}^{(n-m)/2} \frac{(-1)^k\,(n-k)!}{k!\,((n+m)/2-k)!\,((n-m)/2-k)!}\,r^{n-2k}. \tag{7.49}$$

Here, r is the radial distance from the origin $r \in [0,1]$, and m and n are integers, $n \geq m > 0$. The Zernike moments on the 2d unit disk are given by:

$$a_{n\ell} = \frac{(n+1)\sum_{j=0}^{n_r-1}\sum_{k=0}^{n_\theta-1} S^{\mathrm{H2}}_{jn_\theta+k} R_{n\ell}\left(\frac{j}{n_r}\right)\exp\left(-i\ell\frac{2\pi k}{n_\theta}\right)}{\pi\sum_{j=1}^{n_r}\sum_{k=1}^{n_\theta} S^{\mathrm{H2}}_{jn_\theta+k}}. \tag{7.50}$$

The moments are subject to the constraint that $\ell \leq n$ and $(n-\ell)$ is even. Each moment is a complex number. The rotation invariant Zernike moments on the unit disk are given by:

$$|a_{n\ell}| = a_{n\ell}a^*_{n\ell}. \tag{7.51}$$

The 2d Zernike invariants are positive real numbers. A Zernike descriptor can be created by concatenating the desired Zernike moments into a vector, for example:

$$\mathbf{S}^{\mathrm{Z2}} = \left\langle |a_{11}|, |a_{20}|, |a_{22}|, \ldots, |a_{\ell_{max},\ell_{max}}| \right\rangle. \tag{7.52}$$

Again, we have chosen invariant coefficients, but rotation dependent moments may also be used. The Zernike moments on the 3d ball are given by:

$$\mathbf{z}_{n\ell} = \frac{3(n+1)\sum_{j=0}^{n_r-1}\sum_{k=0}^{n_\theta-1}\sum_{l=0}^{n_\phi-1} S^{\mathrm{H3}}_{\substack{jn_\theta n_\phi \\ +n_\phi k \\ +l}} R_{n\ell}\left(\frac{j}{n_r}\right) N_\ell^m P_\ell^m\left[\cos\left(\frac{\pi k}{n_\theta}\right)\right]\exp\left(-im\frac{2\pi l}{n_\phi}\right)}{4\pi\sum_{j=1}^{n_r}\sum_{k=1}^{n_\theta}\sum_{l=1}^{n_\phi} S^{\mathrm{H3}}_{\substack{jn_\theta n_\phi \\ +n_\phi k \\ +l}}}. \tag{7.53}$$

Again, we take $\ell \leq n$ and $(n-\ell)$ is even. Whereas the 2d Zernike moments $a_{n\ell}$ are complex numbers, the 3d Zernike moments $\mathbf{z}_{n\ell}$ are complex vectors with $2\ell+1$ components. The invariant Zernike moments on the unit ball are given by:

$$|\mathbf{z}_{n\ell}| = \sqrt{\frac{4\pi}{2\ell+1}\sum_{m=-\ell}^{\ell} |z^m_{n\ell}|^2}. \tag{7.54}$$

The 3d Zernike invariants are positive real numbers. A 3d Zernike descriptor is created by concatenating the desired moments into a vector, for example:

$$\mathbf{S}^{Z3} = \left\langle |\mathbf{z}_{11}|, |\mathbf{z}_{20}|, |\mathbf{z}_{22}|, \ldots, |\mathbf{z}_{\ell_{max},\ell_{max}}| \right\rangle . \tag{7.55}$$

Again, we have chosen invariant moments, but rotation-dependent moments and other combinations of frequencies may also be used depending on the problem.

Zernike descriptors are best applied to shapes that cannot be described by angles alone, such as certain clusters of nanoparticles, macromolecules, or complex crystals. When computing Zernike moments it is essential that the patterns being compared are normalized consistently on the unit ball or disk. Typically, normalization is performed by translating the centroid of the structure to the origin and rescaling the coordinates such that every point on the pattern has a radial distance less than 1. This scheme is sufficient for the majority of patterns that we encounter in assembled systems.

Combined Descriptors

In many cases, we can create new descriptors by taking linear combinations of the descriptors outlined above. Since the descriptors are represented as vectors, they can be concatenated together to combine their properties. Descriptors may also be multiplied by a weighting vector to (de)emphasize certain components. Other simple descriptor operations, such as averaging or taking probability distributions can also be useful, particularly for describing global structures, as outlined in the following section.

More complex combinations of descriptors can be created for specific applications. For example, one powerful solution to the problem of "partial matching" is given by the "shape contexts" method [327], which combines elements of the point matching descriptor with the shape histogram descriptor. A separate shape histogram is computed for each point in the structure, where the coordinate system is centered at that point. The points in the query structure are then assigned to their corresponding reference structure points by optimizing the match between shape histograms. Outlier points that do not correspond well can be excluded to obtain a partial match. Another powerful combined method is given by the light-field descriptor [328]. This method involves projecting 3D structures onto 2D images from 20 vantage points at the vertexes of a dodecahedron. Each projection is then indexed by a 2d descriptor, and assignment is performed to optimize correspondence between compared structures. In practice, many initial rotations of the dodecahedron are attempted to minimize error. Although the shape contexts and lightfield descriptors are specialized, the method of

combining descriptors to optimize properties is applicable to a wide range of problems.

Other Possible Descriptors

The shape descriptors that we have introduced above are meant to serve as representative examples rather than a complete set. In practice, any quantitative measure of structure can be used as a shape descriptor provided that it can be indexed into an n-dimensional vector. Along this line, there are several structural metrics from the condensed matter literature that could fit into the shape matching framework, and could potentially inspire useful shape descriptors for various applications. For example, in the realm of global structures, diffraction patterns, radial distribution functions, or orientation tensors (e.g. radius of gyration tensor or nematic order tensor [339]) could be indexed into shape descriptors. For local structures, local analysis schemes, such as the common neighbor analysis scheme of reference [263], could be easily be incorporated. Although many of the structural metrics from the literature may not be independently distinguishing for a wide range of problems, they may still yield useful information as part of a more general scheme through linear combination.

Extracting Global Patterns for Shape Descriptors

With the exception of shape distributions, the descriptors defined in the preceding sections are designed to index local structures such as small clusters of atoms or nanoparticles, macromolecules, or large but finite micro or nanoscale assemblies. Describing global structures is more difficult, since local shapes must first be extracted from the infinite system and then combined into patterns that reflect the "global shape" for indexing. The manner in which we construct global patterns depends on the structural properties of the system. In a rough sense, we can group global structures into two different categories: structures with long range orientational ordering (OO), and those without.

For structures with long-range OO, such as crystals and quasicrystals [313], the neighbor directions of all particles in the system are highly correlated. Thus, an intuitive global shape is given by the superposition of all neighbor directions for each local structure in the system [132], sometimes called a "bond order diagram [312]." This is depicted for the diamond structure [29] in Fig. 7.24a, top. As detailed in the previous sections, this type of pattern is best indexed by the shape histogram, or, for rotation-invariant matching, the Fourier descriptors or Zernike descriptors. In the case that it is important to distinguish between particle types, independent global descriptors should be created for each type independently, and combined later via concatenation. An example is given for the tetragonal

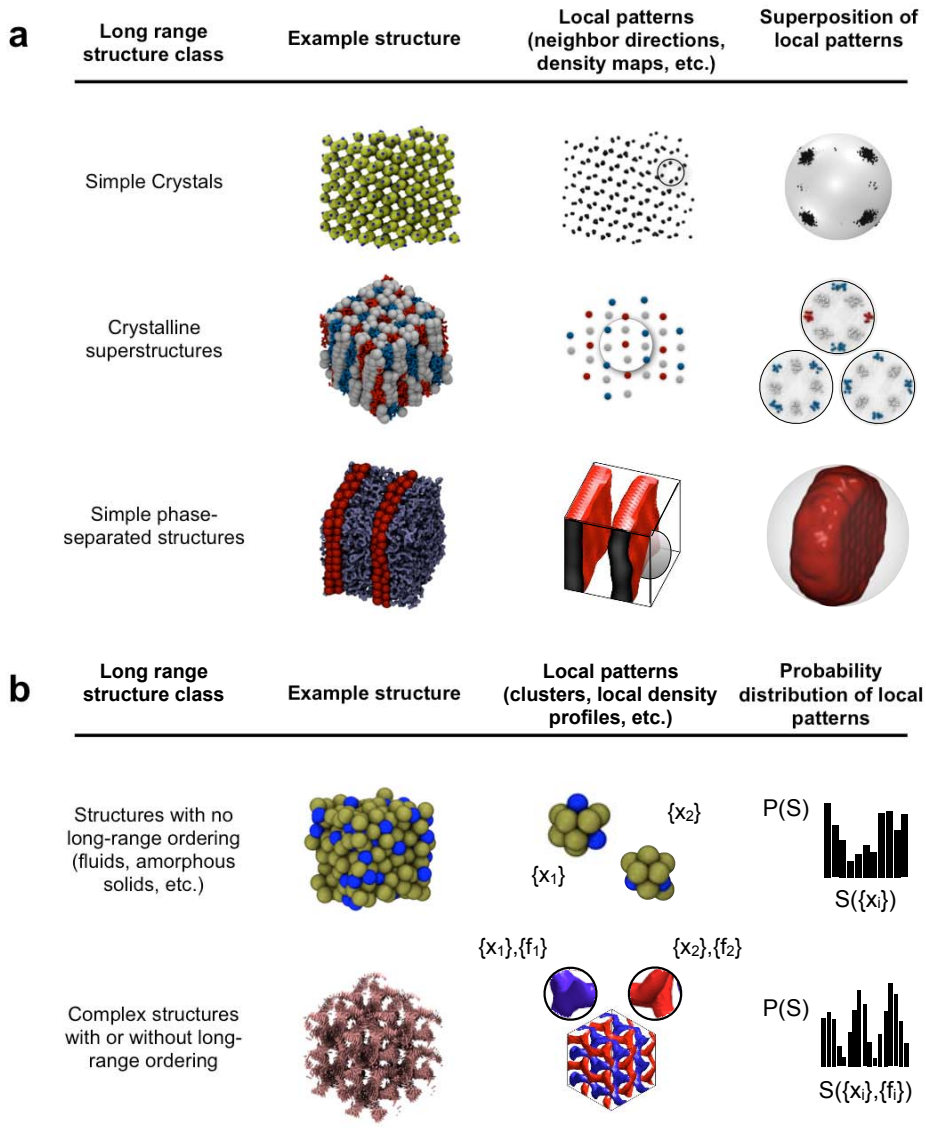


Figure 7.24 Depiction of strategies for extracting global patterns. (*a*) Global patterns by superposition. For structures with long range orientational ordering, a global pattern can be extracted by translating all local clusters or density maps to a common origin. (*b*) For structures with no long range orientational ordering or complex structures with many important directions, a global pattern can be built up from the probability distribution of local patterns.

cylinder structure formed from tethered nanospheres [298] in Fig. 7.24a, middle. Global descriptors based on orientational ordering are applicable to crystalline structures in general, including phase-separated systems arranged in crystalline superstructures [298, 314]. In this case, the neighbor directions are computed for the centers of the micelles, cylinders, etc. rather than the individual particles.

Non-crystalline globally-ordered phase-separated structures such as layered or network structures can be approached in a similar way. However, rather than creating a descriptor based on the superposition of local neighbor directions, a global descriptor is built up based on the superposition of local density maps. An example is given by the lamellar structure formed by tethered nanospheres [331] in Fig. 7.24a, bottom. The resulting patterns can be indexed by shape histograms, Fourier descriptors, Zernike descriptors, etc. in the same way as for crystalline long range order. To capture ordering on a range of lengthscales, descriptors should be created with a radial component that spans the lengthscales of interest.

For systems with no long range ordering such as liquids, gases and amorphous solids, a different approach must be employed. Rather than combining neighbor directions or density maps by superposition, we compute the probability distribution of these local patterns. This method is depicted for a dense liquid [133] in Fig. 7.24b. Since this requires a separate descriptor for every local structure, registration becomes computationally prohibitive. Thus, rotation-invariant descriptors, such as Fourier descriptors or Zernike descriptors, are typically optimal. Computing probability distributions is also useful for complex structures regardless of long range ordering. For example, for the double gyroid structure shown in Fig. 7.24b, bottom [309], the superposition of local density maps may become non-distinguishing for the global sample since there are many different directions, and probability distributions may present a better alternative. As mentioned in section 7.3.3, complex phase-separated structures can often be distinguished by shape distribution descriptors. However, while these descriptors are simple, they yield only a coarse measure of the shape, and thus can be non-distinguishing for similar structures.

7.3.4 Similarity Metrics

The degree to which two shape descriptors match [340] is quantified by a scalar similarity metric $M(\mathbf{S}_i, \mathbf{S}_j)$. Since shape descriptors are vectors by construction, standard vector operations such as the Euclidean distance or vector projection provide natural similarity metrics. Along this line, two standard similarity metrics, $\mathbf{S}_i - \mathbf{S}_j$ and $\mathbf{S}_i \cdot \mathbf{S}_j$, are defined. The

similarity metric based on the Euclidean distance is given by:

$$\mathbf{S}_i - \mathbf{S}_j = \left[\sum_k (S_{i,k} - S_{j,k})^2 \right]^{1/2}. \tag{7.56}$$

Here, k is one component of the shape vector $\mathbf{S}$, which may be a real or complex number. Similarly, a similarity metric based on the projection of one shape descriptor vector onto another is defined by

$$\mathbf{S}_i \cdot \mathbf{S}_j = \left[\sum_k (S_{i,k} S^*_{j,k}) \right]^{1/2}. \tag{7.57}$$

For the sake of comparison, it is useful to define the similarity metrics on the interval $M \in [0,1]$, with $1(0)$ giving the maximum (minimum) match. Thus, we redefine the Euclidean distance similarity metric as:

$$M_{dist}(\mathbf{S}_i, \mathbf{S}_j) = 1 - \left[\left(\mathbf{S}_i - \mathbf{S}_j \right) / \left(|\mathbf{S}_i| + |\mathbf{S}_j| \right) \right]. \tag{7.58}$$

Similarity, we redefine the projection-based similarity metric as:

$$M_{dot}(\mathbf{S}_i, \mathbf{S}_j) = \frac{1}{2} \left[1 + \left(\mathbf{S}_i \cdot \mathbf{S}_j \right) / \left(|\mathbf{S}_i||\mathbf{S}_j| \right) \right]. \tag{7.59}$$

The modified similarity metrics also have simple geometric interpretations. The M_{dist} function is the ratio of the Euclidean distance between vectors and the maximal distance between the vectors (i.e. if the vectors antiparallel). The M_{dot} function is proportional to the degree of spatial alignment between descriptor vectors. If $\mathbf{S}_i$ and $\mathbf{S}_j$ are parallel, M_{dot} has a value of 1. If $\mathbf{S}_i$ and $\mathbf{S}_j$ are antiparallel, M_{dot} has a value of 0. After normalization, the only difference between similarity metrics is the proportional weight given to the two types of differences. Matching functions based on projection are sensitive to differences in the signs of components, whereas distance-based metrics are only sensitive to the magnitude of differences regardless of the sign.

In addition to these metrics, we can define a wide variety of other metrics that are sensitive to particular differences in shape descriptors. For existing metrics, differences or correlations can be dampened or accentuated by applying an arbitrary power p to the component-wise comparison. In some cases, highly specialized similarity metrics can be applied to specific descriptors. An example of a specialized matching scheme is given by the quadratic metric of the shape histogram method of reference [322], which takes into account neighboring histogram bins when computing differences.

7.3.5 Algorithms and Examples

The shape descriptors and similarity metrics described in the previous sections can be used to create various types of order parameters, correlation functions, and other structural metrics. In this section, we describe general algorithms that, when used with the appropriate shape descriptors and similarity metrics, can be applied to characterizing structure for a wide range of particle systems. In some cases, the algorithms are reformulated versions of standard schemes from the condensed matter literature. Additionally, we explore algorithms from the shape matching literature that have not yet been widely applied to particle systems, and present some completely new algorithms that exemplify the future direction of the framework. For all of the algorithms, we provide representative example problems to demonstrate their application. Our examples are mostly drawn from the self-assembly literature; however, in some cases we explore more idealized problems from the condensed matter literature for simplicity. Since our goal is to present important elements of the shape matching framework rather than solve specific problems, the examples should be considered proofs-of-concept rather than optimal solutions.

Simple Structure Identification

The goal in most computer science shape matching applications is to identify unknown structures by searching a database of known reference structures. Structures are identified by the known structure that gives the best match. The algorithm for structure identification is given in pseudocode below.

```
set match_best = 0
set id_best = 'none'
call compute_shape_descriptor(S_i)

for each structure j in reference_database
   call compute_shape_descriptor(S_j)
   set match = M(S_i, S_j)
   if match > match_best
       match_best = match
       set id_best = id_j
   end if
end for
return id_best
```

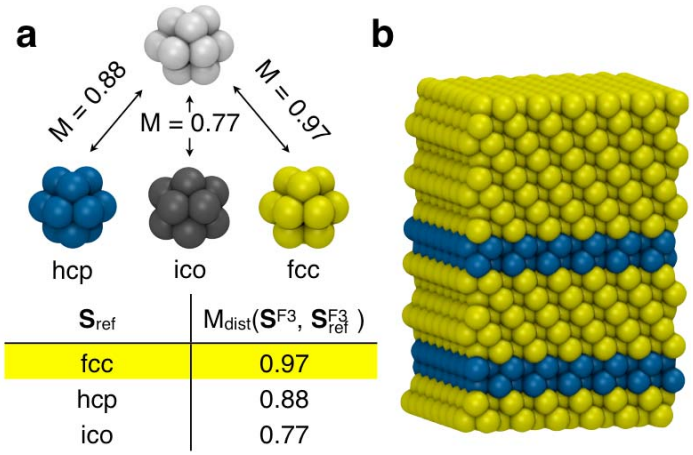
This type of database search has already been applied to particle systems in the context of proteins and macromolecules [329, 343, 330, 344, 345, 310, 311]. Although database searches have been applied in limited cases for condensed matter systems [187, 309], many standard structure identification schemes bear strong resemblance to these types of schemes. For example, the common neighbor analysis (CNA) scheme of reference [263] involves

constructing numerical fingerprints for pairs of atoms based on their local neighbor configurations, and identifying local clusters by matching the distribution of fingerprints with those for predetermined ideal structures. In the language of shape matching, the collection of CNA fingerprints can be considered a shape descriptor, and the catalogue of ideal fingerprints can be considered a database of reference structures. A similar identification scheme is given by the bond order parameters of reference [132]. Here, local structures can be identified by choosing cutoff values for bond order parameters [346]. In this case, the bond order parameters represent shape descriptors, the cutoffs act as similarity metrics, and the ideal structures used to set the cutoffs act as the reference database.

As a minimal example of a structure identification scheme, consider the problem of identifying the small, imperfect cluster in Fig. 7.25a, where particles have been slightly perturbed from their ideal face-centered-cubic (fcc) positions. For the purpose of the present example, we consider a small library of reference structures consisting of fcc, hexagonal-close-packed (hcp) and icosahedral clusters, each with 13 atoms. To differentiate between these clusters, we use rotation invariant Fourier descriptors on the surface of the sphere $\mathbf{S}^{\mathrm{F3}}$, where $\ell = 4, 6$. These coefficients are chosen because they are the leading coefficients for this class of structures [132]. As shown in the table in Fig. 7.25a, the unknown structure best matches with the fcc cluster, followed by the hcp and icosahedral clusters, thus identifying the structure as fcc.

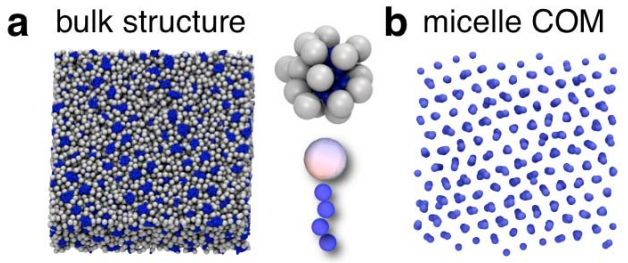
This simple matching scheme can be performed repeatedly to identify local structures in a global sample. Consider the defective fcc crystal shown in Fig. 7.25b, which contains hcp stacking faults [293]. The stacking faults can be identified by finding particles in local hcp configurations rather than fcc. First, local structural patterns must be created for each particle. This is done by clustering all neighboring particles within a cutoff radius r_{cut}. Here, the cutoff is chosen to encompass the first peak in the radial distribution function $g(r)$; first neighbors can alternatively be found by the Voronoi construction [384]. Since ideal fcc and hcp clusters are the only possible structures, our reference library consists of these two structures exclusively. Particles are identified by finding a best match, as in the previous example, and colored based on their local configuration (light corresponds to fcc, dark to hcp) in Fig. 7.25b, highlighting the stacking faults.

Structure identification can also be performed for global samples. Global structure identification can be useful, e.g., when mapping structural phase diagrams (see section 7.3.5). As an example of global structure identification, consider the mono-tethered nanosphere system similar to references [308, 210], whose tethers phase separate into spherical micelles (see Fig. 7.26a). The micelles themselves pack into an ordered crystalline superstructure. The structures of the crystal can be identified by identifying the micelle centers of mass,



S_{ref}	$M_{dist}(S^{F3}, S^{F3}_{ref})$
fcc	0.97
hcp	0.88
ico	0.77

Figure 7.25 Identification of local structures (*a*) Basic identification of a slightly imperfect fcc cluster. The table shows the matching values for the query structure compared to fcc, hcp and icosahedral reference clusters. (*b*) A fcc crystal with hcp stacking faults. The particles are colored based on their first neighbor shell configuration Light (yellow) particles are in the fcc configuration, while dark (blue) particles are in the hcp configuration.



reference	$M_{dot}(S^{F3}, S^{F3}_{ref})$
FCC	0.94
HCP	0.95
BCC	0.99
SC	0.87
Diamond	0.87

Figure 7.26 Identification of global crystalline structures for a system of mono-tethered nanospheres that aggregate into spherical micelles. (*a*) Bulk micelle structure. (*b*)The micelle centers-of-mass are extracted using a Gaussian filter. (*c*) The global pattern is created by superposition of the local patterns. (*d*) Matching is based on a Fourier descriptor (section 7.3.3) that indexes the global superposition of local patterns (section 7.3.3), and identifies the micelles as bcc structured.

which make up the set of positions $\{X\}$ that describe the system (see Fig. 7.26b). The centers of mass are determined by applying a Gaussian filtering algorithm adapted from the colloidal science literature [294, 291]. A global crystalline pattern is determined by computing the superposition of local patterns (see Fig. 7.26c, similar to Fig. 7.24a). The global pattern is then compared to that for several standard candidate crystals, by matching Fourier descriptors for patterns on the surface of the sphere: $M_{dot}(\mathbf{S}^{\mathrm{F3}}, \mathbf{S}^{\mathrm{F3}}_{ref})$. Here, the Fourier descriptor is composed of the leading terms in the harmonic expansion for the standard crystals: $\mathbf{S}^{\mathrm{F3}} =< |\mathbf{q}_4|, |\mathbf{q}_6|, |\mathbf{q}_8|, |\mathbf{q}_{10}|, |\mathbf{q}_{12}| >$. Notice that we use invariant Fourier coefficients for rotation-independent matching. The unknown crystal is identified by the reference structure that gives the best match, in this case BCC (see Fig. 7.26d).

