Title
Theoretical and Computational Investigation of Self-Assembly in Bio-Inspired and Polymeric Materials

Author
Rogers, Carl E.

Publication Date
2013

Peer reviewed|Thesis/dissertation

Theoretical and Computational Investigation of Self-Assembly in Bio-inspired and Polymeric Materials

by

Carl E. Rogers

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Biophysics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Phillip Geissler, Chair
Professor Daniel Fletcher
Professor Berend Smit

Fall 2013

Theoretical and Computational Investigation of Self-Assembly in Bio-inspired and Polymeric Materials

Abstract

Theoretical and Computational Investigation of Self-Assembly in Bio-inspired and Polymeric Materials

by

Carl E. Rogers

Doctor of Philosophy in Biophysics

University of California, Berkeley

Professor Phillip Geissler, Chair

Fabrication of materials through bottom-up, noncovalent self-assembly has the potential to revolutionize many areas of science and engineering that depend upon the precise arrangement and interaction of nanometer-scale particles. Theory and computation provide a useful route to studying self-assembly of these materials due to their ability to study the kinetic and thermodynamic features of simplified systems at a level of detail not possible in experiments. In this thesis we study the self-assembly of two bio-inspired systems using a combination of theory and computation. The first is comprised of archaeal chaperonin proteins that form in-vitro two remarkably different structures: filamentous chains and layers of sheets. Using a quasi-dynamical Monte Carlo algorithm, we show that the binary decision for sheet or string formation can be explained by allowing for conformational changes between a sheet-favoring state and a string-favoring state. Using advanced sampling techniques, we find that the energy gap for this conformational change controls structure formation. The second system is a self assembling cyclic peptide inside a block copolymer matrix. We develop a computationally efficient pseudospectral technique to simulate a Langevin dynamics derived from the block copolymer field-theoretic Hamiltonian and demonstrate two different processes by which nanoparticles may be incorporated into this framework.

In support of the peptide-polymer work, we develop a new algorithm for Metropolis Monte Carlo simulations on high-performance graphics processing units (GPUs) that relies on the local equilibration of non-interacting regions of a lattice system. We show how the technique can better exploit the GPU memory hierarchy resulting in over 100-fold speedups. This technique is well-suited for lattice systems with couplings beyond nearest neighbors and systems with complicated local or dynamical constraints.

Finally, we investigate the implications for excluding volume in a $\phi^4 - \phi^2$ field theory, representative of a block copolymer theory. We provide a simplified derivation for the response of a Gaussian liquid to a volume-excluding solute. We apply this technique to a discrete $\phi^4 - \phi^2$ theory and show that the effects of volume exclusion act independently of the ϕ^4 term. Finally, we show using Monte Carlo simulations that the stabilization provided

by the ϕ^4 term may be replaced by a hard constraint on density fluctuations that imparts stability, while making the model's connections to well-studied Gaussian models of liquids more transparent.

To my family.

Contents

Acknowledgments

First, I would like to thank my advisor, Professor Phillip Geissler, for his help and advice during my time at Berkeley. I deeply appreciate his patience and willingness to help guide me through research problems while giving me the freedom to formulate original questions and solutions. I am also very grateful to Steve Whitelam, whose mentorship was indispensible early in my graduate work. The advice, friendship and support of the entire the Geissler lab made this journey a much more enjoyable one. In particular, I'd like to thank Evan Wang, Anna Schneider, Chris Ryan, Sucheol Shin, Michael Grunwald, Asaph Widmer-Cooper, JiYeon Ku, Todd Gingrich, Laura Armstrong and Will Browne. In addition, David Sivak was a great officemate and brainstorming partner. Jon Forrest's infectious excitement about computers and Kate Chase and Lisa Littlejohn's constant support were essential in navigating through Berkeley.

My family deserves an incredible amount of thanks, especially my parents Ken and Connie. The time and love they invested in my life is a prerequisite to everything in this thesis. Finally, I owe a huge debt of gratitude to Anna Labno. I could not have made it this far without her encouragement and understanding.