The structure identification applications presented in this section are successful because the potential reference structures are known *a priori.* However, the identification schemes fail if the unknown structure is not in the reference database. It is therefore important to carefully choose the appropriate reference structures for a given application. Often, optimal matches are obtained by using imperfect structures from the system rather than mathematically perfect structures for reference structures. As an added consideration, it is sometimes possible to obtain partial structures that are highly ordered, but are missing one or more particles. This is an important factor in phase separated systems, systems with physical boundaries, and systems with high variation in neighbor distances. For proper identification, partial structures must be added to the reference library explicitly [187], unless a shape descriptor capable of partial matching is used, such as a point matching descriptor or shape contexts [327].

Identification Without Reference Structures

As the number of potential structures grows, compiling a complete reference library becomes increasingly difficult. However, if the space of potential reference structures is finite, a reference library can be created on-the-fly (OTF), precluding the need to define reference structures *a priori.* To do so, a new reference structure is added to the library whenever no suitable match is found. The algorithm for identification without known reference structures is given in pseudocode below.

```
set match_best = 0
set id_best = 'none'
call compute_shape_descriptor(S_i)

for each structure j in reference_database
   call compute_shape_descriptor(S_j)
   set match = M(S_i, S_j)
   if match > match_best
       match_best = match
```

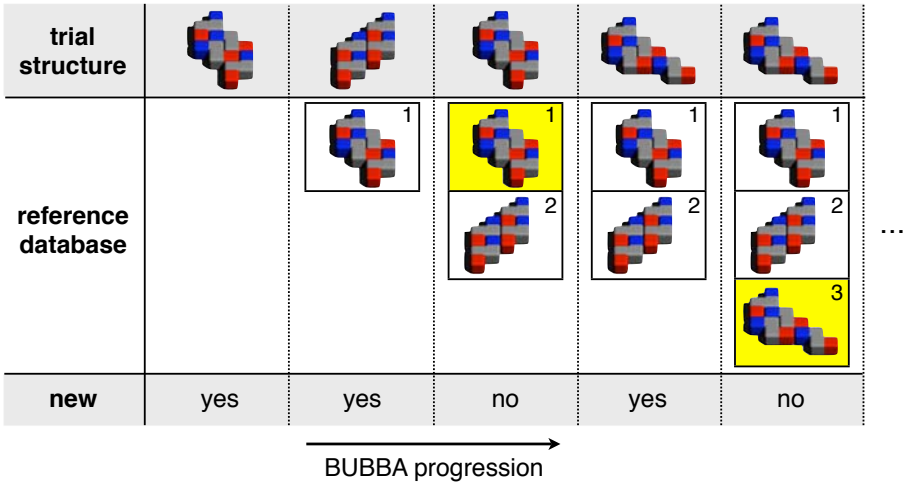


Figure 7.27 On-the-fly reference library with Bottom Up Building Block Assembly (BUBBA) [358]. The BUBBA algorithm involves enumerating unique clusters of a given size N. To ensure that clusters are unique, new clusters are added to the reference library and given a unique identifier, while repeated clusters (yellow) are discarded.

```
        set id_best = id_j
    end if
    if match_best < match_min
        call add_structure_to_database(S_i, counter)
        set id_best = counter
        set counter = counter + 1
    end if
end for
return id_best
```

Notice that the algorithm requires an additional step where the reference structures themselves, which are initially "unnamed," are identified. This can be accomplished by using a standard identification algorithm similar to that outlined in section 7.3.5, or in some cases more simply by visual inspection.

The OTF algorithm is applicable to simulations or algorithms that involve enumerating unique structures. One example is given by Bottom-Up Building Block Assembly (BUBBA) [358]. The BUBBA algorithm efficiently generates low-energy clusters by trying different combinations of smaller low-energy clusters. To ensure that the clusters generated are not redundant, an OTF shape matching scheme is employed (see Fig. 7.27). New clusters are added to a reference library if no match is found, while redundant clusters are ignored. In the end, the reference library contains a list of unique clusters. This type of scheme may also potentially be applied to information compression for mapping structural phase spaces. Since large portions of the parameter space are often redundant for high-resolution mappings, an OTF scheme can be employed to quickly obtain a minimal number of unique structures.

Identification in Systems with Disordered Structures

Creating a comprehensive reference library is nearly impossible when local structures can assume disordered configurations. In this case, the space of potential reference structures is essentially infinite, since "disordered" refers to the vast space of configurations with no particular structure. As a solution, a structure that does not match any structure in the reference library within a certain threshold is considered "disordered [187]." This requires that we choose a cutoff value for a best match. The cutoff must be chosen carefully; in thermal systems, an overly-stringent cutoff might cause a matching scheme to miss highly-ordered structures perturbed slightly from their ideal configurations, whereas an overly-permissive cutoff can misidentify highly disordered structures. In most cases, a sufficiently rigorous cutoff can be defined such that its value does not affect the qualitative results. The algorithm for structure identification with disordered local structures is given in pseudocode below:

```
set match_best = 0
set id_best = 'none'
call compute_shape_descriptor(S_i)

for each structure j in reference_database
   call compute_shape_descriptor(S_j)
   set match = M(S_i, S_j)
   if match > match_best
       match_best = match
       set id_best = id_j
   end if
   if match_best < disordered_cut
      id_best = 'disordered'
   end if
end for
return id_best
```

As an example, consider the hard tetrahedron fluid studied in reference [33] (Fig. 7.28a). In this system, an important local motif, originally identified by visual inspection, is the "pentagonal dipyramid" (PD), formed by five tetrahedra sharing a common edge. The PDs form a spanning network as the system goes through a liquid-liquid transition. To identify PDs, we first cluster all sets of five tetrahedra in the system that form a closed polygon. The shape of each cluster is defined by projecting the directions of the tetrahedra on the surface of a sphere (Fig. 7.28b). An ideal PD gives a 2D pentagonal pattern, which is taken as a reference structure. Although the pentagon reference structure is confined to a plane, the query structures are not. Therefore, this pattern is well-described by Fourier descriptors on the surface of the sphere, with matching given by $M_{dist}(\mathbf{S}^{\mathrm{F3}}_{query}, \mathbf{S}^{\mathrm{F3}}_{pentagon})$. We take rotation-invariant descriptors with frequency parameter $\ell = 5, 6, \ldots 10$. For some systems, there is a clear distinction between ordered and disordered structures. However, for the tetrahedron system we observe a continuous spectrum of PD-like ordering. Thus, we estimate a cutoff based on visual inspection in the range $M_{cut} \sim 0.9$. Although this choice

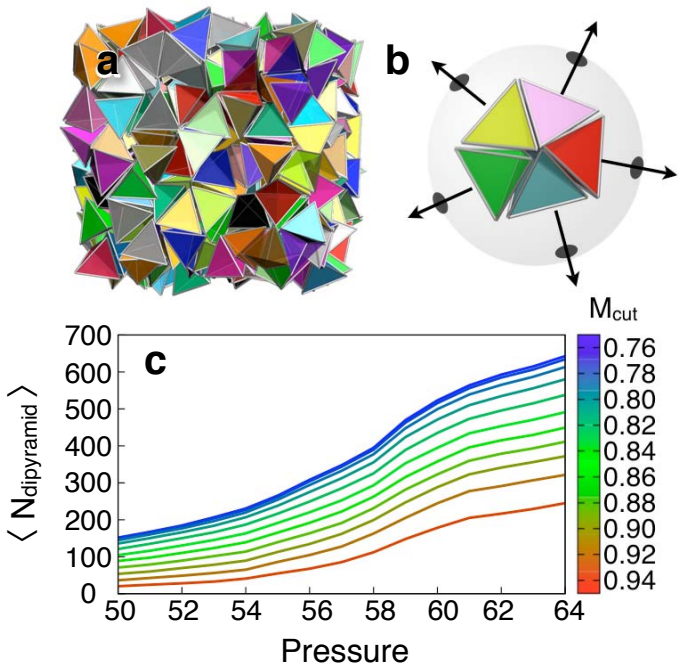


Figure 7.28 Pentagonal dipyramids (PDs) in the hard tetrahedron system [33]. (*a*) A snapshot of the hard tetrahedron liquid at packing density $\phi \approx 0.5$ and pressure $P = 60$. (*b*) A PD-like cluster taken from the system. The arrows depict the pattern of directions $[\theta, \phi]$ on the surface of the sphere indexed for matching. (*c*) The number of PDs as a function of the identification cutoff value. Notice that for all cutoffs, there is an inflection point centered at $P = 58$, which corresponds to a possible liquid-liquid transition marked by a sudden increase in PD-like local ordering.

is arbitrary, it has little effect on structural trends for the system. Fig. 7.28c shows the number of PDs for a wide range of cutoffs. We see that for all cutoffs, the fraction of PDs exhibits a weak crossover, marked by an inflection point, near $P = 58$. This pressure, in turn, corresponds to an interesting thermodynamic transition for the system [33]. Although cutoffs result in different numbers of PDs, the same underlying physical behavior is captured regardless.

Order Parameters and Temporal Correlation Functions

Another standard application of structural metrics is to track structural transitions, either as a function of time or a changing reaction coordinate. This is typically accomplished by monitoring either an order parameter or correlation function as the system changes. In the context of the shape matching framework, the difference between the two cases is largely semantic; while an order parameter typically measures similarity with an ideal structure, a

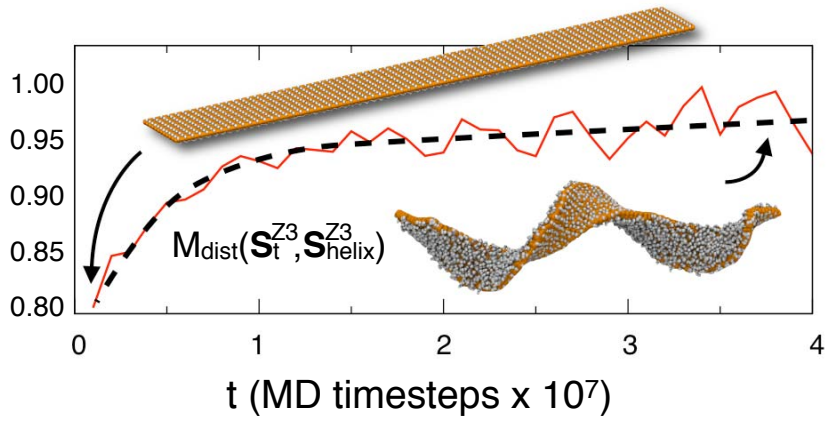


Figure 7.29 Assembly of a helical sheet composed of laterally tethered nano-rods [301]. As time progresses, the initially flat sheet twists into a helix. The matching order parameter $M_{dist}(\mathbf{S}_t, \mathbf{S}_{helix})$ compares the structure at time t with the shape of the final ideal helical structure.

correlation function typically measures similarity between different structures in the system, separated in time and/or space. The simple algorithm for tracking a transition as a function of a changing parameter is given below:

```
call compute_shape_descriptor(S_ref)
for p in changing_parameter
   call compute_shape_descriptor(S_p)
   set order_param[p] = M(S_p, S_ref)
end for
return order_param
```

Here the similarity metric $M \in [0,1]$ serves as a convenient order parameter. Tracking structural transitions is important for a wide variety of applications, including elucidating thermodynamic transitions [57, 348, 349, 350, 351] and assembly pathways [29, 352, 353, 268]. Many of the advanced molecular simulation techniques used to study transitions [156, 157, 60, 354, 158] rely on structural metrics in the context of pseudo-reaction coordinates [157], biasing parameters [156], and collective variables [158] to guide the statistical sampling algorithm. Standard order parameters have been devised for various types of ordering, including bond orientational ordering [48, 54, 132, 355], liquid crystalline ordering [341, 262] such as nematic [356] and smectic [349] phases, chiral ordering [264], and helical ordering [357]. Time correlation functions based on these types of order parameters have been applied to creating structural "memory" functions for glassy liquids [138, 139] and growing quasicrystals [30].

As a simple example of creating an order parameter within the shape matching framework, consider the sheet-like structure self-assembled from laterally tethered nano-rods studied in reference [301], and shown in Fig. 7.29a. Due to an instability, the initial sheet

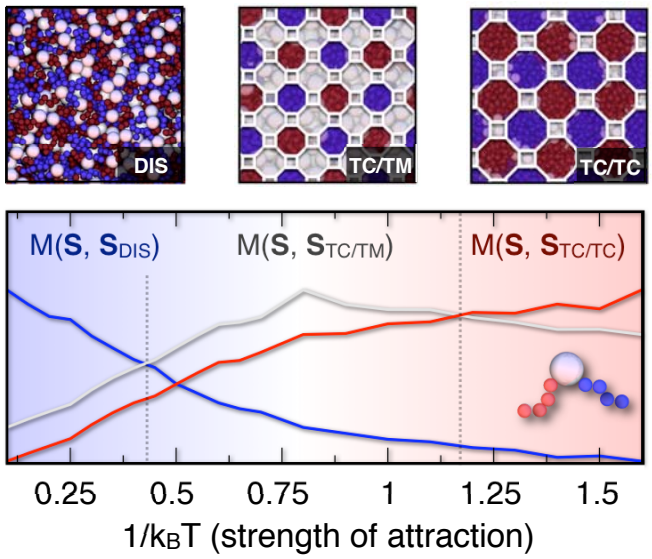


Figure 7.30 Structural transitions in a phase separated system. (*a*) Visual depiction of the three structures formed by a ditethered nanosphere system [314] (left to right: disordered, TC/TM, TC/TC). (*b*) Matching order parameter for the three reference structures as a function of inverse temperature.

relaxes into a helical structure that minimizes the free energy. We can track this structural transition by matching the shape of the sheet at a given time t with the final fully-equilibrated helical structure: $M(\mathbf{S}_t, \mathbf{S}_{helix})$. Since the structure is 3-dimensional and has radial dependence, we can index it using a Zernike descriptor on the unit ball, $\mathbf{S}^{Z3}$. Since the sheet only changes in terms of its twist in space, we save computational effort by only considering points along the backbone of the sheet. To match the shape independent of the orientation of the sheet, we take rotation invariant moments with ℓ in the range $4 \leq \ell \leq 12$. Fig 7.29a shows the helical order parameter as a function of time for a long molecular dynamics run. We observe that over tens of millions of MD steps, the sheet slowly and continuously equilibrates to the final helical structure, in agreement with visual inspection. Our matching order parameter gives a better indication of the structural transition than the more standard helical order parameter $H4$ [357], which is rather insensitive when the pitch of the helix is large compared to the radius [301]. The noise in the data at long t is indicative of the relatively large fluctuations in shape that occur in equilibrium.

As a slightly more complex example, consider the structures formed by the ditethered nanospheres shown in Fig. 7.30a [314]. The system goes through two transitions as a function of inverse temperature or quench depth, first from a disordered structure to a phase-separated structure characterized as a tetragonal cylinder/tetragonal-mesh (TC/TM), and then to a similar structure characterized by tetragonal cylinders(TC/TC) [314]. The

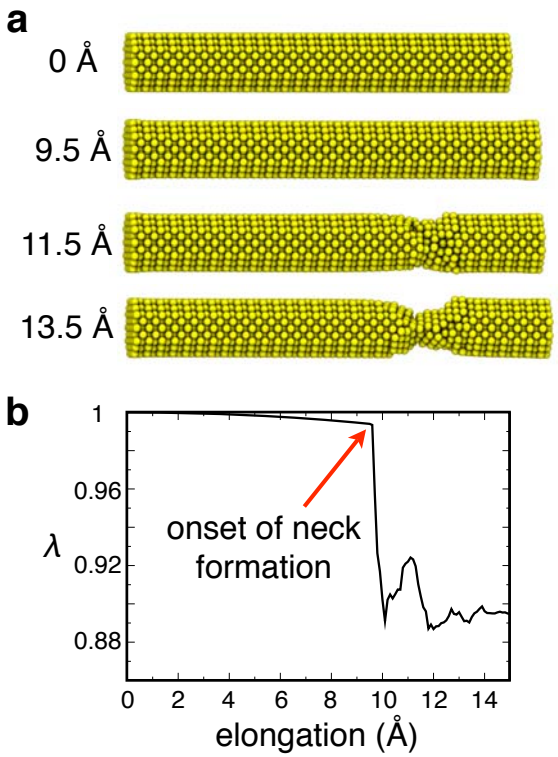


Figure 7.31 Neck formation in a gold nanowire [376, 377]. (*a*) depiction of the nanowire as a function of elongation L. (*b*) The standard deviation of matching values, $\lambda = 1 - < M_{dot}(\mathbf{S}_L^{\mathrm{F3}}, \mathbf{S}_{L=0}^{\mathrm{F3}}) >_{stdev}$ as a function of elongation.

abbreviations indicate the patterns formed by the tethers and nanoparticles, respectively. To obtain a quantitative measure of this behavior, we take three reference points: ideal snapshots from the disordered regime, the TC/TM regime and the TC/TC regime. As outlined in section 7.3.3, several different descriptors are applicable to this type of global structure. Here, we use the shape distribution $\mathbf{S}^{\mathrm{D2}}$, since it is distinguishing between the three compared structures. Separate $\mathbf{S}^{\mathrm{D2}}$ descriptors are created for each of the three "particle" types. These descriptors are then concatenated into an overall descriptor. Rather than considering surface particles exclusively, we use all particle positions, since this is simpler and is still distinguishing for the cylindrical phases under consideration. Fig. 7.30b shows how the character of the system changes as a function of inverse temperature. We see that the structural transition between the three phases is smooth and continuous, as verified by visual inspection. In a separate reference [38], we show that a scheme based on the distribution of Fourier descriptors for local density maps gives identical results.

As a relatively complex example, consider the gold nanowire undergoing tensile elongation shown in Fig. 7.31a [376, 377]. As the wire elongates, a "neck" begins to form, in this

case, at $\sim$10 $\mathring{A}$. This type of structural transition can strongly impact the transport properties of nanowires [385] and thus is important to identify. The neck region can be characterized by the loss of the original FCC structure locally, e.g. a change in orientation, number of neighbors, or overall symmetry. Global crystalline order parameters [132] may not be well suited, since only a subset of the system undergoes a transition. Standard schemes that differentiate between crystal and liquid configurations locally, such as $\mathbf{q}_6 \cdot \mathbf{q}_6$ [57] (see section 7.3.5), are not well suited either, since the finite nature of the nanowire results in neighboring atoms having different local coordinations, even in the ideal FCC configuration (i.e. a mixture of full and partially coordinated FCC clusters). Instead, to detect the onset of necking, we compare an atom's neighbor shell to its initial structure as a function of elongation L: $M(\mathbf{S}_L, \mathbf{S}_{L=0})$. Local neighbor shells are indexed using rotation-dependent Fourier descriptors $\mathbf{S}^{\mathrm{F3}} = \mathbf{q}_6$, where $\ell = 6$ is chosen because it has been shown to describe FCC clusters well without requiring other frequencies [57, 62]. The number of atoms in the neck is small compared to the bulk, thus the average autocorrelation value is not strongly sensitive to neck formation. However, the spread in the data is sensitive to neck formation, and drastically increases when atoms in the neck lose their original structure and yield low matching values. We can therefore create an *ad hoc* order parameter based on $\lambda = 1 - < M_{dot}(S_L^{F3}, S_{L=0}^{F3}) >_{stdev}$, where $\lambda = 1$ for the ideal configuration at $L = 0$, and decreases proportionally as the spread in matching values increases. Fig.7.31b shows that the onset of neck formation occurs at 9.7 $\mathring{A}$, which is consistent with visual inspection.

Spatial Correlation Functions

In addition to characterizing how structures change as a function of time or a reaction coordinate, another common application of structural metrics is to characterize how structures change in space. In the context of the shape matching framework, this involves choosing structures from different points in the system, rather than ideal structures, as reference structures. Spatial correlation functions are often used to measure structural "correlation lengths." The algorithm for computing structural correlation lengths within the matching framework is given in pseudocode below:

```
for i in list_of_local_structures
   call compute_shape_descriptor(S_i)
   for j in list_of_local_structures
       call compute_shape_descriptor(S_j)
       set r =  distance(i, j)
       set correlation_function(r) += M(S_i, S_j)
       set normalization(r) = normalization(r) + 1
end for
return correlation_function / normalization
```

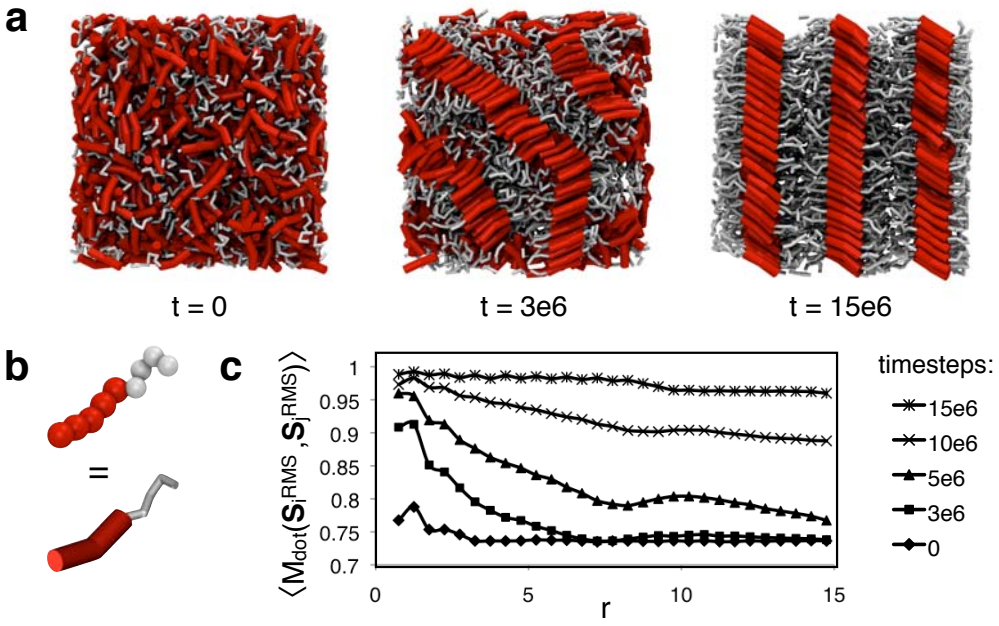


Figure 7.32 Spatial correlations in a system of tethered nano "V's" [306]. (*a*) Depiction of the formation of lamellar phase as the system evolves in time on cooling. (*b*) Depiction of the coarse-grained nanoparticle model. (*c*) Nanoparticle orientational correlations as a function of separation distance *r*.

In the condensed matter literature, structural correlation functions have been defined for crystal-like ordering in 2d [48, 54], and 3d [132, 57], nematic ordering [359], and many other more specialized types of ordering. Other types of spatial correlation functions have been widely applied as well. One example is the $\mathbf{q}_6 \cdot \mathbf{q}_6$ scheme of references [57, 62], which detects ordered crystal nuclei based on spatial correlations between local bond order parameters.

As a simple example of creating a spatial correlation function within the shape matching framework, consider the problem of characterizing the formation of lamella (sheets) in the system depicted in Fig. 7.32a, composed of tethered V-shaped nanoparticles (Fig. 7.32b) [306]. From visual inspection, it is clear that the nanoparticles have long-range orientational correlations in the lamella phase, but not in the disordered phase. This can be quantified by computing an orientational correlation function for nanoparticles as a function of separation distance *r*. It may initially appear that the nematic descriptor can be applied to this problem; however, since the nanoparticles have two directors, we lose important information about particle packing by considering only one angle. Rather, an optimal metric reflects the correspondence between both directors of the nanoparticles. This can be measured by a

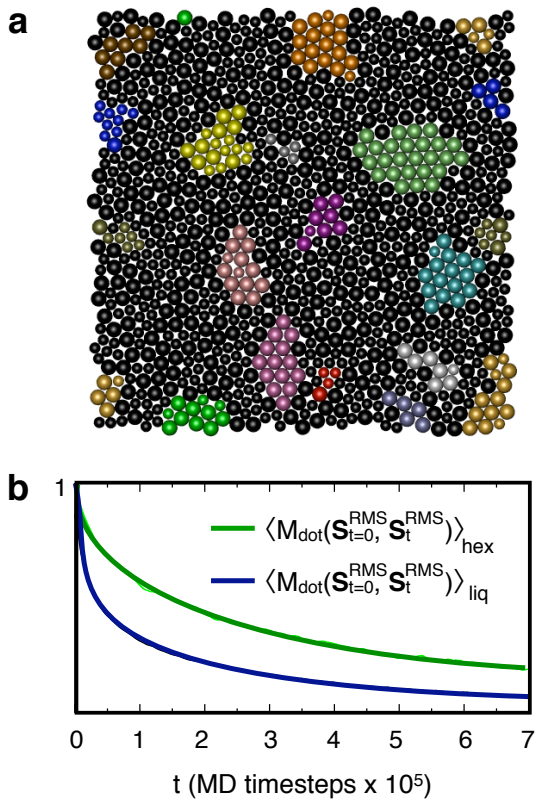


Figure 7.33 Spatial and temporal correlations in hexagonal clusters. *a*) Depiction of the 2d hexagonal grains identified in the system. Each grain is given a different color while particles in the bulk liquid are colored black. *b*) Time-dependent structural decorrelation function for the hexagonal grains and the overall liquid. Notice that the hexagonal grains retain their structure much longer than the overall liquid.

scheme based on the RMS descriptor, where the datapoints $\{X\}$ for each nanoparticle are given by the two directors, pointing from the vertex $\mathbf{x}_v$ to each of the endpoints $\mathbf{x}_1$, $\mathbf{x}_2$: $\{X\} = \{\mathbf{x}_1 - \mathbf{x}_v, \mathbf{x}_2 - \mathbf{x}_v\}$. Assignment of corresponding vectors is performed using the "naive" method (see section 7.3.3), which is exact for this particular problem. Fig. 7.32b shows the average value of the orientational correlation function $\left\langle M_{dot}(\mathbf{S}_i^{RMS}, \mathbf{S}_j^{RMS}(r)\right\rangle$ as a function of the radial separation r for several different snapshots. In the disordered phase, only very short range correlations are present and $\left\langle M_{dot}(\mathbf{S}_i^{RMS}, \mathbf{S}_j^{RMS}(r)\right\rangle$ is small for all r. The correlations quickly grow as the system begins to form sheets, and the range of $\left\langle M_{dot}(\mathbf{S}_i^{RMS}, \mathbf{S}_j^{RMS}(r)\right\rangle$ increases. In the final state, the lengthscale is infinite, spanning the length of the simulation cell.

As a slightly more complex example, consider the problem of measuring time-dependent structural correlations in the 2D binary mixture shown in Fig. 7.33a. The system consists of

a 50:50 mixture of spherical particles with a diameter ratio 1.4:1, and can represent either a model supercooled liquid [135,247] or a granular system near the onset of jamming [218,31]. The system contains of small hexagonal crystal (hex) grains arranged randomly within the disordered bulk liquid. To measure the effect of hex structure on dynamics, we can compare the rate of structural decorrelation in the hex and non-hex regions. To do so, we require two correlation functions: first, a spatial correlation function to identify the hex regions and, second, a temporal correlation function to quantify how closely a given structure "remembers" its initial configuration as a function of time.

Identifying the hex grains within the bulk liquid requires a structural citerion that differentiates between particles in the liquid and hex regions on a per-particle basis. As mentioned previously, the $\mathbf{q}_6 \cdot \mathbf{q}_6$ scheme of reference [57] can be used to monitor nucleation and growth in 3d systems that form fcc, bcc, and hcp crystals [62, 386, 21, 367, 293, 292]. Although the scheme was originally based on the $\ell = 6$ Fourier coefficient $\mathbf{q}_6$, other shape descriptors can be just as easily be substituted [29, 30]. The main physical insight underlying the $\mathbf{q}_6 \cdot \mathbf{q}_6$ scheme is that crystals often contain local particle configurations that match with their neighbors in terms of *both* shape and orientation. Therefore, crystal-like particles can be identified by detecting those that match well with their neighbors. In analogy with reference [57], a good indicator of local crystal-like local ordering is the fraction of solid-like matches with neighbors:

$$f_{solid} = 1/n \sum_{j}^{n} \Theta[M_{dot}(\mathbf{S}_i, \mathbf{S}_j) - M_{cut}]. \tag{7.60}$$

Here Θ is the Heaviside function and "j" is a neighbor of "i," and $\mathbf{S}$ is a rotation-*dependent* shape descriptor. Particles with a minimal fraction of solid-like matches f_{cut} are considered to be locally crystalline. The cutoffs can be determined by viewing plots of the $P(M_{dot}(\mathbf{S}_i, \mathbf{S}_j))$ and $P(f_{solid})$ distributions for the bulk liquid and bulk solid [57, 62], or simply by visual inspection, as we perform here. This scheme holds for crystals in general, provided the neighbor shells all have the same shape. In reference [38], we describe how this scheme can be modified to handle crystals with an assortment of neighbor shells. The algorithm for detecting crystal grains is given in pseudocode below:

```
for i in list_of_particles
   set S[i] = compute_neighbors(i, rcut)
   call compute_rotation_dependent_shape_descriptor(S[i])
end for

for i in list_of_particles
   set n_solid = 0
   set count = 0
   for j within rcut of i
       if M(S[i], S[j]) > M_cut
```

```
            n_solid = n_solid + 1
          end if
          count = count + 1
        end for
        if n_solid / count > f_cut
           append(list_of_solid_particles, i)
        end if
    end for
```

For our example, we identify hex grains using the $\mathbf{S}^{\mathrm{RMS}}$ descriptor with $M_{cut} = 0.99$ and $f_{solid} = 0.5$. Notice that this scheme does not simply detect local hexagons; rather it identifies local hex regions while making the important distinction between isolated hexagons and the intermediate-range hex clusters of interest. Our temporal correlation function is defined by matching the point-matching descriptor for each cluster at time t with itself at a reference time t_0: $M_{dot}[\mathbf{S}^{\mathrm{RMS}}(t), \mathbf{S}^{\mathrm{RMS}}(t_0)]$. Fig. 7.33b shows the correlation function for hexagonal and non-hexagonal particles. We see that the hexagonal particles retain their structure longer on average than non-hexagonal particles. Similar correlation functions have applied to study glassy dynamics [138] and to study how the structure of liquid clusters change as they attach to a growing quasicrystal [30].

The general ideas underlying the correlation functions outlined above can be applied to more abstract problems as well. For example, rather than creating correlation functions in space and time, we can create correlation functions in parameter space. Consider the problem of automatically generating the structural phase diagram for the 2D Lennard-Jones-Gauss (LJG) system [120] shown in Fig. 7.34a from a collection of simulation snapshots for each statepoint. The phase diagram was generated by visual inspection of over 5000 statepoints [120]. We can automate the creation of this phase diagram by using an idea similar to the $\mathbf{q}_6 \cdot \mathbf{q}_6$ scheme outlined above. In this case, rather than finding structural correlations between neighboring particles in real-space, we can calculate correlations between neighboring snapshots in parameter space. For each point on the structural phase diagram, we compute $I(i) = \sum_j M(\mathbf{S}_i, \mathbf{S}_j)$, where "$i$" and "$j$" are neighbors in parameter space, and $\mathbf{S}$ is a global shape descriptor. Here, we take the global descriptor as a combination of a global superposition descriptor, indexed by a shape histogram with $n_\theta = 20$, $n_r = 1$, and a global probability distributions descriptor based on local Fourier descriptors with frequency range $\ell = 6, 7, ...11$: $\mathbf{S}^{\mathrm{global}} = \left\langle \mathbf{S}^{\mathrm{H2}}_{global}, P(\mathbf{S}^{\mathrm{F2}}_{local}) \right\rangle$. Points in stable regions of parameter space match well with neighboring points and have a high value of $I(i)$, whereas transitional points match poorly and have a low value of $I(i)$. The algorithm for creating a visual map of the structural phase boundaries for a parameter space is given in pseudocode below:

```
for each point i in parameter space
   call compute_shape_descriptor(S[i])
end for

for each point i in parameter space
```

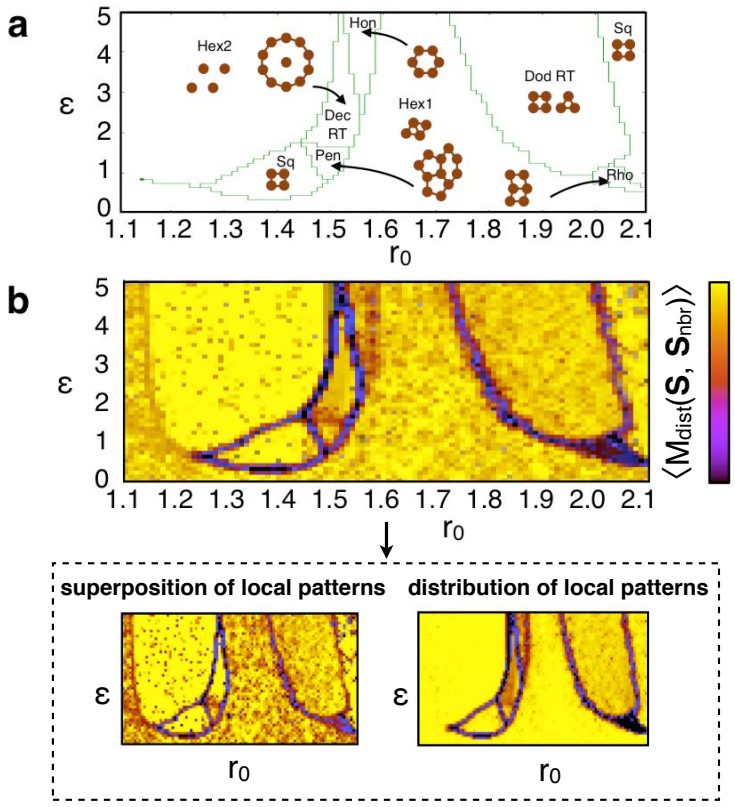


Figure 7.34 Structural phase diagram for the 2d Lennard-Jones Gauss system [120]. (*a*) Structural phase diagram created by visual inspection [120]. (*b*) Structural phase diagram created by shape matching. Each pixel in parameter space is given an intensity based on the average match with structures for neighboring points.

```
    set pixel_intensity[i] = 0
    for each neighboring point j
        set pixel_intensity[i] += M(S[i], S[j])
    end for
end for
```

The transitional points map out structural phase boundaries that look very similar to the diagram created by visual inspection. Notice that the global superposition descriptor detects no difference between the hexagonal and honeycomb crystals, since both structures are six-fold symmetric, and thus yield equivalent combined shape histograms. Additionally, the superposition descriptor picks up a slight artificial "boundary" within the hexagonal region near $r_0 \sim 1.15$, where the hexagonal crystal begins to form multiple grains rather than a single crystal, resulting in different shape histograms. The probability distributions descriptor, on the other hand, gives no distinction between the pentagonal and decagonal phases, since they are nearly identical locally, differing only in long range ordering. Overall, both sets of global information taken in combination are necessary to correctly find all of

the phase boundaries for this particular application. However, in many cases, the space of structures is sufficiently non-degenenate to be described by a single method.

Heat Maps and Grouping

Another common application of shape matching techniques is to the problem of visually grouping or classifying similar structures [360]. Grouping objects based on shape similarity has also been applied recently to macromolecules and proteins [310, 344]. Grouping can be accomplished by plotting the matrix of pairwise matching values known as a "similarity matrix" or "heat map." The algorithm for computing a similarity matrix is given in pseudocode below:

```
for i in list_of_local_structures
   call compute_shape_descriptor(S_i)
   for j in list_of_local_structures
       call compute_shape_descriptor(S_j)
       set matrix[i,j] =  M(S_i, S_j)
end for
return matrix
```

Objects can be grouped or classified based on features of the plot. As an example, consider the TIP4P water clusters [378] shown in Fig. 7.35. The matrix shows the match values obtained for minimum energy clusters of sizes $N = 2-21$, which are available on the Cambridge Cluster Database [379]. The clusters are indexed using rotation-invariant Zernike descriptors with frequency parameters $\ell = 4, 5, ..12$.

The patterns displayed in the heat map require some interpretation. The high correlation along the diagonal $i = j$ is common to all heat maps, and simply indicates that structures match perfectly with themselves. The region $N = 12-15$ displays a bright box, which indicates a group of structures that all match well with one another. The region $N = 7-10$ displays a checkerboard pattern, which indicates that every-other cluster matches well. Cluster $N = 16$, due to its unique non-compact nature, matches poorly with all other clusters, as indicated by the dark (purple) cross at $N = 16$. In addition to grouping and classifying objects, heat maps can be used to visually indicate convergence with respect to a changing parameter. Multiple heat maps based on different descriptors can be constructed for the same set of structures to show the similarities on different levels of ordering. Although we consider local clusters for our example, heat maps can be applied to any structures that can be indexed by shape descriptors, including global structures.

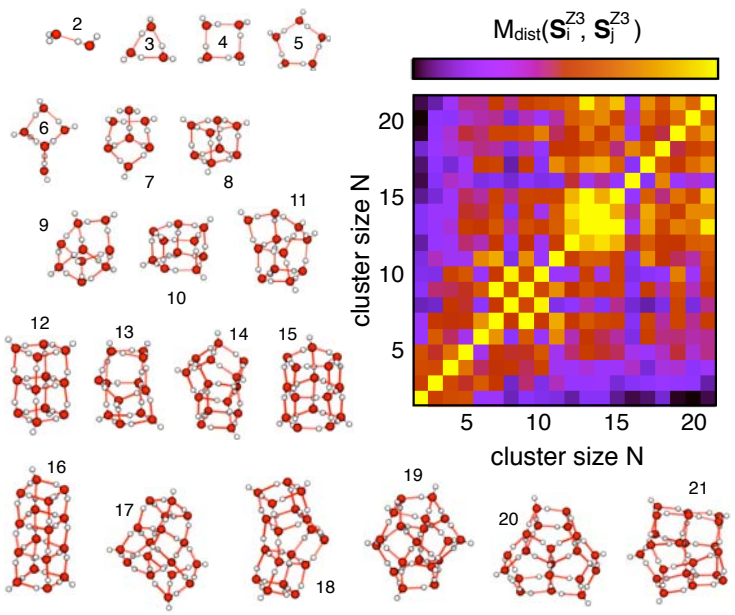


Figure 7.35 Grouping and classifying structures based on shape similarity. The plot shows a similarity matrix for energy-minimized TIP4P water clusters [378]. The matrix simultaneously shows the pairwise matching values for all clusters.

7.3.6 Future Outlook

In summary, we have introduced a shape matching framework for creating new structural metrics for complex patterns, such as those encountered in self-assembly. All of the methods and examples outlined here are accessible online through our C/C++ shape matching library [302]. Although our examples and discussions here are geared towards self-assembly and condensed matter physics, the general ideas underlying the shape matching framework are widely applicable to systems with complex structures, such as those encountered in computational biology. In the future, new shape descriptors and algorithms can be added to the framework to expand its scope to different classes of structures and problems.

The example applications and shape descriptors that we have presented here represent only a small subset of the vast range of possibilities yet to be explored. One obvious area for future study is to test the applicability of the wealth of shape descriptors from the shape matching literature to particle systems. Another promising area to explore is the creation of new abstract order parameters and correlation functions, such as the phase space correlation function of Fig. 7.34. This type of application may represent one of the most important uses for shape matching moving forward; replacing the human element with a computer algorithm to explore parameter space has the potential to greatly expedite self-assembly research.

Chapter 8
Future Outlook

In this thesis, we have explored several unique aspects of the formation of specialized ordered structures, the dynamics of supercooled liquids, and tools for characterizing structural transitions. Our findings have given rise to many new questions that should be addressed through future study. In this chapter, we address the ongoing and future work that relates to the topics presented in the previous chapters.

8.0.7 Generality of Quasicrystal Growth Mechanism

In section 5.1, we investigated the growth mechanism of a model dodecagonal quasicrystal. For this system, we found that the aperiodic growth of the quasicrystal was driven by the ability of the growing quasicrystal nucleus to assimilate kinetically trapped particles into the solid phase with minimal rearrangement, causing the system to form the the quasicyrstal rather than the stable approximant crystal (Fig. 8.1). A similar, coarse-grained version of this mechanism based on tiling models was proposed years earlier in reference [387]. The generality of this mechanism should be investigated for other quasicrystal-forming systems.

One system for which the growth mechanism can potentially be tested is the Lennard-Jones Gauss (LJG) system of reference [120], which forms a 2d decagonal quasicrystal. In this case, the system forms locally stable decagons in the liquid phase (analogous to icosahedra in the Dzugutov system) that may be incorporated into the quasicrystal during the growth phase. The system also has the advantage of being two-dimensional, which could potentially aid the direct visualization of the growth mechanism, and also simplify the mathematical analysis of the underlying particle structure. In addition to the LJG system, the growth mechanism can be tested for the binary LJ system of reference [163], which forms a 3d binary decagonal quasicrystal. In this case, the system has two different particle types, and quasicrystal growth may be dictated by the local concentration of different particle species. Because there are two particle types, the space of locally stable motifs may be more complex, which may also provide an interesting avenue for future investigation.

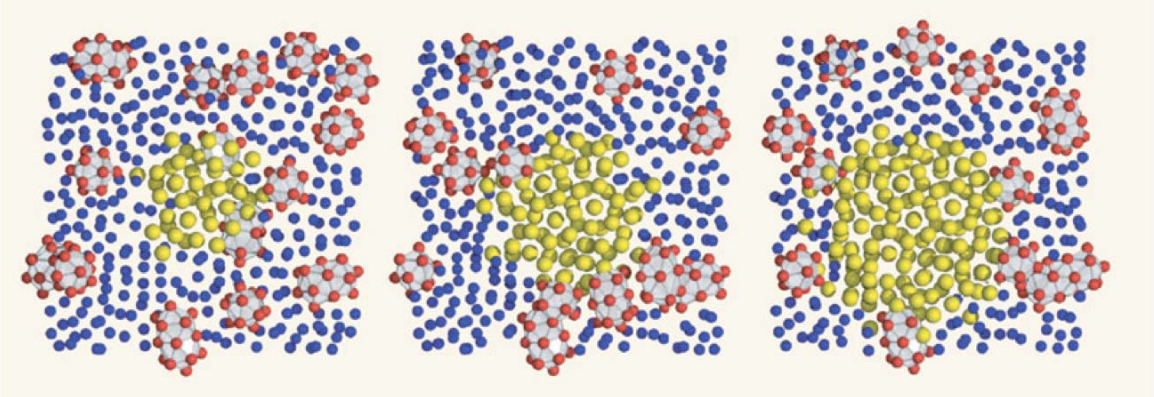


Figure 8.1 Cross-section of growing quasicrystal in the Dzugutov system depicting the growth mechanism proposed in Chapter 5. The image is reproduced from reference [175], and was originally created by Aaron Keys and Chris Iacovella.

8.0.8 Nucleation in Quasicrystals

During the course of our study of quasicrystal growth the Dzugutov system, we learned two interesting facts about the solid "nucleus" that gives rise to the bulk quasicrystal. Namely, we learned that (1) the nucleus has icosahedral local structure and (2), the nucleus is surrounded by a wetting layer of icosahedra. This implies that a potentially interesting avenue for future study involves investigating the effect of local icosahedral ordering on quasicrystal nucleation.

The first question that should be answered is whether or not the formation of the critical nucleus is correlated with icosahedral clusters in the system. Our preliminary results (obtained via umbrella sampling) give strong evidence that this is the case. Fig. 8.2 shows the typical arrangement of icosahedral clusters relative to nuclei of two different structures: a dodecagonal quasicrystal and an fcc crystal. We observe that both nuclei are correlated with the icosahedral clusters but in an opposite way; whereas the fcc nucleus (yellow) tends to arise in a region devoid of icosahedra (purple), the quasicrystal nucleus (red) arises inside an icosahedral cluster (grey). These results make intuitive sense from a packing standpoint; since dodecagonal quasicrystals are polytetrahedral, the quasicrystal nucleus may tend to arise within a region with stronger icosahedral ordering than the bulk liquid. On the other hand, the fcc nucleus, which is not polytetrahedral, may tend to form in the non-icosahedral bulk where the rearrangement of low-energy particles is not required.

A second interesting question to address is how ordered nuclei with icosahedral local ordering (such as the quasicrystal depicted in Figure 8.2a) initially arise within the system. While it is possible that the quasicrystal nucleus forms within an existing icosahedral cluster,

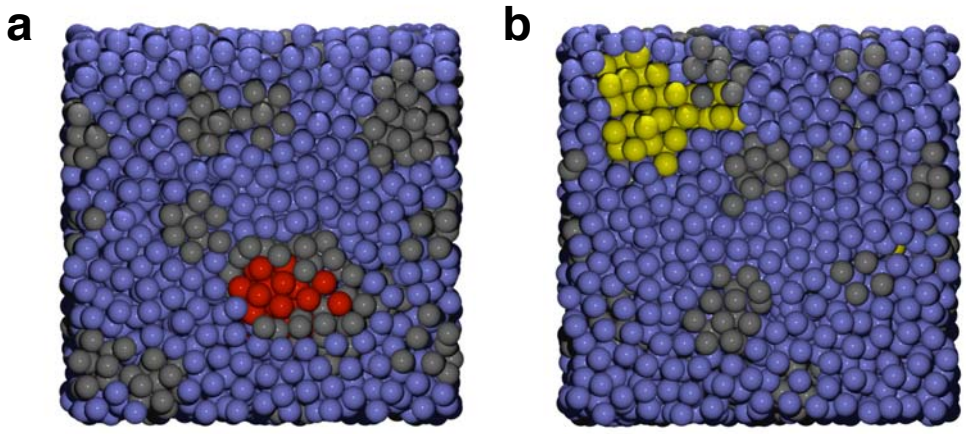


Figure 8.2 Icosahedral clusters and nucleation. (a) A microstate from the Dzugutov system at $T = 0.55$, $P = 3.50$ containing a quasicrystal nucleus (red) of size $N \sim 50$. (b) A microstate from the Dzugutov system at $T = 0.55$, $P = 3.50$ containing an FCC nucleus (yellow) of size $N \sim 100$. Icosahedral clusters are shown in grey. The quasicrystal nucleus preferentially forms inside an icosahedral cluster whereas the FCC nucleus preferentially forms in the non-icosahedral bulk liquid. We use umbrella sampling to sample both nucleus structures at the same state point.

it is also possible that the nucleus arises in the bulk liquid like an ordinary icosahedral cluster and subsequently forms icosahedra about its surface. If the former scenario is the case (i.e., the icosahedral cluster comes first), it is interesting to consider whether the cluster has unique properties that make it a favorable "host" for quasicrystal nuclei. Answering this question would require a method for sampling rare events, such as transition path sampling (TPS) (see section 4.4). Unlike umbrella sampling used to obtain our preliminary results, TPS yields MD trajectories that are sequential in time. Thus, we can extend reactive trajectories backwards in time to determine the origin of the quasicrystal nucleus directly.