Chapter 1

Introduction

Fabrication of materials via the bottom-up, noncovalent self-assembly of constituent nanometer-scale particles has the potential to revolutionize many areas of science and engineering, from more efficient semiconductors to novel optical and plasmonic devices. A mix-and-wait approach to self assembly, in which the nanoparticles are added to a medium and allowed to assemble, is attractive due to its simplicity and time and resource efficiency, but beyond the design of the nanoparticles and the medium, this approach leaves assembly largely up to the vagaries of chance. Without a thorough understanding of how random mixtures of particles spontaneously assemble into extended superstructures, there is little hope for controlling the spatial and orientational distribution of nanoparticles. Although limited progress has been made in designing synthetic experimental systems that successfully self-assemble over meaningful length-scales at a meaningful level of complexity, living systems clearly demonstrate that reliable self organization can be achieved over nearly any length scale, and can be utilized for a stunning variety of purpose. But beyond biomimicry, it is still unclear what general principles and concepts are necessary to design de-novo systems that self assemble. Straightforwardly varying parameters of a model system is akin to searching for a needle in a high-dimensional haystack and typically results in disordered or defect-laden structures. On the other hand, theory and computation provide a useful route to studying self-assembly due to their ability to study the kinetic and thermodynamic features of simplified systems at a level of detail not possible in experiments. Since interactions are non-covalent, pairwise interaction energies in these systems are usually on the order of kT, and statistical physics provides an ideal framework to describe the equilibrium and non-equilibrium behavior of these systems. Thus, such an approach has the potential to elucidate mechanisms responsible for defect-free assembly in simplified systems that could serve as foundational principles to designing successful experimental systems.

In this thesis we explore the self-assembly of two bio-inspired systems, and we extend various techincal issues related to their study. In Chapter 2, we study a system of archaeal chaperonin proteins, which are known to self-assemble into micron-scale structures when purified in-vitro. Two remarkably different structures form in a largely binary fashion: either bundles of filamentous chains, or layers of hexagonally packed sheets. Both structures

are being investigated as templates for nanoscale devices. Using a newly developed quasi-dynamical Monte Carlo simulation technique, we reproduce the assembled structures and show that the binary decision for sheet or string formation can be explained by allowing for conformational changes between a sheet-favoring state and a string-favoring state, and that the energy gap for this conformational change controls which structure will eventually form. Further, we show classical nucleation theory can explain this assembly as a competition to form structures with differing nucleation barriers. Finally, the dynamical aspects of the Monte Carlo technique allow us to illustrate differences in the structural features of equilibrium ground states and kinetically accessible assemblies.

In Chapter 3, we examine a system of self assembling cyclic peptides inside a block copolymer matrix. Most studies to date have focused on the self assembly properties of nanoparticle systems in a structureless solvent, whose only effect is to introduce a random 'buffeting' force on the nanoparticles. The system studied in this chapter takes its inspiration from materials science, where an increasingly popular trend has been to employ structured solvents to aid in nanoscale assembly of materials. Block copolymers have been a solvent of choice due to their ability to undergo mesoscopic phase separation on 1-100 nm length scales, into a variety of self assembled phases. The basic theory of block copolymers is well-known, and a variety of computational techniques are available to investigate complicated polymer architectures, where closed-form solutions are not known. The situation is less clear for nanocomposites, since the field-theoretic description of the polymer matrix must be topologically constrained to the regions outside of the (assumed) hard-core nanoparticles. We delay a detailed discussion of the theory unil later chapters and instead assume that a sufficient condition on the block copolymer is that compositional fluctuations vanish inside hard-core nanoparticles. Under this assumption we develop an pseudospectral approach to simulating nanocomposities and investigate nanoparticle self assembly with this approach. We find that a major obstacle to this simulation approach is the enormous interfacial penalties associated with excluding volume for nanoparticles. Such findings indicate future work should be directed at developing efficient constraint algorithms.