8.0.9 Phase Diagrams for Model Micelles

In our study of soft matter quasicrystals in Chapter 5, we introduced a minimal micelle model to study the effect of surface particle mobility on quasicrystal stability. Specifically, our model consists of small spheres tethered to a rigid spherical scaffolding (Fig. 8.4a). In some cases, we also introduced "dimerized" micelles by bonding the centers of two spheres together. Using this model, we were able to capture the important physical characteristics of two different micelle-forming systems based on tethered nanosphere building blocks. We showed that, for a particular set of parameters, the unique mobile surface particle entropic mechanism exhibited by these systems can stabilize Frank-Kasper (FK) phases, such as

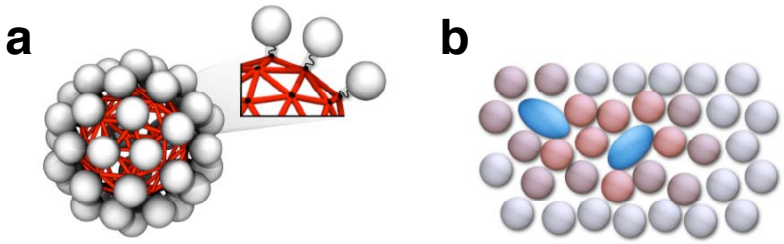


Figure 8.3 Model systems for future assembly studies of quasicrystal assembly. (a) Model micelle with mobile surface atoms. (b) Mixture of spherical and aspherical particles.

dodecagonal quasicrystals as well as more common structures such as bcc crystals.

In the future, the phase behavior of these model micelles can be studied in more detail. Since the micelle model has numerous variable parameters, it presents a vast parameter space. Parameters for the model include the size, shape, and interactions of the attached mobile particles, the type and stiffness of the spring, and the shape of the rigid scaffolding that provides the attachment points. In some cases, polymers might be tethered to the attachment points rather than particles. It is possible that the entropic mechanism of surface area minimization might help to stabilize unique structures other than quasicrystals for different sets of the parameters. Once desirable target structures are obtained, explicit micellar systems can then be designed based on simpler nanoparticle building blocks that assemble micelles resembling the parameterized model. In addition to micellar systems, nanosatellite systems may also be designed in some cases, such as the systems of references [201] and [388].

8.0.10 Sphere-Dimer Binary Mixtures

In our study of soft matter quasicrystals in Chapter 5, we introduced model binary mixtures of pill-shaped dimers and spheres to study the effect of particle shape polydispersity on assembled crystal structures (Fig. 8.4b). For certain parameter ranges, we found that FK phases exhibit lower free energies than standard fcc, hcp, and bcc crystals for the binary mixtures. Since other untested crystal structures may be more stable, our results do not prove that the FK phases minimize the global free energy; rather, our results prove that particle shape polydispersity lends relative stability to FK phases over standard crystals.

Of course, our results still strongly imply that it is possible for FK structures to minimize the global free energy under particular conditions. Oblong particles can already be fabricated in a variety of colloidal systems [389, 390, 391]. Thus any binary mixtures with standard colloidal interaction potentials that assemble FK structures in simulations would

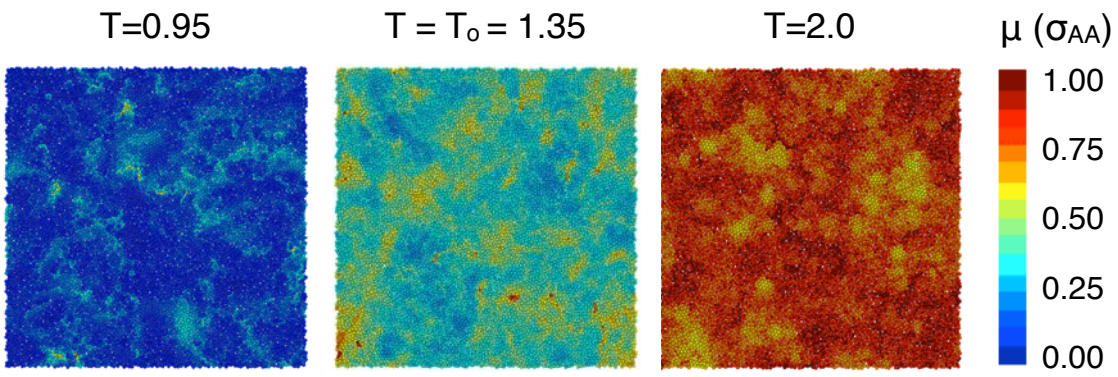


Figure 8.4 Excitations and particle structure. The panel shows isoconfigurationally-averaged mobility fields [137] for a 2d binary WCA supercooled liquid at different temperatures. The averaging window is given by a typical timescale of an excitation, as defined in Chapter 6. The presence of maxima throughout the averaged mobility field (i.e., the different colors) indicates that the dynamics are correlated to particle structure.

be potentially viable experimentally. Molecular simulations can be employed to perform a preliminary investigation of the parameter space over which systems might form FK crystals and quasicrystals, as well as other interesting crystal structures. Our results from section 5.2 indicate that it may be difficult to assemble ordered phases for sphere-dimer mixtures for relatively hard interactions, such as the WCA potential. For example, self-assembly runs using parameters in the "stable" range for the FK structures shown in Fig. 5.9 typically result in disordered non-ergodic phases (i.e., glasses). This indicates that either nucleation is suppressed due to glassy dynamics (see section 2.3.3) or the free energy minimizing structure is not ordered. The former scenario may be circumvented by choosing a potential with softer interactions, thus allowing particles to better escape from local minima. One standard colloidal soft interaction is given by the Yukawa potential for systems with screened electrostatic interactions [392].

8.0.11 Elementary Excitations and Structural Weaknesses

In our study of fundamental particle motions in Chapter 6, we observed strong parallels between particle dynamics in atomistic supercooled liquids and kinetically constrained models (KCMs) (see section 2.4.2). The theory of KCMs is partially based on the underlying idea of dynamical facilitation (DF); that is, the motion of a particle begets or facilitates the subsequent motion of another particle at a later time. Such facilitated motion may have an associated waiting time; that is, the motion of one particle may lower the activation barrier

for a local rare particle motion event, making it more likely to occur, but not guaranteeing so.

In this sense, the motion of particles may be loosely tied to the underlying structure. Structural weaknesses that arise due to nearby particle motion allow subsequent particles to move. This idea can be tested by applying a standard path sampling framework to activated events characterized by particle motion, similar to that applied in section 6.2. Structural transitions in local particle configurations can then be characterized using the standard methodology used to study rare events. For example, we can measure the free energy for a local particle jump, and identify the transition state and reaction coordinate(s) characterizing local particle motion. Doing so would provide an important step forward in understanding the glass transition, since it would, for the first time, draw a concrete connection between local particle structures and the dynamics of supercooled liquid dynamics.

8.0.12 Structural Correlations with String-Like Motion

A more specific subset of the broader question posed in the previous section, is: what is the atomistic mechanism underlying strings of mobile particles? That is, what is it about the local structure of particles that causes them to follow one another in quasi-one dimensional strings? In 3d systems, such as the Dzugutov system, strings tend to occur at the surface of icosahedral clusters [70]. This is surprising, since one might naively expect particle motion to occur as far from the locally-ordered regions as possible. Rather, this result implies that a tendency for particles to attach to the stable icosahedral clusters locally may provide a driving force for string-like motion. Such attachment would increase the frustration of the cluster, causing, on average, a subsequent detachment of particles elsewhere, such that the cluster size distribution remains in equilibrium.

This idea could be tested by considering a simple model, such as a 2d liquid. Imagine a tagged particle in such a system situated next to a relatively large ordered (say hexagonal) cluster. We can ask ourselves the following questions: (1) how are the forces on this particle different than those in a crystallizing system? (2) How will the resulting motion of our tagged particle differ from a particle in a crystallizing system? Certainly in a global sense, a glass-forming system and a crystallizing system differ strongly; the glass-forming system can only form ordered grains of a certain small size before packing frustration sets in and breaks the structure apart. However, locally, the forces and motions of particles may not significantly differ between particles in a glass or in a nucleating crystal. Thus, particle motion may occur via the same basic mechanisms. A related idea has been put forth by Granato and coworkers, based around interstitial defects [393, 394].

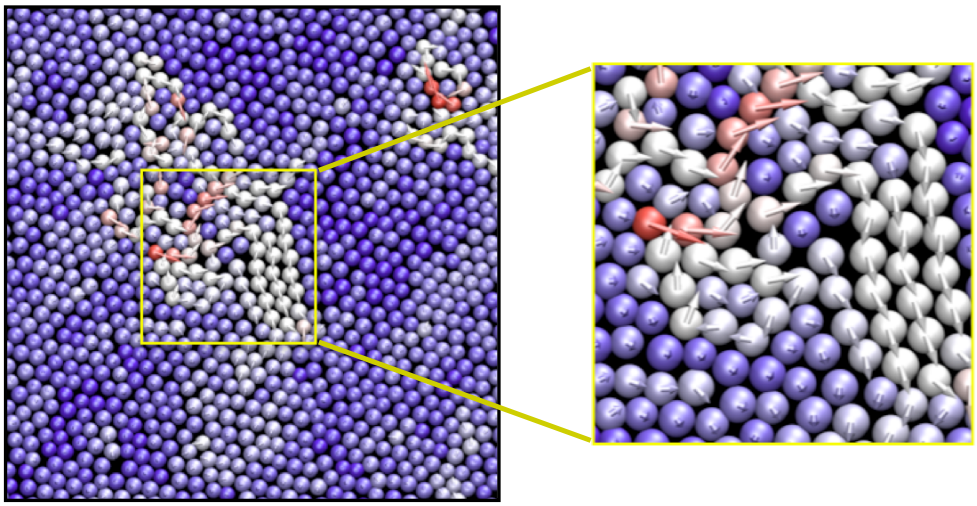


Figure 8.5 String-like motion in crystallization. The panel depicts a 2d crystalline system with a structural defect, imposed by a small degree of particle size polydispersity. The displacement vectors depict how the particles move to anneal the defect, causing another defect to form.

One potentially simple way to test our idea is to observe the particle dynamics in a system with a growing crystal. Another potentially simple method is to observe how a crystalline structure fixes local structural defects. Fig. 8.5 shows particle dynamics within a hexagonal crystal grain with a small defect that occurs due to the presence of a small amount of particle size polydispersity in the system. We observe that the particles move as strings to fix the defect locally. The defect moves throughout the system when the defective regions crystallize, causing nearby regions to become defective. This very preliminary result is sufficiently interesting, that in the future, a more thorough study of the relationship between string-like motion and crystallization can potentially be conducted.

8.0.13 Shape Matching For Experimental Systems

In our study of the shape matching methods for characterizing transitions in assembled systems in Chapter 7, we introduced several general metrics that can be applied to particle structures in general. As described in sections 7.2 and 7.3, this includes experimental images, such those obtained from TEM or confocal microscopy. Thus, many of the same shape matching algorithms to characterize experimental data as well, with the caveat that in some systems, we often only obtain 2d snapshots of a 3d system.

In section 7.2.12, we demonstrated that shape matching can be used to identify unknown structures based on experimental images, which is useful, for example, in the context of

data mining algorithms (see, Fig. 7.18). Another application for which shape matching techniques may prove particularly useful is in the case of quantifying structural perfection in experimental systems. Images of particles can be compared with ideal mathematically constructed reference images to give a quantitative degree of perfection, or a pseudo-order parameter.

8.0.14 Targeted Parameter Space Searches Using Shape Matching

In a very simplified sense, there are two generic problems in the field of computational self-assembly: (1) given a desirable target structure, what is a suitable range of parameters for which the structure can be obtained, and (2) given a system consisting of a building block and set of feasible parameters, what is the space of unique assembled structures that we can obtain? Both problems are typically solved by a combination of physical intuition and feedback from simulation results. For example, for a problem of the form (1), we might guess a set of parameters that will yield the desired target structure and adjust these parameters based on molecular simulation results. For a problem of the form (2), we might run simulations for a range of parameters and then use post-processing to construct a structural phase diagram. In many cases, we attempt to optimize our study to consider the largest space possible with the fewest simulations.

Both types of problems mentioned above can potentially be automated and optimized by using shape matching methods described in Chapter 7. For example, the problem of identifying parameter sets that produce a particular target structure can potentially be solved by using shape matching techniques within the context of standard optimization algorithms, such as genetic algorithms [375]. In this case, the measure of "fitness" might be proportional to the match between an ideal structure, and that obtained for a trial parameter set. Shape matching schemes may also be applied to the problem of finding unique structures for a parameter space. For example, in section 7.3.5, we automatically defined structural phase boundaries for a system sampled on a grid in parameter space. An algorithm based on this type of matching scheme could be used to, for example, iterate along a phase boundary to efficiently map the parameter space without simulating the regions within the boundaries, thus saving significant computational effort.

Chapter 9
Conclusions

In summary, we have investigated solidification transitions in several different model and experimental liquids, with a varying degrees of structural complexity. To supplement our studies, as well as other future studies in the field of colloidal and nanoscale assembly, we have expanded the breadth of the mathematical toolset that can be used to characterize particle structures and structural transitions. Our most important findings can be summarized as follows:

1. Particle dynamics in dense granular materials approaching a jamming transition are indistinguishable from particle dynamics in supercooled liquids approaching a glass transition.

2. The theory developed based on coarse-grained kinetically constrained models (KCMs) can be accurately applied to describe the particle dynamics and bulk relaxation properties of fully-atomistic model glass formers.

3. The detailed particle-level mechanism by which quasicrystals form involves the incorporation of locally-stable motifs from the liquid-solid interface into the growing solid nucleus with minimal structural rearrangement.

4. Quasicrystals can be consistently assembled from systems of micelles that exhibit both shape polydispersity and mobile surface entities.

5. Computer science shape matching algorithms, commonly used for electronic database searches, can be applied to create highly flexible order parameters for complex assembled structures.

Our findings are modest within the grand scope of liquid-solid transitions; however, in a more focused context, they provide some important incremental steps forward.

Our study regarding the connection between the glass transition and "jamming" in granular materials is one of several studies on this topic that have demonstrated the universality

of glassy phenomena (see also references [66, 65]). Because the granular system that we investigate consists of hard particles, our results highlight the important role played by excluded volume interactions in producing glassy phenomena, as predicted by Weeks, Chandler and Andersen over 30 years prior [118]. Our study regarding the physical connection between KCMs and fully-atomistic glass formers provides an important step forward in our understanding of the nature of the glass transition. While KCMs are known to exhibit the hallmarks of glassy dynamics in coarse-grained systems, the connection between these models and fully atomistic systems has, until recently, been ambiguous. Our work, along with other recent studies relating coarse-grained models to more realistic systems [111, 109, 107], has helped elucidate this connection. This highlights the important role of the minimal physics built into coarse-grained glass formers in producing glassy dynamics in realistic systems. The picture of glass forming liquids that arises from our study allows us to explain previously unexplained behaviors, such as the origin of the growing dynamical length-scale for strings of mobile particles, and the exponential probability distribution of string lengths [73, 71, 31].

Our study regarding quasicrystal growth has a similar flavor; in this case, several coarse-grained models had been proposed to describe quasicrystal growth [175, 387], but our study was the first to investigate quasicrystal growth within a particle-based simulation. Our results demonstrate the applicability of a particular growth model [175] proposed in reference [387], and also provide important insight into how quasicrystals might form via self-assembly. In particular, quasicrystals should be formed in rapidly quenched systems, in which there is a sizable barrier to rearranging stable structures at the liquid-solid interface. This is in agreement with the well-known experimental observation that quasicrystals typically form in rapidly quenched melts and are often thermodynamically metastable with respect to a crystalline stable state [189]. This study has already aided us in investigating the thermodynamic mechanism by which quasicrystals arise in a system of dendrimers [191], and understanding how this type of mechanism can be widely exploited to assemble quasicrystals in micro and nanoscale systems.

Our study of shape matching techniques introduces a powerful new paradigm for characterizing complex particle structures. The framework that we introduce describes highly-general algorithms that can be used with arbitrary shape descriptors to characterize virtually any type of structure. The harmonic shape descriptors that we introduce can be applied to a sizable subset of the space of structures that are currently known to occur in nanoscale and colloidal systems. Shape matching techniques have already been applied to several studies within the Glotzer group (see references [29, 30, 187, 33]), and, in the future, may potentially be applied broadly within the computational and experimental self-assembly

communities.

Perhaps more importantly than the direct implications of our findings, our studies provide many interesting avenues for future investigation. Our study of glassy dynamics in granular matter has already given rise to several follow-up studies, based on exploring the generality of our findings for different types of granular and colloidal systems, and systems with different types of applied driving forces. Our findings regarding particle dynamics in glassy liquids will not only aid future studies concerned with relating atomistic and molecular liquids to KCMs, but will also aid future theoretical studies by highlighting important aspects of glassy liquids that theoretical descriptions of the glass transition must account for. Our collective findings regarding quasicrystal formation imply that a wide range of systems might form quasicrystals under the proper conditions, including micellar systems composed of nanoparticle or block-copolymer building blocks, nano-satellite systems, or simple binary colloidal mixtures with particle shape polydispersity. These systems will be studied in more detail using computer simulations in the immediate future, and may also be realized experimentally in the not-too-distant future. Our study concerning shape matching methods will aid both computational and experimental studies of assembled systems by providing a robust numerical mechanism for characterizing complex structures. In the longer-term, these types of methods might be applied to unique computational algorithms, such as data mining applications, or genetic algorithms to automate searches of parameter space for desired target structures.

In summary, our findings provide some general insights into liquids-solid transitions, which, we predict, will be most applicable to the problem of self-assembly on the nano and colloidal scale, where a detailed understanding of these mechanisms is particularly important. In a sense, our work represents only a tiny subset of the huge collection of studies aimed at bridging the gap between standard materials and devices of yesterday and the new tiny materials and devices of tomorrow. The future of nanotechnology is bright, and in the coming decades we will witness huge new innovations on a tiny scale.

Appendices

Appendix A
Additional Studies

The work presented in this thesis is largely fundamental in nature, dealing primarily with the physics of liquid-solid phase transitions in a general sense. Thus, for our work to have lasting impact, our ideas must be applied by other scientists to facilitate their own research. In some specific cases, we have done so directly. The results of these studies, which fall into the field of self-assembly, are briefly outlined below.

A.1 Self-assembly of patchy particles into diamond structures through molecular mimicry

The diamond structure is desirable for photonic applications, because it exhibits a complete 3d photonic bandgap [395]. In [Z. Zhang, A.S. Keys, T. Chen, and S.C. Glotzer, *Langmuir* (2005)] [29], we demonstrate that the diamond structure can be self-assembled from a model system of patchy particles (Fig. A.1). We also show that the nucleation rate can be greatly enhanced by seeding the system with a static diamond-structured nucleus. Additionally, we show that adding a directional interaction, we reduce defects in the final structure.

In addition to demonstrating the assembly of a unique and desirable structure, this study represents the first (implicit) use of our shape matching methods, which are applied to create a specialized order parameter to detect diamond crystalline ordering. This application is discussed in further detail in reference [38].

A.2 Icosahedral packing of polymer-tethered nanospheres and stabilization of the gyroid phase

The double gyroid phase is a unique bi-continuous network structure that is known to occur in systems of block copolymers [307, 210]. When functionalized with nanoparticles, this

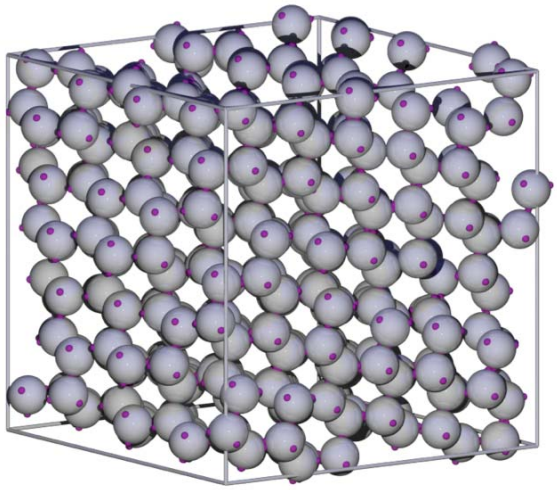


Figure A.1 Depiction of the diamond structure assembled from patchy particles studied in reference [29].

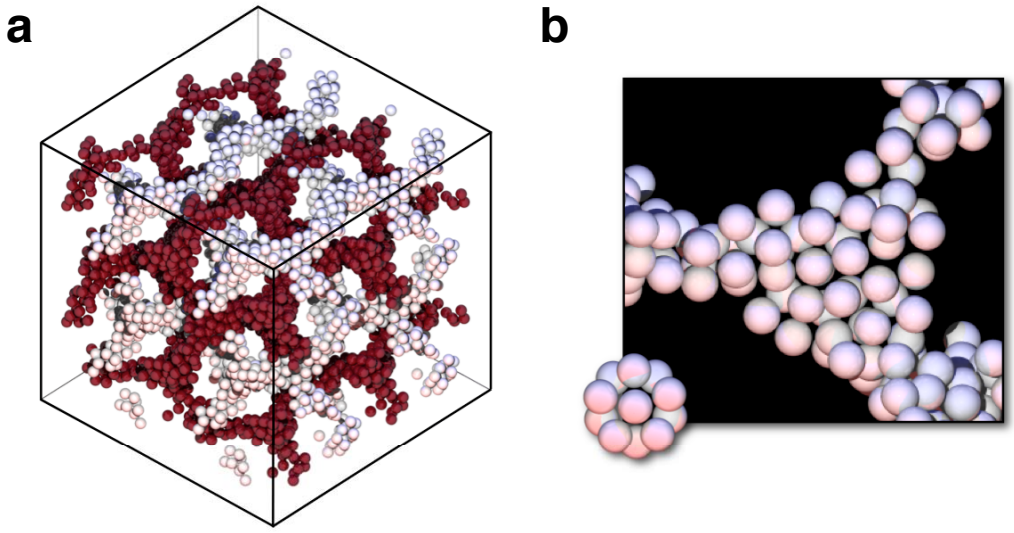


Figure A.2 Double gyroid structure assembled from tethered nanospheres [187]. (a) The double gyroid global structure. The two domains are colored differently for ease of viewing (b) Nanoparticles at the gyroid node, exhibiting local icosahedral ordering.

type of structure may have important applications in photonics [396] and drug delivery [397]. In [C.R. Iacovella, A.S. Keys, M.A. Horsch, and S.C. Glotzer, *Physical Review E* (2007)], we demonstrate that double gyroid phases can be assembled from systems of tethered nanosphere building blocks (Fig. A.2).

In addition to reporting the formation of the double gyroid, we demonstrate that stability of the gyroid can be rationalized by studying the local packing of nanoparticles within the

structure. We find that nanoparticles exhibit local polytetrahedral ordering, much like a liquid of spherical particles (see Chapter 3). However, since the nanoparticles are packed into confined domains between the tethers, the icosahedral packing is not frustrated, and thus stabilizes the system. This highlights the important difference between packing particles in confined domains and in the bulk. In addition to these findings our study represents the first instance of applying an explicit shape matching scheme to detecting particle structures, which was applied for characterizing the local ordering of nanoparticles.

A.3 Materials science: A tale of two tilings

Only in the past decade have quasicrystals been assembled on the colloidal and nanoscale. In [S.C. Glotzer and A.S. Keys, *Nature* (2008)] [32], we highlight the recent discovery by Mikhael *et al.* [124] of decagonal quasicrystals in a 2d systems of colloidal particles with a laser-induced potential field exhibiting 10-fold symmetry. For strong fields, the system forms a decagonal quasicrystal, while for weak fields, the system forms an ordinary hexagonal crystal. Interestingly, for intermediate field strengths, the system forms a Fibonacci chain composed of of Archimedean (square and triangle) tiles that can be considered a unique type of 1d quasicrystal that is periodic in one dimension and quasiperiodic in the other (Fig: A.3). This finding not only highlights the interesting connection between quasicrystals and Archimedean tiles, but also hints at a unique connection between quasiperiodic structures with different dimensionalities.

A.4 Disordered, quasicrystalline and crystalline phases of densely packed tetrahedra

Until recently, all of known quasicrystals consist of spherical particles, or roughly spherical entities, such spherical micelles. In [A. Haji-Akbari *et al.*, *Nature* (2009)] [33], we report the formation of a unique dodecagonal quasicrystal from hard tetrahedral particles (Fig: A.4. We demonstrate that the quasicrystal and its approximants represent a dense packing paradigm for tetrahedra that, at the time, easily set the record for the densest known packing of tetrahedra. (This record has since been broken several times by a class of structures with a dimer packing motif).

In addition to reporting the formation of the quasicrystal, we investigate several other properties of the hard tetrahedron system, including a possible liquid-liquid transition at

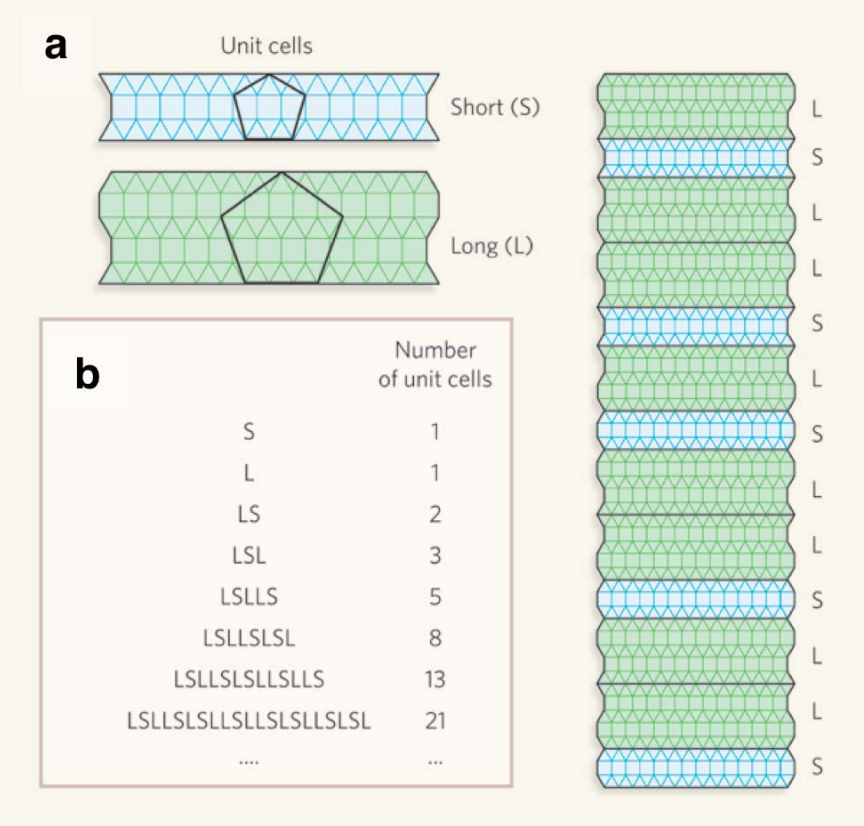


Figure A.3 Schematic of a Fibonacci chain observed in a system of colloidal particles [124]. (a) The two unit cells that are stacked to make up the chain (b) The Fibonacci sequence up to 21 elements.

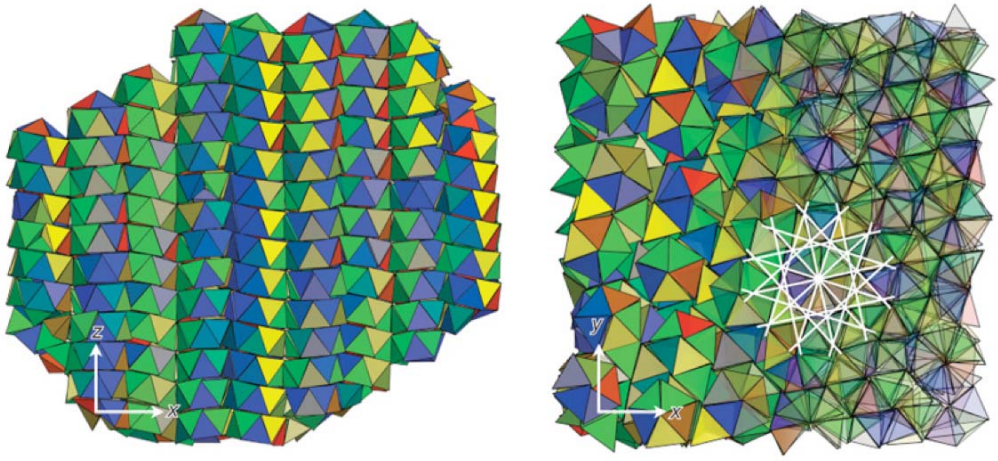


Figure A.4 Dodecagonal quasicrystal assembled from hard tetrahedra [33]. The quasicrystal is shown from two directions: (left) along the periodic z direction, (right) along the aperiodic 12-fold direction.

intermediate packing densities. To do so, we perform a detailed characterization of local motifs within the system using shape matching. This represents our first application of our particle shape matching framework to a system with highly non-spherical particles.

Appendix B
Simulations and Analysis Codes

During the course of this dissertation, we developed several codes for performing molecular simulations and data analysis. Although the source codes are too long to include in this document, we provide a brief overview of the codes and their basic purposes in this appendix. All of the codes listed below are currently available via the web.

B.1 LibTPS – Transition Path Sampling Library

LibTPS is an object-orientated library for performing transition path sampling and related methods for an arbitrary molecular simulation code. The primary goal of the project is to expose path sampling techniques to a wider audience for whom the time required to code a path sampling algorithm poses a significant barrier to using the methods. A secondary aim of the library is to provide a collaborative development environment for researchers working on path sampling techniques to test and disseminate new path sampling algorithms. Currently, the library implements transition path sampling (TPS) [60] and related methods, such as transition interface sampling (TIS) [398] and Aimless Shooting [43]. Additionally, the library provides several standard prototype systems with the aim of facilitating the comparison of results and testing new algorithms. The library currently includes built-in support for the molecular simulation API (see section B.6 below), LAMMPS [148], and HOOMD-blueLAMMPS [250]; however, the library can be used with any molecular simulation code written in either C/C++, Java or python. In the future, we also plan to provide interactive path-sampling demos via the web.

B.2 SMAC – Shape Matching Analysis Code

SMAC is a library of routines for computing the shape descriptors and performing the analyses described in Chapter 7. The library provides routines for computing Fourier de-

scriptors, Zernike descriptors, shape contexts, shape distributions, and shape histograms. Additionally, the library provides algorithms for several standard applications, such as local and global structural identification and crystal grain detection. The routines are packaged as a C++/python library that can be included within an existing code.

B.3 LibSHD – Spatially Heterogeous Dynamics Library

LibSHD is a library of routines for characterizing spatially heterogeneous dynamics in supercooled liquids and related systems based on the analyses outlined in section 4.6. The library includes routines for computing standard dynamical functions such as the mean-squared displacement, the self-intermediate scattering function, and the the self-Van Hove correlation function. Additionally, the library includes routines for computing more specialized measures of heterogeneous dynamics, including the non-Gaussian parameter, clusters of mobile particles, strings of mobile particles, and various four-point correlation functions. The library also provides routines for creating movies of dynamical data from simulation trajectories.

B.4 LibSLAC – Library of Standard Liquid Analysis Codes

LibSLAC provides routines for computing standard condensed matter analyses (see section 4.5, including the radial distribution function, the isotropic structure factor, and the simulated diffraction pattern. Although these analyses are fairly simple to implement, we provide the library as a means to implement these standard functions quickly and correctly, or to validate the results for new analysis codes.

B.5 VMDStream – Communicate with VMD from C++ using a TCP connection

VMDStream is a simple, lightweight library to control VMD [373] remotely from a C++ code. This is particularly useful for generating custom renderings for data analysis, controlling movies, or making animations. All that is required is some knowledge of VMDs Tcl interface. The code consists of a single header file (only $\sim$ 300 lines of code) and can easily

be incorporated into any existing C/C++ code.

B.6 Additional Codes and Projects

In addition to the codes outlined above, we have been involved in the development of projects that benefit the molecular simulation community as a whole. We have developed several additions to the HOOMD-blue code, including time-reversible integration routines and binary restart read/write capabilities (both of which are necessary for transition path sampling), potential energy-minimization, and the ability to run simulations in 2d. We have contributed to the development of a molecular simulation application programmer interface (API), which provides a common interface for interacting with molecular simulation codes. Using this API, developers can write standardize diagnostics and algorithms based on an arbitrary underlying molecular simulation engine. Our transition path sampling library, LibTPS, mentioned above is one example of a library built around this API.

Bibliography

[1] O. D. Sherby and J. Wadsworth, "Ancient blacksmiths, the Iron Age, Damascus steels, and modern metallurgy," *Journal of Materials Processing Technology*, vol. 117, no. 3, pp. 347–353, 2001.

[2] J. Henderson, "Tradition and experiment in first millennium AD glass production: The emergence of early Islamic glass technology in late antiquity," *Accounts of Chemical Research*, vol. 35, no. 8, pp. 594–602, 2002.

[3] M. Kumar, "Nano and microparticles as controlled drug delivery devices," *Journal of Pharmacy and Pharmaceutical Sciences*, vol. 3, no. 2, pp. 234–258, 2000.

[4] G. F. Paciotti, L. Myer, D. Weinreich, D. Goia, N. Pavel, R. E. McLaughlin, and L. Tamarkin, "Colloidal gold: A novel nanoparticle vector for tumor directed drug delivery," *Drug Delivery*, vol. 11, no. 3, pp. 169–183, 2004.

[5] W. Lu and C. M. Lieber, "Nanoelectronics from the bottom up," *Nature Materials*, vol. 6, no. 11, pp. 841–850, 2007.

[6] R. Waser, H. Keller, and U. Erb, "Nanoelectronics and information technology: Advanced electronic materials and novel devices," *MRS Bulletin*, 2004.

[7] G. Subramanian, V. N. Manoharan, J. D. Thorne, and D. J. Pine, "Ordered macroporous materials by colloidal assembly: A possible route to photonic bandgap materials," *Advanced Materials*, vol. 11, no. 15, pp. 1261–1265, 1999.

[8] A. Stein and R. C. Schroden, "Colloidal crystal templating of three-dimensionally ordered macroporous solids: Materials for photonics and beyond," *Current Opinion in Solid State and Materials Science*, vol. 5, no. 6, pp. 553–564, 2001.

[9] G. M. Whitesides and B. Grzybowski, "Self-assembly at all scales," in *Macromolecular Symposia*, vol. 77, p. 283, 1994.

[10] S. C. Glotzer, "Some assembly required," *Science*, vol. 306, pp. 419–420, Oct 15 2004.

[11] S. C. Glotzer and M. J. Solomon, "Anisotropy of building blocks and their assembly into complex structures," *Nature Materials*, vol. 6, pp. 557–562, 2007.

[12] Z. Nie, A. Petukhova, and E. Kumacheva, "Properties and emerging applications of self-assembled structures made from inorganic nanoparticles," *Nature Nanotechnology*, vol. 5, no. 1, pp. 15–25, 2010.

[13] L. E. Euliss, J. A. DuPont, S. Gratton, and J. DeSimone, "Imparting size, shape, and composition control of materials for nanomedicine," *Chemical Society Reviews*, vol. 35, no. 11, pp. 1095–1104, 2006.

[14] D. Chandler, *Introduction to Modern Statistical Mechanics*. Oxford University Press, USA, 1987.

[15] E. D. Zanotto, "Do cathedral glasses flow?," *American Journal of Physics*, vol. 66, no. 5, pp. 392–395, 1998.

[16] W. Ostwald, "Studien über die Bildung und Umwandlung fester Körper," *Zeitschrift für Physikalische Chemie*, vol. 22, p. 289, 1897.

[17] M. Volmer and A. Weber, "Keimbildung in übersättigten Gebilden," *Zeitschrift für Physikalische Chemie*, vol. 119, pp. 277–301, 1926.

[18] R. Becker and W. Döring, "Kinetische Behandlung der Keimbildung in übersättigten Dämpfen," *Annalen der Physik*, vol. 416, no. 8, pp. 719–752, 1935.

[19] D. Moroni, P. R. ten Wolde, and P. G. Bolhuis, "Interplay between structure and size in a critical crystal nucleus," *Physical Review Letters*, vol. 94, p. 235703, Jun 2005.

[20] A. C. Pan and D. Chandler, "Dynamics of nucleation in the Ising model," *Journal of Physical Chemistry B*, vol. 108, pp. 19681–19686, 2004.

[21] A. Cacciuto, S. Auer, and D. Frenkel, "Onset of heterogeneous crystal nucleation in colloidal suspensions," *Nature*, vol. 428, no. 6981, pp. 404–406, 2004.

[22] A. Cacciuto and D. Frenkel, "Stresses inside critical nuclei," *Journal of Physical Chemistry B*, vol. 109, no. 14, pp. 6587–6594, 2005.

[23] C. A. Angell, K. L. Ngai, G. B. McKenna, P. F. McMillan, and S. W. Martin, "Relaxation in glassforming liquids and amorphous solids," *Journal of Applied Physics*, vol. 88, p. 3113, 2000.

[24] P. G. Debenedetti and F. H. Stillinger, "Supercooled liquids and the glass transition," *Nature*, vol. 410, no. 6825, pp. 259–67, 2001.

[25] P. G. Debenedetti, *Metastable Liquids: Concepts and Principles*. Princeton University Press, 1996.

[26] M. D. Ediger, C. A. Angell, and S. R. Nagel, "Supercooled liquids and glasses," *Journal of Physical Chemistry*, vol. 100, no. 31, pp. 13200–13212, 1996.

[27] M. E. Zoorob, M. D. B. Charlton, G. J. Parker, J. J. Baumberg, and M. C. Netti, "Complete photonic bandgaps in 12-fold symmetric quasicrystals," *Nature*, vol. 404, no. 6779, pp. 740–743, 2000.

[28] W. Man, M. Megens, P. J. Steinhardt, and P. M. Chaikin, "Experimental measurement of the photonic properties of icosahedral quasicrystals," *Nature*, vol. 436, no. 7053, pp. 993–996, 2005.

[29] Z. Zhang, A. S. Keys, T. Chen, and S. C. Glotzer, "Self-assembly of patchy particles into diamond structures through molecular mimicry," *Langmuir*, vol. 21, no. 25, pp. 11547–11551, 2005.

[30] A. S. Keys and S. C. Glotzer, "How do quasicrystals grow?," *Physical Review Letters*, vol. 99, no. 23, p. 235503, 2007.

[31] A. S. Keys, A. R. Abate, S. C. Glotzer, and D. J. Durian, "Measurement of growing dynamical length scales and prediction of the jamming transition in a granular material," *Nature Physics*, vol. 3, no. 4, pp. 260–264, 2007.

[32] S. C. Glotzer and A. S. Keys, "Materials science: A tale of two tilings," *Nature*, vol. 454, no. 7203, pp. 420–421, 2008.

[33] A. Haji-Akbari, M. Engel, A. S. Keys, X. Zheng, R. G. Petschek, P. Palffy-Muhoray, and S. C. Glotzer, "Disordered, quasicrystalline and crystalline phases of densely packed tetrahedra," *Nature*, vol. 462, no. 7274, pp. 773–777, 2009.

[34] A. S. Keys, C. R. Iacovella, and S. C. Glotzer, "Characterizing structure through shape matching and applications to self-assembly," *Annual Reviews of Condensed Matter Physics*, vol. 2, 2011.

[35] C. R. Iacovella, A. S. Keys, and S. C. Glotzer, "Self-assembly of soft matter quasicrystals." preprint, 2010.

[36] A. S. Keys, H. L. O., J. P. Garrahan, S. C. Glotzer, and D. Chandler, "Structure of localized excitations and relaxation in supercooled glass-forming liquids." preprint, 2011.

[37] A. S. Keys, C. R. Iacovella, and S. C. Glotzer, "Characterizing structure in assembled systems using shape matching : Algorithms and applications." preprint, 2010.

[38] A. S. Keys, C. R. Iacovella, and S. C. Glotzer, "Harmonic shape descriptors applied to characterizing assembled structures." preprint, 2010.

[39] P. Jenniskens and D. F. Blake, "Structural transitions in amorphous water ice and astrophysical implications," *Science*, vol. 265, no. 5173, p. 753, 1994.

[40] J. C. Angus and C. C. Hayman, "Low-pressure, metastable growth of diamond and" diamondlike" phases," *Science*, vol. 241, no. 4868, p. 913, 1988.

[41] L. Onsager, "Crystal statistics. I. A two-dimensional model with an order-disorder transition," *Physical Review*, vol. 65, no. 3-4, pp. 117–149, 1944.

[42] J. W. Cahn and J. E. Hilliard, "Spinodal decomposition: A reprise," *Acta Metallurgica*, vol. 19, no. 2, pp. 151–161, 1971.

[43] B. Peters and B. L. Trout, "Obtaining reaction coordinates by likelihood maximization," *The Journal of Chemical Physics*, vol. 125, p. 054108, 2006.

[44] F. C. Frank and J. S. Kasper, "Complex alloy structures regarded as sphere packings. I. Definitions and basic principles," *Acta Crystallographica*, vol. 11, no. 3, pp. 184–190, 1958.

[45] J. L. Finney, "Random packings and the structure of simple liquids. I. The geometry of random close packing," *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, vol. 319, no. 1539, pp. 479–493, 1970.

[46] D. R. Nelson, "Order, frustration, and defects in liquids and glasses," *Physical Review B*, vol. 28, no. 10, pp. 5515–5535, 1983.

[47] D. R. Nelson and F. Spaepen, "Polytetrahedral order in condensed matter," *Solid State Physics*, vol. 42, pp. 1–90, 1989.

[48] B. I. Halperin and D. R. Nelson, "Theory of two-dimensional melting," *Physical Review Letters*, vol. 41, no. 2, pp. 121–124, 1978.

[49] F. C. Frank, "Supercooling of liquids," *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, vol. 215, no. 1120, pp. 43–46, 1952.

[50] T. C. Hales, "A proof of the Kepler conjecture," *Annals of Mathematics-Second Series*, vol. 162, no. 3, pp. 1065–1186, 2005.

[51] M. O. Robbins, K. Kremer, and G. S. Grest, "Phase diagram and dynamics of Yukawa systems," *The Journal of Chemical Physics*, vol. 88, p. 3286, 1988.

[52] J. Roth and A. R. Denton, "Solid-phase structures of the Dzugutov pair potential," *Physical Review E*, vol. 61, pp. 6845–6857, Jun 2000.

[53] N. D. Mermin, "Crystalline order in two dimensions," *Physical Review*, vol. 176, no. 1, pp. 250–254, 1968.

[54] D. R. Nelson and B. I. Halperin, "Dislocation-mediated melting in two dimensions," *Physical Review B*, vol. 19, no. 5, pp. 2457–2484, 1979.

[55] J. W. Gibbs, *The Scientific Papers*. Dover Publications, 1961.

[56] D. W. Oxtoby, "Homogeneous nucleation: Theory and experiment," *Journal of Physics: Condensed matter*, vol. 4, no. 38, pp. 7627–7650, 1992.

[57] P. R. ten Wolde, M. J. Ruiz-Montero, and D. Frenkel, "Numerical calculation of the rate of crystal nucleation in a Lennard-Jones system at moderate undercooling," *The Journal of Chemical Physics*, vol. 104, pp. 9932–9947, Jun 1996.

[58] A. Cacciuto, S. Auer, and D. Frenkel, "Onset of heterogeneous crystal nucleation in colloidal suspensions," *Nature*, vol. 428, no. 6981, pp. 404–406, 2004.

[59] P. R. ten Wolde, D. Frenkel, and S. Matter, "Homogeneous nucleation and the Ostwald step rule," *Physical Chemistry Chemical Physics*, vol. 1, pp. 2191–2196, 1999.

[60] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, "Transition path sampling: Throwing ropes over rough mountain passes, in the dark," *Annual Review of Physical Chemistry*, vol. 53, pp. 291–318, 2002.

[61] S. Auer and D. Frenkel, "Prediction of absolute crystal-nucleation rate in hard-sphere colloids," *Nature*, vol. 409, no. 6823, pp. 1020–1023, 2001.

[62] S. Auer and D. Frenkel, "Numerical prediction of absolute crystallization rates in hard-sphere colloids," *The Journal of Chemical Physics*, vol. 120, pp. 3015–3029, Feb 2004.

[63] S. C. Glotzer, "Spatially heterogeneous dynamics in liquids: Insights from simulation," *Journal of Non-Crystalline Solids*, vol. 274, no. 1-3, pp. 342–355, 2000.

[64] E. R. Weeks, J. C. Crocker, A. C. Levitt, A. Schofield, and D. A. Weitz, "Three-dimensional direct imaging of structural relaxation near the colloidal glass transition," *Science*, vol. 287, no. 5453, p. 627, 2000.

[65] O. Dauchot, G. Marty, and G. Biroli, "Dynamical heterogeneity close to the jamming transition in a sheared granular material," *Physical Review Letters*, vol. 95, no. 26, pp. 265701–265701, 2005.

[66] G. Marty and O. Dauchot, "Subdiffusion and cage effect in a sheared granular material," *Physical Review Letters*, vol. 94, no. 1, p. 15701, 2005.

[67] M. D. Ediger, "Spatially heterogeneous dynamics in supercooled liquids," *Annual Review of Physical Chemistry*, vol. 51, pp. 99–128, 2000.

[68] H. Sillescu, "Heterogeneity at the glass transition: A review," *J. Non-Cryst. Solids*, vol. 243, no. 2-3, pp. 81–108, 1999.

[69] N. Lačević, F. W. Starr, T. B. Schrøder, and S. C. Glotzer, "Spatially heterogeneous dynamics investigated via a time-dependent four-point density correlation function," *The Journal of Chemical Physics*, vol. 119, p. 7372, 2003.