In Chapter 4 we develop a method for using Metropolis Monte Carlo lattice simulations on graphics processing units (GPUs). A number of authors have previously adapted parallel checkerboard lattice techniques to simulating the Ising model on the GPU [123, 94]. This approach is tractable because the Ising model contains only nearest-neighbor couplings in its Hamiltonian, meaning the decomposition produces two separate lattices that are independently updated. The block copolymer Hamiltonian contains higher-order gradient terms that introduce couplings beyond nearest neighbors, making straightforward lattice decomposition techniques neither feasible to program, nor even remotely efficient at utilizing the GPU memory hierarchy that frequently is a bottleneck in simulations. We develop an approach that first partitions the lattice into non-interacting subregions, which are sent to individual multiprocessors on the GPU, where a checkerboard decomposition is able to run quickly since data has been cached in fast, on-chip memory. We demonstrate that our approach is able to achieve above a 100-fold increase in the number of Monte Carlo steps per second and provide evidence that the rate of configurational decorrelation is even greater.

Finally, in Chapter 5, we develop a theory for nanocomposite systems. This requires two modifications to the well known derivation for the block copolymer Hamiltonian: first a volume is excluded for each nanoparticle and then polymer-particle interactions are turned on. The polymer-particle interactions may be effectively treated at a coarse-grained level similar to Flory Huggins theory for polymer melts by adding an interaction term to the Hamiltonian. Excluding volume is much more difficult and amounts to a restricted functional integration over both A-type and B-type density fields such that both densities vanish within nanoparticle hard cores. A derivation analogous to the original block copolymer Hamiltonian derivation would apply the random phase approximation to the intractable integral formulas, using a system of non-interacting chains in a volume-excluded system as a reference state. Instead, we make an analogy to simple liquids, where the effect of volume exclusion is well known in the context of Gaussian field theories, by showing that the nonlinear terms in the block copolymer Hamiltonian, traditionally truncated and replaced by a Landau-like quartic term, essentially amount to a sharp restoring force that constrains nonphysical fluctuations in the order parameter. With this shown, we approximate the system as a GFT with constraints placed on the real-space compositional fluctuations and demonstrate the validity of this approach with Monte Carlo simulations.

Chapter 2

Impact of Conformational Fluctuations on Self-Assembly Exemplified by Self Assembly of Group II Chaperonin Complexes

Chaperonins are high molecular mass double-ring structures composed of 60-kDa protein subunits [124]. Group II chaperonin complexes from the hyperthermophilic archaeon *Sulfolobus*, called rosettasomes, spontaneously assemble *in vitro* into micron-scale structures [90, 81]. Two remarkably different structures can be formed by the self-assembled, roughly spherical rosettasomes: bundles of filamentous chains bound pole-to-pole, and hexagonally-packed sheets of units bound equator-to-equator. The prospect of switching between the assembly of different structures by controlling only environmental conditions offers great promise for materials science. Dual assembly poses new fundamental questions regarding the statistical dynamics of organized growth: How do multiple modes of assembly cooperate or interfere? Are hybrid structures an unavoidable endpoint when forces favoring different structural motifs combine? Conceptually, these are general questions that may be asked for many complicated, self-assembling systems. In later chapters, we address these same questions for other bio-inspired systems, whereas in this chapter, we attempt to answer these questions by detailed computational exploration of group II chaperonin assembly.

2.1 Importance of Group II chaperonin complexes and their assembly

The Role of Chaperonins in *Sulfolobus shibatae*

Sulfolobus shibatae is a hyperthermophilic archaeon that was first identified living in acidic geothermal hot springs. When *S. shibatae* is subjected to higher than normal temperatures

it activates a heatshock response. The major component of the heatshock response is a formation of a large ring structure called archaeosome or rosettasome, which is composed of two different subunits ($TF55\alpha$ and TF55β), bearing resemblance to to the prokaryotic GroEL [98]. Unlike their bacterial counterparts, however, *S. shibatae* chaperonins are not able to fold more than a few specific proteins in vitro [40, 118]. This may be due to technical problems associated with in vitro experiments or specialization of these chaperonins for folding specific proteins. Alternatively, it suggests that chaperonins may have other functions *in vivo*.