[70] M. N. J. Bergroth and S. C. Glotzer, "Structural signatures of strings and propensity for mobility in a simulated supercooled liquid above the glass transition," *Arxiv preprint cond-mat/0610395*, 2006.

[71] Y. Gebremichael, "Particle dynamics and the development of string-like motion in a simulated monoatomic supercooled liquid," *The Journal of Chemical Physics*, vol. 120, no. 9, p. 4415, 2004.

[72] M. Vogel and S. C. Glotzer, "Spatially heterogeneous dynamics and dynamic facilitation in a model of viscous silica," *Physical Review Letters*, vol. 92, no. 25, p. 255901, 2004.

[73] C. Donati, J. F. Douglas, W. Kob, S. J. Plimpton, P. H. Poole, and S. C. Glotzer, "Stringlike cooperative motion in a supercooled liquid," *Physical Review Letters*, vol. 80, no. 11, pp. 2338–2341, 1998.

[74] C. Donati, S. C. Glotzer, P. H. Poole, W. Kob, and S. J. Plimpton, "Spatial correlations of mobility and immobility in a glass-forming Lennard-Jones liquid," *Physical Review E*, vol. 60, no. 3, pp. 3107–3119, 1999.

[75] C. Donati, S. Franz, S. C. Glotzer, and G. Parisi, "Theory of non-linear susceptibility and correlation length in glasses and liquids," *Journal of Non-Crystalline Solids*, vol. 307, pp. 215–224, 2002.

[76] S. C. Glotzer, V. N. Novikov, and T. Schrøder, "Time-dependent, four-point density correlation function description of dynamical heterogeneity and decoupling in supercooled liquids," *The Journal of Chemical Physics*, vol. 112, no. 2, p. 509, 2000.

[77] F. H. Stillinger and J. A. Hodgdon, "Translation-rotation paradox for diffusion in fragile glass-forming liquids," *Physical Review E*, vol. 50, no. 3, pp. 2064–2068, 1994.

[78] C. A. Angell, "Perspective on the glass transition," *Journal of Physics and Chemistry of Solids*, vol. 49, no. 8, pp. 863–871, 1988.

[79] H. Vogel, "Temperature dependence of viscosity of melts," *Physikalische Zeitschrift*, vol. 22, pp. 645–646, 1921.

[80] G. Tammann and W. Hesse, "Dependence of viscosity on the temperature of supercooled liquids Z," *Anorganische Allgemeine Chemie*, vol. 156, pp. 245–257, 1926.

[81] G. S. Fulcher, "Analysis of recent measurements of the viscosity of glasses," *Journal of the American Ceramic Society*, vol. 75, no. 5, pp. 1043–1055, 1992.

[82] J. T. Bendler and M. F. Shlesinger, "Generalized Vogel law for glass-forming liquids," *Journal of Statistical Physics*, vol. 53, no. 1, pp. 531–541, 1988.

[83] H. Bässler, "Viscous flow in supercooled liquids analyzed in terms of transport theory for random media with energetic disorder," *Physical Review Letters*, vol. 58, no. 8, pp. 767–770, 1987.

[84] R. Richert and H. Bässler, "Dynamics of supercooled melts treated in terms of the random-walk concept," *Journal of Physics: Condensed Matter*, vol. 2, p. 2273, 1990.

[85] R. H. Colby, "Dynamic scaling approach to glass formation," *Physical Review E*, vol. 61, no. 2, pp. 1783–1792, 2000.

[86] M. H. Cohen and G. S. Grest, "Liquid-glass transition, a free-volume approach," *Physical Review B*, vol. 20, no. 3, pp. 1077–1098, 1979.

[87] W. Kauzmann, "The nature of the glassy state and the behavior of liquids at low temperatures.," *Chemical Reviews*, vol. 43, no. 2, pp. 219–256, 1943.

[88] M. Goldstein, "Viscous liquids and the glass transition: A potential energy barrier picture," *The Journal of Chemical Physics*, vol. 51, p. 3728, 2003.

[89] S. Sastry, P. G. Debenedetti, and F. H. Stillinger, "Signatures of distinct dynamical regimes in the energy landscape of a glass-forming liquid," *Nature*, vol. 393, p. 555, 1998.

[90] S. Büchner and A. Heuer, "Potential energy landscape of a model glass former: Thermodynamics, anharmonicities, and finite size effects," *Physical Review E*, vol. 60, no. 6, pp. 6507–6518, 1999.

[91] S. Sastry, "The relationship between fragility, configurational entropy and the potential energy landscape of glass-forming liquids," *Nature*, vol. 409, no. 6817, pp. 164–167, 2001.

[92] W. Gotze and L. Sjogren, "Relaxation processes in supercooled liquids," *Reports on Progress in Physics*, vol. 55, no. 3, pp. 241–376, 1992.

[93] D. R. Reichman and P. Charbonneau, "Mode-coupling theory," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2005, p. P05013, 2005.

[94] G. Biroli, J. P. Bouchaud, K. Miyazaki, and D. R. Reichman, "Inhomogeneous mode-coupling theory and growing dynamic length in supercooled liquids," *Physical Review Letters*, vol. 97, no. 19, p. 195701, 2006.

[95] G. Biroli and J. P. Bouchaud, "Critical fluctuations and breakdown of the Stokes–Einstein relation in the mode-coupling theory of glasses," *Journal of Physics: Condensed Matter*, vol. 19, p. 205101, 2007.

[96] G. Adam and J. H. Gibbs, "On the temperature dependence of cooperative relaxation properties in glass-forming liquids," *The Journal of Chemical Physics*, vol. 43, p. 139, 1965.

[97] N. Giovambattista, S. V. Buldyrev, F. W. Starr, and H. E. Stanley, "Connection between Adam-Gibbs theory and spatially heterogeneous dynamics," *Physical Review Letters*, vol. 90, no. 8, p. 85506, 2003.

[98] F. H. Stillinger and P. G. Debenedetti, "Energy landscape diversity and supercooled liquid properties," *The Journal of Chemical Physics*, vol. 116, p. 3353, 2002.

[99] B. Doliwa and A. Heuer, "Hopping in a supercooled Lennard-Jones liquid: Metabasins, waiting time distribution, and diffusion," *Physical Review E*, vol. 67, no. 3, p. 30501, 2003.

[100] X. Xia and P. G. Wolynes, "Fragilities of liquids predicted from the random first order transition theory of glasses," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 7, p. 2990, 2000.

[101] J. D. Stevenson and P. G. Wolynes, "Thermodynamic- kinetic correlations in supercooled liquids: A critical survey of experimental data and predictions of the random first-order transition theory of glasses," *Journal of Physical Chemistry B*, vol. 109, no. 31, pp. 15093–15097, 2005.

[102] J. S. Langer, "Dynamics and thermodynamics of the glass transition," *Physical Review E*, vol. 73, no. 4, p. 41504, 2006.

[103] G. H. Fredrickson and H. C. Andersen, "Kinetic Ising model of the glass transition," *Physical Review Letters*, vol. 53, no. 13, pp. 1244–1247, 1984.

[104] R. G. Palmer, D. L. Stein, E. Abrahams, and P. W. Anderson, "Models of hierarchically constrained dynamics for glassy relaxation," *Physical Review Letters*, vol. 53, no. 10, pp. 958–961, 1984.

[105] J. P. Garrahan and D. Chandler, "Geometrical explanation and scaling of dynamical heterogeneities in glass forming systems," *Physical Review Letters*, vol. 89, no. 3, p. 35704, 2002.

[106] J. P. Garrahan and D. Chandler, "Coarse-grained microscopic model of glass formers," *Proceedings of the National Academy of Sciences*, vol. 100, no. 17, p. 9710, 2003.

[107] Y. S. Elmatad, D. Chandler, and J. P. Garrahan, "Corresponding states of structural glass formers," *Journal of Physical Chemistry B*, vol. 113, no. 16, pp. 5563–5567, 2009.

[108] D. Chandler and J. P. Garrahan, "Dynamics on the way to forming glass: Bubbles in space-time," *Annual Review of Physical Chemistry*, vol. 61, no. 1, pp. 191–217, 2010.

[109] L. O. Hedges, R. L. Jack, J. P. Garrahan, and D. Chandler, "Dynamic order-disorder in atomistic models of structural glass formers," *Science*, vol. 323, no. 5919, p. 1309, 2009.

[110] A. C. Pan, J. P. Garrahan, and D. Chandler, "Decoupling of self-diffusion and structural relaxation during a fragile-to-strong crossover in a kinetically constrained lattice gas," *ChemPhysChem*, vol. 6, no. 9, pp. 1783–1785, 2005.

[111] L. O. Hedges, L. Maibaum, D. Chandler, and J. P. Garrahan, "Decoupling of exchange and persistence times in atomistic models of glass formers," *The Journal of Chemical Physics*, vol. 127, p. 211101, 2007.

[112] M. H. Cohen and D. Turnbull, "Molecular transport in liquids and glasses," *The Journal of Chemical Physics*, vol. 31, p. 1164, 1959.

[113] D. Turnbull and M. H. Cohen, "Free-volume model of the amorphous phase: Glass transition," *The Journal of Chemical Physics*, vol. 34, p. 120, 1961.

[114] S. A. Kivelson, X. Zhao, D. Kivelson, T. M. Fischer, and C. M. Knobler, "Frustration-limited clusters in liquids," *The Journal of Chemical Physics*, vol. 101, no. 3, p. 2391, 1994.

[115] D. Kivelson and G. Tarjus, "The Kauzmann paradox interpreted via the theory of frustration-limited-domains," *The Journal of Chemical Physics*, vol. 109, p. 5481, 1998.

[116] G. Tarjus, D. Kivelson, and P. Viot, "The viscous slowing down of supercooled liquids as a temperature-controlled super-Arrhenius activated process: A description in terms of frustration-limited domains," *Journal of Physics: Condensed Matter*, vol. 12, no. 29, pp. 6497–6508, 2000.

[117] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*. Clarendon Press, 1989.

[118] J. D. Weeks, D. Chandler, and H. C. Andersen, "Role of repulsive forces in determining the equilibrium structure of simple liquids," *The Journal of Chemical Physics*, vol. 54, no. 12, p. 5237, 1971.

[119] M. Dzugutov, "Formation of a dodecagonal quasicrystalline phase in a simple monatomic liquid," *Physical Review Letters*, vol. 70, no. 19, pp. 2924–2927, 1993.

[120] M. Engel and H. R. Trebin, "Self-assembly of monatomic complex crystals and quasicrystals with a double-well interaction potential," *Physical Review Letters*, vol. 98, no. 22, p. 225505, 2007.

[121] F. C. Frank and J. S. Kasper, "Complex alloy structures regarded as sphere packings. II. Analysis and classification of representative structures," *Acta Crystallographica*, vol. 12, no. 7, pp. 483–499, 1959.

[122] D. Shechtman, I. Blech, D. Gratias, and J. W. Cahn, "Metallic phase with long-range orientational order and no translational symmetry," *Physical Review Letters*, vol. 53, no. 20, pp. 1951–1953, 1984.

[123] C. Janot, *Quasicrystals: A Primer*. Clarendon Press, 1994.

[124] J. Mikhael, J. Roth, L. Helden, and C. Bechinger, "Archimedean-like tiling on decagonal quasicrystalline surfaces," *Nature*, vol. 454, no. 7203, pp. 501–504, 2008.

[125] D. V. Talapin, E. V. Chevchenko, M. I. Bodnarchuk, Y. X., J. Chen, and C. B. Murray, "Quasicrystalline order in self-assembled binary nanoparticle superlattices," *Nature*, vol. 461, pp. 964–967, 10 2009.

[126] A. I. Goldman and R. F. Kelton, "Quasicrystals and crystalline approximants," *Reviews of Modern Physics*, vol. 65, pp. 213–230, Jan 1993.

[127] M. Dzugutov, "Glass-formation in a simple monatomic liquid," *Physical Review A*, vol. 46, pp. R2984–R2987, Sep 1992.

[128] M. Dzugutov, S. I. Simdyankin, and F. H. M. Zetterling, "Decoupling of diffusion from structural relaxation and spatial heterogeneity in a supercooled simple liquid," *Physical Review Letters*, vol. 89, p. 195701, Nov 2002.

[129] F. H. M. Zetterling, M. Dzugutov, and S. I. Simdyankin, "Formation of large-scale icosahedral clusters in a simple liquid approaching the glass transition," *Journal of Non-Crystalline Solids*, vol. 293, pp. 39–44, Nov 2001.

[130] F. H. M. Zetterling, M. Dzugutov, and S. Lidin, "Gamma-brass crystallization in a simple monatomic liquid," in *Materials Research Society Symposium Proceedings*, vol. 643, pp. 9–9, 2001.

[131] M. Engel and H. R. Trebin, "Structural complexity in monodisperse systems of isotropic particles," *Zeitschrift für Kristallographie*, vol. 223, no. 11-12, pp. 721–725, 2008.

[132] P. J. Steinhardt, D. R. Nelson, and M. Ronchetti, "Bond-orientational order in liquids and glasses," *Physical Review B*, vol. 28, no. 2, pp. 784–805, 1983.

[133] W. Kob and H. C. Andersen, "Testing mode-coupling theory for a supercooled binary Lennard-Jones mixture I: The van Hove correlation function," *Physical Review E*, vol. 51, no. 5, pp. 4626–4641, 1995.

[134] G. Wahnström, "Molecular-dynamics study of a supercooled two-component Lennard-Jones system," *Physical Review A*, vol. 44, no. 6, pp. 3752–3764, 1991.

[135] D. N. Perera and P. Harrowell, "Stability and structure of a supercooled liquid mixture in two dimensions," *Physical Review E*, vol. 59, no. 5, pp. 5721–5743, 1999.

[136] M. M. Hurley and P. Harrowell, "Kinetic structure of a two-dimensional liquid," *Physical Review E*, vol. 52, no. 2, pp. 1694–1698, 1995.

[137] A. Widmer-Cooper, P. Harrowell, and H. Fynewever, "How reproducible are dynamic heterogeneities in a supercooled liquid?," *Physical Review Letters*, vol. 93, no. 13, p. 135701, 2004.

[138] T. Kawasaki, T. Araki, and H. Tanaka, "Correlation between dynamic heterogeneity and medium-range order in two-dimensional glass-forming liquids," *Physical Review Letters*, vol. 99, no. 21, p. 215701, 2007.

[139] H. Shintani and H. Tanaka, "Frustration on the way to crystallization in glass," *Nature Physics*, vol. 2, no. 3, pp. 200–206, 2006.

[140] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*. Academic Press, Inc. Orlando, F. L., U. S.A, 2002.

[141] J. S. Liu and R. Chen, "Sequential Monte Carlo methods for dynamic systems," *Journal of the American Statistical Association*, vol. 93, no. 443, pp. 1032–1044, 1998.

[142] S. Duane, A. D. Kennedy, B. Pendleton, and D. Roweth, "Hybrid Monte Carlo," *Physics Letters B*, vol. 195, no. 2, pp. 216–222, 1987.

[143] G. J. Martyna, M. L. Klein, and M. Tuckerman, "Nosé-Hoover chains: The canonical ensemble via continuous dynamics," *The Journal of Chemical Physics*, vol. 97, no. 4, 1992.

[144] D. J. Evans and B. L. Holian, "The Nosé-Hoover thermostat," *The Journal of Chemical Physics*, vol. 83, p. 4069, 1985.

[145] W. G. Hoover, "Constant-pressure equations of motion," *Physical Review A*, vol. 34, no. 3, pp. 2499–2500, 1986.

[146] R. Car and M. Parrinello, "Unified approach for molecular dynamics and density-functional theory," *Physical Review Letters*, vol. 55, no. 22, pp. 2471–2474, 1985.

[147] M. Grünwald, E. Rabani, and C. Dellago, "Mechanisms of the wurtzite to rock-salt transformation in CdSe nanocrystals," *Physical Review Letters*, vol. 96, no. 25, p. 255701, 2006.

[148] S. J. Plimpton, "Fast parallel algorithms for short-range molecular dynamics," *Journal of Computational Physics*, vol. 117, pp. 1–19, 1995.

[149] J. A. Anderson, C. D. Lorenz, and A. Travesset, "General purpose molecular dynamics simulations fully implemented on graphics processing units," *Journal of Computational Physics*, vol. 227, no. 10, pp. 5342–5359, 2008.

[150] J. R. Shewchuk, "An introduction to the conjugate gradient method without the agonizing pain." Carnegie Mellon University, Pittsburg, PA, Technical Report CMU-CS-94-125, 1994.

[151] E. Bitzek, P. Koskinen, F. Gähler, M. Moseler, and P. Gumbsch, "Structural relaxation made simple," *Physical Review Letters*, vol. 97, no. 17, p. 170201, 2006.

[152] C. Jarzynski, "Nonequilibrium equality for free energy differences," *Physical Review Letters*, vol. 78, no. 14, pp. 2690–2693, 1997.

[153] C. Jarzynski, "Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach," *Physical Review E*, vol. 56, no. 5, pp. 5018–5035, 1997.

[154] J. G. Kirkwood, "Statistical mechanics of pure fluids," *The Journal of Chemical Physics*, vol. 3, pp. 300–313, 1935.

[155] R. W. Zwanzig, "High-temperature equation of state by a perturbation method. I. Nonpolar gases," *The Journal of Chemical Physics*, vol. 22, p. 1420, 1954.

[156] G. M. Torrie and J. P. Valleau, "Non-physical sampling distributions in Monte-Carlo free-energy estimation - Umbrella Sampling," *Journal of Computational Physics*, vol. 23, no. 2, pp. 187–199, 1977.

[157] C. Dellago, P. G. Bolhuis, F. S. Csajka, and D. Chandler, "Transition path sampling and the calculation of rate constants," *The Journal of Chemical Physics*, vol. 108, p. 1964, 1998.

[158] A. Laio and M. Parrinello, "Escaping free-energy minima," *Proceedings of the National Academy of Sciences*, vol. 99, no. 20, pp. 12562–12566, 2002.

[159] C. H. Bennett, A. S. Nowick, and J. J. Burton, "Diffusion in solids, recent developments," 1975.

[160] D. Chandler, "Statistical mechanics of isomerization dynamics in liquids and the transition state approximation," *The Journal of Chemical Physics*, vol. 68, p. 2959, 1978.

[161] J. L. Yarnell, M. J. Katz, R. G. Wenzel, and S. H. Koenig, "Structure factor and radial distribution function for liquid argon at 85 K," *Physical Review A*, vol. 7, no. 6, pp. 2130–2144, 1973.

[162] B. Cullity and S. Stock, *Elements of X-ray Diffraction*. Prentice Hall, Upper Saddle River, NJ, 2001.

[163] J. W. Roth, R. Schilling, and H. R. Trebin, "Nucleation of quasicrystals by rapid cooling of a binary melt: A molecular-dynamics study," *Physical Review B*, vol. 51, no. 22, pp. 15833–15840, 1995.

[164] G. D. Bergland, "A guided tour of the fast Fourier transform," *IEEE Spectrum*, vol. 6, no. 7, pp. 41–52, 2009.

[165] J. Roth and A. R. Denton, "Solid-phase structures of the Dzugutov pair potential," *Physical Review E*, vol. 61, no. 6, pp. 6845–6857, 2000.

[166] W. Kob, C. Donati, S. J. Plimpton, P. H. Poole, and S. C. Glotzer, "Dynamical heterogeneities in a supercooled Lennard-Jones liquid," *Physical Review Letters*, vol. 79, no. 15, pp. 2827–2830, 1997.

[167] C. Bennemann, C. Donati, J. Baschnagel, and S. C. Glotzer, "Growing range of correlated motion in a polymer melt on cooling towards the glass transition," *Nature*, vol. 399, no. 6733, pp. 246–249, 1999.

[168] M. D. Ediger, "Movies of the glass transition," *Science*, vol. 287, no. 5453, pp. 604–605, 2000.

[169] A. Rahman, "Correlations in the motion of atoms in liquid argon," *Physical Review*, vol. 136, no. 2A, pp. 405–411, 1964.

[170] Y. Gebremichael, T. B. Schrøder, F. W. Starr, and S. C. Glotzer, "Spatially correlated dynamics in a simulated glass-forming polymer melt: Analysis of clustering phenomena," *Physical Review E*, vol. 64, no. 5, p. 51503, 2001.

[171] G. Biroli and J. P. Bouchaud, "Diverging length scale and upper critical dimension in the Mode-Coupling Theory of the glass transition," *Europhysics Letters*, vol. 67, p. 21, 2004.

[172] M. Merolle, J. P. Garrahan, and D. Chandler, "Space–time thermodynamics of the glass transition," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 31, p. 10837, 2005.

[173] C. Donati, S. C. Glotzer, P. H. Poole, W. Kob, and S. J. Plimpton, "Spatial correlations of mobility and immobility in a glass-forming Lennard-Jones liquid," *Physical Review E*, vol. 60, no. 3, pp. 3107–3119, 1999.

[174] D. Chandler, J. P. Garrahan, R. L. Jack, L. Maibaum, and A. C. Pan, "Lengthscale dependence of dynamic four-point susceptibilities in glass formers," *Physical Review E*, vol. 74, no. 5, p. 51501, 2006.

[175] P. J. Steinhardt, "Solid-state physics - How does your quasicrystal grow?," *Nature*, vol. 452, no. 7183, pp. 43–44, 2008.

[176] K. F. Kelton, "Crystal nucleation in liquids and glasses," *Solid State Physics-Advances in Research and Applications*, vol. 45, pp. 75–177, 1991.

[177] U. Grimm and D. Joseph, *Quasicrystals: An Introduction to Structure, Physical Properties, and Applications*, ch. Modelling Quasicrystal Growth, pp. 49–66. Springer, 2002.

[178] D. Levine and P. J. Steinhardt, "Quasicrystals - a new class of ordered structures," *Physical Review Letters*, vol. 53, no. 26, pp. 2477–2480, 1984.

[179] H. C. Jeong and P. J. Steinhardt, "Constructing penrose-like tilings from a single prototile and the implications for quasicrystals," *Physical Review B*, vol. 55, pp. 3520–3532, Feb 1997.

[180] V. Elser, "Quasicrystals - a new class of ordered structures - comment," *Physical Review Letters*, vol. 54, no. 15, pp. 1730–1730, 1985.

[181] M. Oxborrow and C. L. Henley, "Random square-triangle tilings - a model for twelvefold-symmetrical quasi-crystals," *Physical Review B*, vol. 48, pp. 6966–6998, Sep 1993.

[182] J. P. K. Doye, D. J. Wales, F. H. M. Zetterling, and M. Dzugutov, "The favored cluster structures of model glass formers," *The Journal of Chemical Physics*, vol. 118, pp. 2792–2799, Feb 2003.

[183] M. Dzugutov, "Formation of a dodecagonal quasi-crystalline phase in a simple monatomic liquid," *Physical Review Letters*, vol. 70, pp. 2924–2927, May 1993.

[184] F. Trudu, D. Donadio, and M. Parrinello, "Freezing of a Lennard-Jones fluid: From nucleation to spinodal regime," *Physical Review Letters*, vol. 97, no. 10, p. 105701, 2006.

[185] J. E. S. Socolar, T. C. Lubensky, and P. J. Steinhardt, "Phonons, phasons, and dislocations in quasi-crystals," *Physical Review B*, vol. 34, pp. 3345–3360, Sep 1986.

[186] S. I. Simdyankin, S. N. Taraskin, M. Dzugutov, and S. R. Elliott, "Vibrational properties of the one-component sigma phase," *Physical Review B*, vol. 62, pp. 3223–3231, Aug 2000.

[187] C. R. Iacovella, A. S. Keys, M. A. Horsch, and S. C. Glotzer, "Icosahedral packing of polymer-tethered nanospheres and stabilization of the gyroid phase," *Physical Review E*, vol. 75, p. 040801(R), 2007.

[188] P. J. Steinhardt, D. R. Nelson, and M. Ronchetti, "Bond-orientational order in liquids and glasses," *Physical Review B*, vol. 28, no. 2, pp. 784–805, 1983.

[189] C. Janot, *Quasicrystals: A Primer*. Oxford University Press, U. S.A, 1997.

[190] J. M. Dubois, "New prospects from potential applications of quasicrystalline materials," *Materials Science & Engineering A*, vol. 294, pp. 4–9, 2000.

[191] X. Zeng, G. Ungar, Y. Liu, V. Percec, A. E. Dulcey, and J. K. Hobbs, "Supramolecular dendritic liquid quasicrystals," *Nature*, vol. 428, no. 6979, pp. 157–160, 2004.

[192] Y. Roichman and D. G. Grier, "Holographic assembly of quasicrystalline photonic heterostructures," *Optics Express*, vol. 13, pp. 5434–5439, Jul 11 2005.

[193] J. Mikhael, J. Roth, L. Helden, and C. Bechinger, "Archimedean-like tiling on decagonal quasicrystalline surfaces," *Nature*, vol. 454, pp. 501–504, Jul 24 2008.

[194] K. Hayashida, T. Dotera, A. Takano, and Y. Matsushita, "Polymeric quasicrystal: Mesoscopic quasicrystalline tiling in ABC star polymers," *Physical Review Letters*, vol. 98, May 11 2007.

[195] S. Lee, M. J. Bluemle, and F. S. Bates, "Discovery of a Frank-Kasper σ phase in sphere-forming block copolymer melts," *Science*, vol. 330, no. 6002, p. 349, 2010.

[196] G. Ungar, Y. Liu, X. Zeng, V. Percec, and W. D. Cho, "Giant supramolecular liquid crystal lattice," *Science*, vol. 299, no. 5610, pp. 1208–1211, 2003.

[197] G. M. Grason and R. D. Kamien, "Self-consistent field theory of multiply branched block copolymer melts," *Physical Review E*, vol. 71, p. 051801, May 2005.

[198] C. R. Iacovella, M. A. Horsch, Z. Zhang, and S. C. Glotzer, "Phase diagrams of self-assembled mono-tethered nanospheres from molecular simulation and comparison to surfactants," *Langmuir*, vol. 21, no. 21, pp. 9488–9494, 2005.

[199] C. R. Iacovella and S. C. Glotzer, "Complex crystal structures formed by the self assembly of di-tethered nanospheres," *Nano Letters*, vol. 9, no. 3, pp. 1206–1211, 2009.

[200] R. C. Mucic, J. J. Storhoff, C. A. Mirkin, and R. L. Letsinger, "DNA-directed synthesis of binary nanoparticle network materials," *Journal of the American Chemical Society*, vol. 120, pp. 12674–12675, 11 1998.

[201] J. Lee, A. O. Govorov, and N. A. Kotov, "Nanoparticle assemblies with molecular springs: A nanoscale thermometer," *Angewandte Chemie-International Edition*, vol. 44, no. 45, pp. 7439–7442.

[202] M. A. Horsch, Z. Zhang, and S. C. Glotzer, "Simulation studies of self-assembly of end-tethered nanorods in solution and role of rod aspect ratio and tether length," *The Journal of Chemical Physics*, vol. 125, Nov 2006.

[203] D. Weaire and R. Phelan, "A counterexample to Kelvin's conjecture on minimal surfaces," *Philosophical Magazine Letters*, vol. 69, no. 2, pp. 107–110, 1994.

[204] N. Lu, J. K. Singh, and D. A. Kofke, "Appropriate methods to combine forward and reverse free-energy perturbation averages," *The Journal of Chemical Physics*, vol. 118, p. 2977, 2003.

[205] E. J. Meijer, D. Frenkel, R. A. Lesar, and A. J. C. Ladd, "Location of melting-point at 300-K of nitrogen by Monte Carlo simulation," *The Journal of Chemical Physics*, vol. 92, pp. 7570–7575, 6 1990.

[206] X. Zeng and G. Ungar, "Inflation rules of square-triangle tilings: From approximants to dodecagonal liquid quasicrystals," *Philosophical Magazine*, vol. 86, no. 6, pp. 1093–1103, 2006.

[207] J. A. Champion, Y. K. Katare, and S. Mitragotri, "Making polymeric micro- and nanoparticles of complex shapes," *Proceedings of the National Academy of Sciences*, vol. 104, pp. 11901–11904, 07 2007.

[208] Y.-S. Cho, G.-R. Yi, J.-M. Lim, S.-H. Kim, V. N. Manoharan, D. J. Pine, and S.-M. Yang, "Self-organization of bidisperse colloids in water droplets," *Journal of the American Chemical Society*, vol. 127, no. 45, pp. 15968–15975, 2005.

[209] C. R. Iacovella and S. C. Glotzer, "Complex crystal structures formed by the self assembly of di-tethered nanospheres," *Nano Letters*, vol. 9, no. 3, pp. 1206–1211, 2009.

[210] Z. Zhang, M. A. Horsch, M. H. Lamm, and S. C. Glotzer, "Tethered nano building blocks: Toward a conceptual framework for nanoparticle self-assembly," *Nano Letters*, vol. 3, no. 10, pp. 1341–1346, 2003.

[211] G. Biroli, "Jamming: A new kind of phase transition?," *Nature Physics*, vol. 3, no. 4, pp. 222–223, 2007.

[212] W. K. Kegel and A. van Blaaderen, "Direct observation of dynamical heterogeneities in colloidal hard-sphere suspensions," *Science*, vol. 287, no. 5451, pp. 290–3, 2000.

[213] O. Pouliquen, M. Belzons, and M. Nicolas, "Fluctuating particle motion during shear induced granular compaction," *Physical Review Letters*, vol. 91, no. 1, p. 14301, 2003.

[214] C. S. O'Hern, S. A. Langer, A. J. Liu, and S. R. Nagel, "Random packings of frictionless particles," *Physical Review Letters*, vol. 88, no. 7, p. 75507, 2002.

[215] C. S. O'Hern, L. E. Silbert, A. J. Liu, and S. R. Nagel, "Jamming at zero temperature and zero applied stress: The epitome of disorder," *Physical Review E*, vol. 68, no. 1, p. 11306, 2003.

[216] A. J. Liu and S. R. Nagel, "Jamming is not just cool any more," *Nature*, vol. 396, no. 6706, pp. 21–22, 1998.

[217] J. D. Stevenson, J. Schmalian, and P. G. Wolynes, "The shapes of cooperatively rearranging regions in glass-forming liquids," *Nature Physics*, vol. 2, no. 4, pp. 268–274, 2006.

[218] A. R. Abate and D. J. Durian, "Approach to jamming in an air-fluidized granular bed," *Physical Review E*, vol. 74, no. 3, p. 31308, 2006.

[219] M. Aichele, Y. Gebremichael, F. W. Starr, J. Baschnagel, and S. C. Glotzer, "Polymer-specific effects of bulk relaxation and stringlike correlated motion in the dynamics of a supercooled polymer melt," *The Journal of Chemical Physics*, vol. 119, p. 5290, 2003.

[220] M. B. Stone, D. P. Bernstein, R. Barry, M. D. Pelc, Y. K. Tsui, and P. Schiffer, "Stress propagation: Getting to the bottom of a granular medium," *Nature*, vol. 427, no. 6974, pp. 503–504, 2004.

[221] E. I. Corwin, H. M. Jaeger, and S. R. Nagel, "Structural signature of jamming in granular media," *Nature*, vol. 435, no. 7045, pp. 1075–1078, 2005.

[222] L. E. Silbert, A. J. Liu, and S. R. Nagel, "Vibrations and diverging length scales near the unjamming transition," *Physical Review Letters*, vol. 95, no. 9, p. 98301, 2005.

[223] C. Angell, "Formation of glasses from liquids and biopolymers," *Science*, vol. 267, no. 5206, pp. 1924–1924, 1995.

[224] P. G. Debenedetti and F. H. Stillinger, "Supercooled liquids and the glass transition," *Nature*, vol. 410, no. 6825, pp. 259–267, 2001.

[225] M. D. Ediger, C. A. Angell, and S. R. Nagel, "Supercooled liquids and glasses," *Journal of Physical Chemistry*, vol. 100, no. 31, pp. 13200–13212, 1996.

[226] X. Xia and P. G. Wolynes, "Fragilities of liquids predicted from the random first order transition theory of glasses," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 7, p. 2990, 2000.

[227] R. G. Palmer, D. L. Stein, E. Abrahams, and P. W. Anderson, "Models of hierarchically constrained dynamics for glassy relaxation," *Physical Review Letters*, vol. 53, no. 10, pp. 958–961, 1984.

[228] S. H. Glarum, "Dielectric relaxation of polar liquids," *The Journal of Chemical Physics*, vol. 33, p. 1371, 1960.

[229] Y. Gebremichael, M. Vogel, and S. C. Glotzer, "Particle dynamics and the development of string-like motion in a simulated monoatomic supercooled liquid," *The Journal of Chemical Physics*, vol. 120, p. 4415, 2004.

[230] J. P. Garrahan and D. Chandler, "Coarse-grained microscopic model of glass formers," *Proceedings of the National Academy of Sciences*, vol. 100, no. 17, p. 9710, 2003.

[231] F. H. Stillinger and T. A. Weber, "Packing structures and transitions in liquids and solids," *Science*, vol. 225, no. 4666, p. 983, 1984.

[232] C. Dellago, P. G. Bolhuis, and D. Chandler, "On the calculation of reaction rate constants in the transition path ensemble," *The Journal of Chemical Physics*, vol. 110, p. 6617, 1999.

[233] P. G. Bolhuis, "Rare events via multiple reaction channels sampled by path replica exchange," *The Journal of Chemical Physics*, vol. 129, p. 114108, 2008.

[234] D. Chandler, "Statistical mechanics of isomerization dynamics in liquids and the transition state approximation," *The Journal of Chemical Physics*, vol. 68, p. 2959, 1978.

[235] P. Hanggi, P. Talkner, and M. Borkovec, "Reaction-rate theory: Fifty years after Kramers," *Reviews Modern Physics*, vol. 62, no. 2, pp. 251–341, 1990.

[236] R. Candelier, A. Widmer-Cooper, J. K. Kummerfeld, O. Dauchot, G. Biroli, P. Harrowell, and D. R. Reichman, "Avalanches and Dynamical Correlations in supercooled liquids," *Arxiv preprint arXiv:0912.0193*, 2009.

[237] F. Ritort and P. Sollich, "Glassy dynamics of kinetically constrained models," *Advances in Physics*, vol. 52, no. 4, pp. 219–342, 2003.

[238] A. Widmer-Cooper and P. Harrowell, "Central role of thermal collective strain in the relaxation of structure in a supercooled liquid," *Physical Review E*, vol. 80, no. 6, p. 61501, 2009.

[239] T. Keyes, "Neighborship structure and dynamics in supercooled liquids," *The Journal of Chemical Physics*, vol. 110, p. 1097, 1999.

[240] W. Kob and H. C. Andersen, "Testing mode-coupling theory for a supercooled binary Lennard-Jones mixture. II. Intermediate scattering function and dynamic susceptibility," *Physical Review E*, vol. 52, no. 4, pp. 4134–4153, 1995.

[241] D. J. Ashton, L. O. Hedges, and J. P. Garrahan, "Fast simulation of facilitated spin models," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2005, p. P12010, 2005.

[242] M. Vogel and S. C. Glotzer, "Spatially heterogeneous dynamics and dynamic facilitation in a model of viscous silica," *Physical Review Letters*, vol. 92, no. 25, p. 255901, 2004.

[243] M. N. J. Bergroth, M. Vogel, and S. C. Glotzer, "Examination of dynamic facilitation in molecular dynamics simulations of glass-forming liquids," *Journal of Physical Chemistry B*, vol. 109, no. 14, pp. 6748–6753, 2005.

[244] C. Donati, J. F. Douglas, W. Kob, S. J. Plimpton, P. H. Poole, and S. C. Glotzer, "Stringlike cooperative motion in a supercooled liquid," *Physical Review Letters*, vol. 80, no. 11, pp. 2338–2341, 1998.

[245] E. R. Weeks, J. C. Crocker, A. C. Levitt, A. Schofield, and D. A. Weitz, "Three-dimensional direct imaging of structural relaxation near the colloidal glass transition," *Science*, vol. 287, no. 5453, p. 627, 2000.

[246] M. Vogel, B. Doliwa, A. Heuer, and S. C. Glotzer, "Particle rearrangements during transitions between local minima of the potential energy landscape of a binary Lennard-Jones liquid," *The Journal of Chemical Physics*, vol. 120, p. 4404, 2004.

[247] A. Widmer-Cooper, P. Harrowell, and H. Fynewever, "How reproducible are dynamic heterogeneities in a supercooled liquid?," *Physical Review Letters*, vol. 93, no. 13, p. 135701, 2004.

[248] A. Widmer-Cooper and P. Harrowell, "Predicting the long-time dynamic heterogeneity in a supercooled liquid on the basis of short-time heterogeneities," *Physical Review Letters*, vol. 96, no. 18, p. 185701, 2006.

[249] J. D. Weeks, D. Chandler, and H. C. Andersen, "Role of repulsive forces in determining the equilibrium structure of simple liquids," *The Journal of Chemical Physics*, vol. 54, no. 12, p. 5237, 1971.

[250] J. A. Anderson, C. D. Lorenz, and A. Travesset, "General purpose molecular dynamics simulations fully implemented on graphics processing units," *Journal of Computational Physics*, vol. 227, no. 10, pp. 5342–5359, 2008.

[251] L. O. Hedges, L. Maibaum, D. Chandler, and J. P. Garrahan, "Decoupling of exchange and persistence times in atomistic models of glass formers," *The Journal of Chemical Physics*, vol. 127, p. 211101, 2007.

[252] M. M. Hurley and P. Harrowell, "Kinetic structure of a two-dimensional liquid," *Physical Review E*, vol. 52, no. 2, pp. 1694–1698, 1995.

[253] S. C. Glotzer, M. A. Horsch, C. R. Iacovella, Z. Zhang, E. R. Chan, and X. Zhang, "Self-assembly of anisotropic tethered nanoparticle shape amphiphiles," *Current Opinion in Colloid and Interface Science*, vol. 10, pp. 287–295, Dec 2005.

[254] M. C. Daniel and D. Astruc, "Gold nanoparticles: Assembly, supramolecular chemistry, quantum-size-related properties, and applications toward biology, catalysis, and nanotechnology," *Chemical Reviews*, vol. 104, pp. 293–346, Jan 2004.

[255] C. P. Collier, T. Vossmeyer, and J. R. Heath, "Nanocrystal superlattices," *Annual Review of Physical Chemistry*, vol. 49, no. 1, pp. 371–404, 1998.

[256] J. J. Storhoff, A. A. Lazarides, R. C. Mucic, C. A. Mirkin, R. L. Letsinger, and G. C. Schatz, "What controls the optical properties of DNA-linked gold nanoparticle assemblies?," *Journal of the American Chemical Society*, vol. 122, no. 19, pp. 4640–4650, 2000.

[257] K. Kempa, B. Kimball, J. Rybczynski, Z. P. Huang, P. F. Wu, D. Steeves, M. Sennett, M. Giersig, D. V. G. L. N. Rao, D. L. Carnahan, D. Z. Wang, J. Y. Lao, W. Z. Li, and Z. F. Ren, "Photonic crystals based on periodic arrays of aligned carbon nanotubes," *Nano Letters*, vol. 3, no. 1, pp. 13–18, 2002.

[258] B. M. Reinhard, M. Siu, H. Agarwal, A. P. Alivisatos, and J. Liphardt, "Calibration of dynamic molecular rulers based on plasmon coupling between gold nanoparticles," *Nano Letters*, vol. 5, no. 11, pp. 2246–2252, 2005.

[259] P. J. Flory, "Phase equilibria in solutions of rod-like particles," *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, vol. 234, no. 1196, pp. 73–89, 1956.

[260] L. Onsager, "The effects of shape on the interaction of colloidal particles," *Annals of the New York Academy of Sciences*, vol. 51, no. Molecular Interaction, pp. 627–659, 1949.

[261] J. D. Bernal, "A geometrical approach to the structure of liquids," *Nature*, vol. 183, pp. 141–147, 1959.

[262] R. G. Larson, *The Structure and Rheology of Complex Fluids*. Oxford University Press, U. S.A, 2000.

[263] J. D. Honeycutt and H. C. Andersen, "Molecular dynamics study of melting and freezing of small Lennard-Jones clusters," *Journal of Physical Chemistry*, vol. 91, no. 19, pp. 4950–4963, 1987.

[264] A. B. Harris, R. D. Kamien, and T. C. Lubensky, "Molecular chirality and chiral parameters," *Reviews of Modern Physics*, vol. 71, pp. 1745–1757, Oct 1999.

[265] B. I. Halperin and D. R. Nelson, "Theory of two-dimensional melting," *Physical Review Letters*, vol. 41, no. 2, pp. 121–124, 1978.

[266] P. Akcora, H. Liu, S. K. Kumar, J. Moll, Y. Li, B. C. Benicewicz, L. S. Schadler, D. Acehan, A. Z. Panagiotopoulos, V. Pryamitsyn, V. Ganesan, J. Ilavsky, P. Thiyagarajan, R. H. Colby, and J. F. Douglas, "Anisotropic self-assembly of spherical polymer-grafted nanoparticles," *Nature Materials*, vol. 8, no. 4, pp. 354–359, 2009.

[267] Z. H. Nie, D. Fava, E. Kumacheva, S. Zou, G. C. Walker, and M. Rubinstein, "Self-assembly of metal-polymer analogues of amphiphilic triblock copolymers," *Nature Materials*, vol. 6, pp. 609–614, Aug 2007.

[268] Z. Y. Tang, Z. Zhang, Y. Wang, S. C. Glotzer, and N. A. Kotov, "Self-assembly of CdTe nanocrystals into free-floating sheets," *Science*, vol. 314, pp. 274–278, Oct 2006.

[269] Z. Zhang, Z. Y. Tang, N. A. Kotov, and S. C. Glotzer, "Simulations and analysis of self-assembly of CdTe nanoparticles into wires and sheets," *Nano Letters*, vol. 7, pp. 1670–1675, Jun 2007.

[270] S. Srivastava, A. Santos, K. Critchley, K.-S. Kim, P. Podsiadlo, K. Sun, J. Lee, C. Xu, G. D. Lilly, S. C. Glotzer, and N. A. Kotov, "Light-controlled self-assembly of semiconductor nanoparticles into twisted ribbons," *Science*, vol. 327, no. 5971, pp. 1355–1359, 2010.

[271] A. Ahniyaz, Y. Sakamoto, and L. Bergström, "Magnetic field-induced assembly of oriented superlattices from maghemite nanocubes," *Proceedings of the National Academy of Sciences*, vol. 104, no. 45, pp. 17570–17574, 2007.

[272] X. Zhang, Z. Zhang, and S. C. Glotzer, "Simulation study of dipole-induced self-assembly of nanocubes," *The Journal of Physical Chemistry C*, vol. 111, no. 11, pp. 4132–4137, 2007.

[273] B. S. John and F. A. Escobedo, "Phase behavior of colloidal hard tetragonal parallelepipeds (cuboids): A Monte Carlo simulation study," *The Journal of Physical Chemistry B*, vol. 109, no. 48, pp. 23008–23015, 2005.

[274] V. N. Manoharan, M. T. Elsesser, and D. J. Pine, "Dense packing and symmetry in small clusters of microspheres," *Science*, vol. 301, no. 5632, pp. 483–487, 2003.

[275] Y.-S. Cho, G.-R. Yi, J.-M. Lim, S.-H. Kim, V. N. Manoharan, D. J. Pine, and S.-M. Yang, "Self-organization of bidisperse colloids in water droplets," *Journal of the American Chemical Society*, vol. 127, no. 45, pp. 15968–15975, 2005.

[276] D. J. Kraft, J. Goenewold, and K. Kegel, Willem, "Colloidal molecules with well-controlled bond angles," *Soft Matter*, vol. 5, pp. 3823–3826, 2009.

[277] S. A. Vanapalli, C. R. Iacovella, K. E. Sung, D. Mukhija, J. M. Millunchick, M. A. Burns, S. C. Glotzer, and M. J. Solomon, "Fluidic assembly and packing of microspheres in confined channels," *Langmuir*, vol. 24, no. 7, pp. 3661–3670, 2008.

[278] K. E. Sung, S. A. Vanapalli, D. Mukhija, H. A. McKay, J. Mirecki Millunchick, M. A. Burns, and M. J. Solomon, "Programmable fluidic production of microparticles with configurable anisotropy," *Journal of the American Chemical Society*, vol. 130, no. 4, pp. 1335–1340, 2008.

[279] G. Zhang, D. Y. Wang, and H. Mohwald, "Decoration of microspheres with gold nanodots-giving colloidal spheres valences," *Angewandte Chemie-International Edition*, vol. 44, no. 47, pp. 7767–7770, 2005.

[280] Z. Zhang and S. C. Glotzer, "Self-assembly of patchy particles," *Nano Letters*, vol. 4, pp. 1407–1413, Aug 2004.

[281] G. A. DeVries, M. Brunnbauer, Y. Hu, A. M. Jackson, B. Long, B. T. Neltner, O. Uzun, B. H. Wunsch, and F. Stellacci, "Divalent metal nanoparticles," *Science*, vol. 315, pp. 358–361, Jan 2007.

[282] A. M. Jackson, J. W. Myerson, and F. Stellacci, "Spontaneous assembly of subnanometre-ordered domains in the ligand shell of monolayer-protected nanoparticles," *Nature Materials*, vol. 3, pp. 330–336, May 2004.

[283] A. B. Pawar and I. Kretzschmar, "Patchy particles by glancing angle deposition," *Langmuir*, vol. 24, no. 2, pp. 355–358, 2008.

[284] J.-W. Park and E. L. Thomas, "Anisotropic micellar nanoobjects from reactive liquid crystalline rodcoil diblock copolymers," *Macromolecules*, vol. 37, no. 10, pp. 3532–3535, 2004.

[285] E. Reister and G. H. Fredrickson, "Phase behavior of a blend of polymer-tethered nanoparticles with diblock copolymers," *The Journal of Chemical Physics*, vol. 123, no. 21, p. 214903, 2005.

[286] A. Jayaraman and K. S. Schweizer, "Structure and assembly of dense solutions and melts of single tethered nanoparticles," *The Journal of Chemical Physics*, vol. 128, no. 16, 2008.

[287] A. J. Waddon, L. Zheng, R. J. Farris, and E. B. Coughlin, "Nanostructured polyethylene-POSS copolymers: Control of crystallization and aggregation," *Nano Letters*, vol. 2, pp. 1149–1155, Oct 2002.

[288] A. K. Boal, F. Ilhan, J. E. DeRouchey, T. Thurn-Albrecht, T. P. Russell, and V. M. Rotello, "Self-assembly of nanoparticles into structured spherical and network aggregates," *Nature*, vol. 404, no. 6779, pp. 746–748, 2000.

[289] T. Chen, Z. Zhang, and S. C. Glotzer, "A precise packing sequence for self-assembled convex structures," *Proceedings of the National Academy of Sciences*, vol. 104, no. 3, pp. 717–722, 2007.

[290] D. C. Rapaport, "Self-assembly of polyhedral shells: A molecular dynamics study," *Physical Review E*, vol. 70, no. 5, p. 51905, 2004.

[291] J. C. Crocker and D. G. Grier, "Methods of digital video microscopy for colloidal studies," *Journal of Colloid and Interface Science*, vol. 179, pp. 298–310, Apr 1996.

[292] L. T. Shereda, R. G. Larson, and M. J. Solomon, "Local stress control of spatiotemporal ordering of colloidal crystals in complex flows," *Physical Review Letters*, vol. 101, no. 3, p. 38301, 2008.

[293] T. Solomon and M. J. Solomon, "Stacking fault structure in shear-induced colloidal crystallization," *The Journal of Chemical Physics*, vol. 124, no. 13, p. 134905, 2006.

[294] P. Varadan and M. J. Solomon, "Direct visualization of long-range heterogeneous structure in dense colloidal gels," *Langmuir*, vol. 19, no. 3, pp. 509–512, 2003.

[295] P. Coura, S. Legoas, A. Moreira, and F. Sato, "On the structural and stability features of linear atomic suspended chains formed from gold nanowires stretching," *Nano Letters*, Jan 2004.

[296] D. V. Talapin, E. V. Shevchenko, M. I. Bodnarchuk, X. Ye, J. Chen, and C. B. Murray, "Quasicrystalline order in self-assembled binary nanoparticle superlattices," *Nature*, vol. 461, no. 7266, pp. 964–967, 2009.

[297] E. V. Shevchenko, D. V. Talapin, N. A. Kotov, S. O'Brien, and C. B. Murray, "Structural diversity in binary nanoparticle superlattices," *Nature*, vol. 439, pp. 55–59, Jan 2006.

[298] C. R. Iacovella and S. C. Glotzer, "Complex crystal structures formed by the self-assembly of ditethered nanospheres," *Nano Letters*, vol. 9, no. 3, pp. 1206–1211, 2009.

[299] M. E. Leunissen, C. G. Christova, A. P. Hynninen, C. P. Royall, A. I. Campbell, A. Imhof, M. Dijkstra, R. van Roij, and A. van Blaaderen, "Ionic colloidal crystals of oppositely charged particles," *Nature*, vol. 7056, pp. 235–240, Sep 2005.

[300] A. Mohraz and M. J. Solomon, "Direct visualization of colloidal rod assembly by confocal microscopy," *Langmuir*, vol. 21, no. 12, pp. 5298–5306, 2005.

[301] T. D. Nguyen and S. C. Glotzer, "Switchable helical structures formed by the hierarchical self-assembly of laterally tethered nanorods," *Small*, vol. 5, no. 18, pp. 2092–2098, 2009.

[302] A. S. Keys and C. R. Iacovella, "Particle shape matching library and examples," *URL http://www. glotzerlab. engin. umich. edu/shapematching. html*, 2010.

[303] R. C. Veltkamp and M. Hagedoorn, "State of the art in shape matching," *Principles of Visual Information Retrieval*, p. 87, 2001.

[304] X. Lu, D. Colbry, and A. K. Jain, "Three-dimensional model based face recognition," *Pattern Recognition*, vol. 1, pp. 362–366, 2004.

[305] H. D. Tagare, C. C. Jaffe, and J. Duncan, "Medical image databases: A content-based retrieval approach," *Journal of the American Medical Informatics Association*, vol. 4, no. 3, p. 184, 1997.

[306] T. D. Nguyen, Z. Zhang, and S. C. Glotzer, "Molecular simulation study of self-assembly of tethered v-shaped nanoparticles," *The Journal of Chemical Physics*, vol. 129, p. 244903, 2008.

[307] M. A. Horsch, Z. Zhang, and S. C. Glotzer, "Self-assembly of polymer-tethered nanorods," *Physical Review Letters*, vol. 95, no. 5, p. 056105, 2005. 056105.

[308] C. R. Iacovella, M. A. Horsch, Z. Zhang, and S. C. Glotzer, "Phase diagrams of self-assembled mono-tethered nanospheres from molecular simulation and comparison to surfactants," *Langmuir*, vol. 21, no. 21, pp. 9488–9494, 2005.

[309] C. R. Iacovella, M. A. Horsch, and S. C. Glotzer, "Local ordering of polymer-tethered nanospheres and nanorods and the stabilization of the double gyroid phase," *The Journal of Chemical Physics*, vol. 129, p. 044902, 2008.

[310] L. Mak, S. Grandison, and R. J. Morris, "An extension of spherical harmonics to region-based rotationally invariant descriptors for molecular shape description and comparison," *Journal of Molecular Graphics and Modelling*, vol. 26, no. 7, pp. 1035–1045, 2008.

[311] S. Grandison, C. Roberts, and R. J. Morris, "The application of 3D Zernike moments for the description of model-free molecular structure, functional motion, and structural reliability," *Journal of Computational Biology*, vol. 16, no. 3, pp. 487–500, 2009.

[312] J. Roth and A. R. Denton, "Solid-phase structures of the Dzugutov pair potential," *Physical Review E*, vol. 61, no. 6, pp. 6845–6857, 2000.

[313] D. Shechtman, I. Blech, D. Gratias, and J. W. Cahn, "Metallic phase with long-range orientational order and no translational symmetry," *Physical Review Letters*, vol. 53, no. 20, pp. 1951–1953, 1984.

[314] C. R. Iacovella and S. C. Glotzer, "Phase behavior of ditethered nanospheres," *Soft Matter*, vol. 5, pp. 4492 – 4498, 2009.

[315] P. J. Besl and H. D. McKay, "A method for registration of 3-D shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239–256, 1992.

[316] G. H. Dunteman, *Principal Components Analysis*. Sage Publications, Inc., Newbury Park, CA, U.S.A., 1989.

[317] N. Iyer, S. Jayanti, K. Lou, Y. Kalyanaraman, and K. Ramani, "Three-dimensional shape searching: State-of-the-art review and future trends," *Computer-Aided Design*, vol. 37, no. 5, pp. 509–530, 2005.

[318] J. W. H. Tangelder and R. C. Veltkamp, "A survey of content based 3d shape retrieval methods," *Multimedia Tools and Applications*, vol. 39, no. 3, pp. 441–471, 2008.

[319] D. Zhang and G. Lu, "Review of shape representation and description techniques," *Pattern Recognition*, vol. 37, no. 1, pp. 1–19, 2004.

[320] H. Chui and A. Rangarajan, "A new point matching algorithm for non-rigid registration," *Computer Vision and Image Understanding*, vol. 89, no. 2-3, pp. 114–141, 2003.

[321] M. Ankerst, G. Kastenmuller, H. P. Kriegel, and T. Seidl, "3d shape histograms for similarity search and classification in spatial databases," *Lecture Notes in Computer Science*, pp. 207–228, 1999.

[322] R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin, "Shape distributions," *ACM Transactions on Graphics (TOG)*, vol. 21, no. 4, pp. 807–832, 2002.

[323] C. T. Zahn and R. Z. Roskies, "Fourier descriptors for plane closed curves," *IEEE Transactions on Computers*, vol. 21, no. 3, pp. 269–281, 1972.

[324] M. Kazhdan, T. Funkhouser, and S. Rusinkiewicz, "Rotation invariant spherical harmonic representation of 3d shape descriptors," in *Proceedings of the 2003 Eurographics/ACM SIGGRAPH Symposium on Geometry Processing*, p. 164, Eurographics Association, 2003.

[325] A. Khotanzad and Y. H. Hong, "Invariant image recognition by Zernike moments," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 5, pp. 489–497, 1990.

[326] M. Novotni and R. Klein, "3D Zernike descriptors for content based shape retrieval," in *Proceedings of the Eighth ACM Symposium on Solid Modeling and Applications*, pp. 216–225, ACM New York, N. Y., U. S.A, 2003.

[327] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 509–522, 2002.

[328] D. Y. Chen, X. P. Tian, Y. T. Shen, and M. Ouhyoung, "On visual similarity based 3d model retrieval," in *Computer Graphics Forum*, vol. 22, pp. 223–232, Amsterdam: North Holland, 1982-, 2003.

[329] A. M. Lesk, "A toolkit for computational molecular biology. II. On the optimal superposition of two sets of coordinates," *Acta Crystallographica Section A: Foundations of Crystallography*, vol. 42, no. 2, pp. 110–113, 1986.

[330] R. Blankenbecler, M. Ohlsson, C. Peterson, and M. Ringnér, "Matching protein structures with fuzzy alignments," *Proceedings of the National Academy of Sciences*, vol. 100, no. 21, p. 11936, 2003.

[331] C. R. Iacovella, M. A. Horsch, Z. Zhang, and S. C. Glotzer, "Phase diagrams of self-assembled mono-tethered nanospheres from molecular simulation and comparison to surfactants," *Langmuir*, vol. 21, no. 21, pp. 9488–9494, 2005.

[332] A. S. Keys, C. R. Iacovella, and S. C. Glotzer, "Characterizing structure through shape matching: Applications and algorithms." preprint, 2010.

[333] M. Körtgen, G. J. Park, M. Novotni, and R. Klein, "3D shape matching with 3D shape contexts," in *The 7th Central European Seminar on Computer Graphics*, vol. 3, Citeseer, 2003.

[334] M. Hilaga, Y. Shinagawa, T. Kohmura, and T. L. Kunii, "Topology matching for fully automatic similarity estimation of 3D shapes," in *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 203–212, ACM, 2001.

[335] H. Sundar, D. Silver, N. Gagvani, and S. Dickinson, "Skeleton based shape matching and retrieval," in *Shape Modeling International*, vol. 130, p. 139, 2003.

[336] S. Biasotti, S. Marini, M. Mortara, G. Patane, M. Spagnuolo, and B. Falcidieno, "3D shape matching through topological structures," in *Discrete Geometry for Computer Imagery*, pp. 194–203, Springer, 2003.

[337] M. Kazhdan, T. Funkhouser, and S. Rusinkiewicz, "Symmetry descriptors and 3D shape matching," in *Proceedings of the 2004 Eurographics/ACM SIGGRAPH Symposium on Geometry Processing*, pp. 115–123, ACM, 2004.

[338] P. Jiantao, L. Yi, X. Guyu, Z. Hongbin, L. Weibin, and Y. Uehara, "3D model retrieval based on 2D slice similarity measurements," in *2nd International Symposium on 3D Data Processing, Visualization and Transmission, 2004. 3DPVT 2004. Proceedings*, pp. 95–101, 2004.

[339] M. Cosentino Lagomarsino, M. Dogterom, and M. Dijkstra, "Isotropic-nematic transition of long, thin, hard spherocylinders confined in a quasi-two-dimensional planar geometry," *The Journal of Chemical Physics*, vol. 119, pp. 3535–3540, August 2003.

[340] R. C. Veltkamp, "Shape matching: Similarity measures and algorithms," in *Shape Modeling and Applications, S. M.I 2001 International Conference on.*, pp. 188–197, 2001.

[341] P. G. de Gennes and J. Prost, *The Physics of Liquid Crystals*. Oxford University Press, U.S.A, 1995.

[342] A. S. Keys, C. R. Iacovella, and S. C. Glotzer, "Harmonic shape descriptors applied to characterizing assembled structures." preprint, 2010.

[343] E. Krissinel and K. Henrick, "Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions," *Acta Crystallographica Section D: Biological Crystallography*, vol. 60, no. 12, pp. 2256–2268, 2004.

[344] J. S. Yeh, D. Y. Chen, B. Y. Chen, and M. Ouhyoung, "A web-based three-dimensional protein retrieval system by matching visual similarity," *Bioinformatics*, vol. 21, no. 13, p. 3056, 2005.

[345] V. Venkatraman, L. Sael, and D. Kihara, "Potential for protein surface shape analysis using spherical harmonics and 3D Zernike descriptors," *Cell Biochemistry and Biophysics*, vol. 54, no. 1, pp. 23–32, 2009.

[346] U. Gasser, A. Schofield, and D. A. Weitz, "Local order in a supercooled colloidal fluid observed by confocal microscopy," *Journal of Physics: Condensed Matter*, vol. 15, p. S375, 2003.

[347] A. I. Campbell, V. J. Anderson, J. S. Van Duijneveldt, and P. Bartlett, "Dynamical arrest in attractive colloids: The effect of long-range repulsion," *Physical Review Letters*, vol. 94, no. 20, p. 208301, 2005.

[348] B. I. Halperin, T. C. Lubensky, and S. K. Ma, "First-order phase transitions in superconductors and smectic-a liquid crystals," *Physical Review Letters*, vol. 32, no. 6, pp. 292–295, 1974.

[349] W. L. McMillan, "Simple molecular model for the smectic a phase of liquid crystals," *Physical Review A*, vol. 4, no. 3, pp. 1238–1246, 1971.

[350] R. Radhakrishnan, K. E. Gubbins, and M. Sliwinska-Bartkowiak, "Effect of the fluid-wall interaction on freezing of confined fluids: Toward the development of a global phase diagram," *The Journal of Chemical Physics*, vol. 112, no. 24, pp. 11048–11057, 2000.

[351] P. T. Cummings, H. Docherty, C. R. Iacovella, and J. K. Singh, "Phase transitions in nanoconfined fluids: The evidence from simulation and theory," *AIChE Journal*, vol. 56, no. 4, pp. 842–848, 2010.

[352] M. Yamaki, J. Higo, and K. Nagayama, "Size-dependent separation of colloidal particles in two-dimensional convective self-assembly," *Langmuir*, vol. 11, no. 8, pp. 2975–2978, 1995.

[353] P. Bladon and A. C. Griffin, "Self-assembly in living nematics," *Macromolecules*, vol. 26, no. 24, pp. 6604–6610, 1993.

[354] R. J. Allen, D. Frenkel, and P. R. ten Wolde, "Forward flux sampling-type schemes for simulating rare events: Efficiency analysis," *The Journal of Chemical Physics*, vol. 124, p. 194111, 2006.

[355] P. J. Steinhardt, D. R. Nelson, and M. Ronchetti, "Icosahedral bond orientational order in supercooled liquids," *Physical Review Letters*, vol. 47, no. 18, pp. 1297–1300, 1981.

[356] M. J. Freiser, "Ordered states of a nematic liquid," *Physical Review Letters*, vol. 24, no. 19, pp. 1041–1043, 1970.

[357] J. P. Kemp and J. Z. Y. Chen, "Helical structures in proteins," *Biomacromolecules*, vol. 2, no. 2, pp. 389–401, 2001.

[358] E. Jankowski and S. C. Glotzer, "A comparison of new methods for generating energy-minimizing configurations of patchy particles," *The Journal of Chemical Physics*, vol. 131, p. 104104, 2009.

[359] M. P. Allen and M. A. Warren, "Simulation of structure and dynamics near the isotropic-nematic transition," *Physical Review Letters*, vol. 78, no. 7, pp. 1291–1294, 1997.

[360] P. Shilane, P. Min, M. Kazhdan, and T. Funkhouser, "The Princeton shape benchmark," 2004.

[361] D. Ortiz, E. P. Jankowski, and S. C. Glotzer, "Assembly propensity in a system of 2d colloidal tetrominoes." preprint, 2010.

[362] M. Meyers, P. Chen, A. Lin, and Y. Seki, "Biological materials: Structure and mechanical properties," *Progress in Materials Science*, vol. 53, no. 1, pp. 1–206, 2008.

[363] Y. Wang, S. Jian, S. Han, S. Feng, Z. Feng, B. Cheng, and D. Zhang, "Photonic band-gap engineering of quasiperiodic photonic crystals," *Journal of Applied Physics*, vol. 97, no. 10, p. 106112, 2005.

[364] J. Yoon, W. Lee, and E. L. Thomas, "Self-assembly of block copolymers for photonic-bandgap materials," *MRS Bulletin*, vol. 30, no. 10, pp. 721–726, 2005.

[365] A. H. Marcus and S. A. Rice, "Observations of first-order liquid-to-hexatic and hexatic-to-solid phase transitions in a confined colloid suspension," *Physical Review Letters*, vol. 77, no. 12, pp. 2577–2580, 1996.

[366] R. M. Ernst, S. R. Nagel, and G. S. Grest, "Search for a correlation length in a simulation of the glass transition," *Physical Review B*, vol. 43, no. 10, pp. 8070–8080, 1991.

[367] U. Gasser, E. R. Weeks, A. Schofield, P. N. Pusey, and D. A. Weitz, "Real-space imaging of nucleation and growth in colloidal crystallization," *Science*, vol. 292, no. 5515, p. 258, 2001.

[368] M. F. Hagan and D. Chandler, "Dynamic pathways for viral capsid assembly," *Biophysical Journal*, vol. 91, no. 1, pp. 42–54, 2006.

[369] H. W. Kuhn, "The Hungarian method for the assignment problem," *Naval Research Logistics Quarterly*, vol. 2, no. 1-2, pp. 83–97, 1955.

[370] M. J. Mohlenkamp, "A fast transform for spherical harmonics," *Journal of Fourier Analysis and Applications*, vol. 5, no. 2, pp. 159–184, 1999.

[371] A. Hershko and A. Ciechanover, "The Ubiquitin system," *Annual Review of Biochemistry*, vol. 67, no. 1, pp. 425–479, 1998.

[372] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten, "Scalable molecular dynamics with NAMD," *Journal of Computational Chemistry*, vol. 26, no. 16, p. 1781, 2005.

[373] W. Humphrey, A. Dalke, and K. Schulten, "VMD: Visual molecular dynamics," *Journal of Molecular Graphics*, vol. 14, no. 1, pp. 33–38, 1996.

[374] C. R. Iacovella, A. S. Keys, and S. C. Glotzer, "Image processing techniques for structural analysis of particle systems." preprint, 2011.

[375] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley Reading Menlo Park, 1989.

[376] Q. Pu, Y. Leng, and P. T. Cummings, "Rate-dependent energy release mechanism of gold nanowires under elongation," *Journal of the American Chemical Society*, vol. 130, pp. 17907–12, Dec 2008.

[377] C. R. Iacovella, W. R. French, and P. T. Cummings, "The formation of polyhedral structures in elongated gold nanowires." preprint, 2010.

[378] V. D. Blank, S. G. Buga, N. R. Serebryanaya, G. A. Dubitsky, B. N. Mavrin, M. Y. Popov, R. H. Bagramov, V. M. Prokhorov, S. N. Sulyanov, B. A. Kulnitskiy, Y. V. Tatyanin, L. Zhou, Y. Zhou, D. J. Wales, and M. P. Hodges, "Global minima of water clusters (H2O)n, n less than or equal to 21, described by an empirical potential," *Chemical Physics Letters*, vol. 286, no. 1-2, pp. 65–72, 1998.

[379] D. J. Wales, J. P. Doye, A. Dullweber, M. P. Hodges, F. Y. Naumkin, F. Calvo, J. Hernández-Rojas, and T. F. Middleton, "The cambridge cluster database," *URL http://www-wales. ch. cam. ac. uk/CCD. html*, 2001.

[380] F. Calvo, J. P. K. Doye, and D. J. Wales, "Quantum partition functions from classical distributions: Application to rare-gas clusters," *The Journal of Chemical Physics*, vol. 114, p. 7312, 2001.

[381] M. L. Boas, *Mathematical Methods in the Physical Sciences*. John Wiley & Sons New York, 2002.

[382] R. Sadourny, "Conservative finite-difference approximations of the primitive equations on quasi-uniform spherical grids," *Monthly Weather Review*, vol. 100, pp. 136–144, 1972.

[383] D. V. Vranic, D. Saupe, and J. Richter, "Tools for 3D-object retrieval: Karhunen-Loeve transform and spherical harmonics," in *IEEE 2001 Workshop Multimedia Signal Processing*, pp. 293–298, Citeseer, 2001.

[384] A. Goede, R. Preissner, and C. Froemmel, "Voronoi cell: New method for allocation of space among atoms: Elimination of avoidable errors in calculation of atomic volume and density," *Journal of Computational Chemistry*, vol. 18, no. 9, pp. 1113–1123, 1997.

[385] R. N. Barnett and U. Landman, "Cluster-derived structures and conductance fluctuations in nanowires," *Nature*, vol. 387, no. 6635, pp. 788–791, 1997.

[386] P. R. ten Wolde and D. Frenkel, "Enhancement of protein crystal nucleation by critical density fluctuations," *Science*, vol. 277, no. 5334, p. 1975, 1997.

[387] G. Y. Onoda, P. J. Steinhardt, D. P. DiVincenzo, and J. E. S. Socolar, "Growing perfect quasicrystals," *Physical Review Letters*, vol. 60, no. 25, pp. 2653–2656, 1988.

[388] D. S. Sebba, J. J. Mock, D. R. Smith, T. H. LaBean, and A. A. Lazarides, "Reconfigurable core-satellite nanoassemblies as molecularly-driven plasmonic switches," *Nano Letters*, vol. 8, no. 7, pp. 1803–1808, 2008.

[389] A. Mohraz and M. J. Solomon, "Direct visualization of colloidal rod assembly by confocal microscopy," *Langmuir*, vol. 21, no. 12, pp. 5298–5306, 2005.

[390] J. A. Champion, Y. K. Katare, and S. Mitragotri, "Making polymeric micro-and nanoparticles of complex shapes," *Proceedings of the National Academy of Sciences*, vol. 104, no. 29, p. 11901, 2007.

[391] Y. S. Cho, G. R. Yi, J. M. Lim, S. H. Kim, V. N. Manoharan, D. J. Pine, and S. M. Yang, "Self-organization of bidisperse colloids in water droplets," *Journal of the American Chemical Society*, vol. 127, no. 45, pp. 15968–15975, 2005.

[392] H. Yukawa, "On the interaction of elementary particles," in *Proceedings of the Physico-Mathematical Society of Japan*, vol. 17, p. 48, 1935.

[393] A. V. Granato, "Self-interstitials as basic structural units of liquids and glasses," *Journal of Physics and Chemistry of Solids*, vol. 55, no. 10, pp. 931–939, 1994.

[394] A. V. Granato and V. A. Khonik, "An interstitialcy theory of structural relaxation and related viscous flow of glasses," *Physical Review Retters*, vol. 93, no. 15, p. 155502, 2004.

[395] J. D. Joannopoulos, P. R. Villeneuve, and S. Fan, "Photonic crystals," *Solid State Communications*, vol. 102, no. 2-3, pp. 165–173, 1997.

[396] M. Maldovan, A. M. Urbas, N. Yufa, W. C. Carter, and E. L. Thomas, "Photonic properties of bicontinuous cubic microphases," *Physical Review B*, vol. 65, no. 16, p. 165123, 2002.

[397] K. Kataoka, A. Harada, and Y. Nagasaki, "Block copolymer micelles for drug delivery: Design, characterization and biological significance," *Advanced Drug Delivery Reviews*, vol. 47, no. 1, pp. 113–131, 2001.

[398] T. S. van Erp and P. G. Bolhuis, "Elaborating transition interface sampling methods," *Journal of Computational Physics*, vol. 205, no. 1, pp. 157–181, 2005.