The two chaperonin proteins represent 4% of *S. shibatae* total protein content and have a combined intracellular concentration of >30 mg/ml which is notable because at concentrations 0.5 mg/ml purified chaperonins form filaments which occurs at physiological relevant conditions. These observations suggest that chaperonin filaments may exist *in vivo* and form an extensive cytostructure. For example, Trent and coworkers observed filamentous structures in unfixed, uranyl-acetate-stained *S. shibatae* cells, which resemble the chaperonin filaments in size and appearance. ImmunoGold labeling using chaperonin antibodies indicated that many chaperonins are associated with insoluble cellular structures and these structures appear to be filamentous in some areas. Moreover, in *in vitro* experiments they observed that as the concentration of purified chaperonins is approaching the concentration found in cells, the isolated double-ring structures assembled into ordered filaments.

The formation of chaperonin filaments in cells has profound implications on their function as it will affect their role in both protein folding and as a cytoskeletal unit. In protein folding, chaperonin filaments may serve a regulatory function. It is believed that protein folding happens inside the chaperonin's cavity [32, 126] so if chaperonins form filaments access to the active cavity will be blocked for majority of non-terminal units. This can be used by *S. shibatae* as a mechanism to sequester chaperonins and limit their activities. Analogously, filaments can dissociate to free the individual chaperonin rings to increase their protein folding ability rapidly, much faster than if additional protein synthesis were required. This may allow the arcea to respond rapidly to changes in environmental conditions. Additionally it is possible that chaperonin filaments themselves form a cytoplasmic structure in Archaea that is functionally similar to the cytoskeleton in eukaryotes. Sulfolobus, similarly to other Archaeas, lack rigid cell walls, and such an internal cytoskeleton could be used to maintain and change their shape in solution [49, 113]. Moreover TCP1, a related eukaryotic cytoplasmic protein, is associated with the eukaryotic cytoskeleton [124].

The Role of Chaperonins in Nanofabricated Active Materials

Both filamentous- and planar-assembled chaperonin materials separately display great promise for nanotechnology. The controlled organization of materials into multi-dimensional structured arrays is a goal of many nano fabrication efforts and could potentially impact the engineering of many nano devices including semiconductor materials used in logic and memory units, as well as photonic bandgap (PBG) crystals used in nonlinear optical and sensing devices [136, 144]. Many of these devices are currently fabricated using lithographic patterning processes, which offer great control but are very expensive, time consuming and can

only be used for fabrication of devices smaller than mm. Moreover for devices with features smaller than 100 nm, ion and electron beam lithography becomes prohibitively expensive and time consuming, and more importantly, at these scales quantum effects fundamentally change the properties of devices. [110].

Arrays of nanoparticles formed by top-down self assembly methods are being explored for use as viable alternatives to standard lithographically patterned devices. Two-dimensional (2D) arrays of QDs with nanoscale resolution have a potential to form the basis of future generations of electronic and photonic devices [77, 78]. Biomolecules, such as chaperonins, are capable of self-assembling into a wide diversity of structures with nanoscale architecture. Proteins in particular can form intricate structures that can be readily manipulated and functionalized because their synthesis is genetically directed. The utility of chaperonins as scaffolds depends on their structure and their ability to self-assemble into double-rings and higher-order structures, such as filaments and two-dimensional arrays. Moreover the subunits can be radically engineered to affect their hierarchical self-assembly into rings, filaments and two dimensional arrays without compromising their structure-forming ability. They can be engineered by metallization of single chains of the D1 mutant to make nanowires by coating them with nickel/palladium [72]. Modifiyng β subunits by adding chemically reactive sites can guide the subunits to assemble into 2D crystals and guide the assembly of preformed metal and semiconductor nanoparticle QDs into ordered arrays [81]. Developments in nano science require alternative fabrication methods and new insights into the behaviour of materials on nanometer scales will be crucial.

2.2 Previous Experimental and Theoretical Approaches

Emulating the complex self-assembly properties of rosettasomes with synthetic particles requires interactions of a 'patchy' nature. It may seem sufficient to situate a strongly sticky region at each particle's equator, and a distinct kind of stickiness at its poles. Several computer simulation studies [51, 142, 55, 47, 130, 133, 9, 88, 104] and a few ambitious experiments have adopted an analogous perspective, attempting to design specific aggregate geometries through anisotropy of microscopic forces. For example Glotzer and coworkers performed molecular simulations to study the self-assembly of nanoparticles functionalized with oligomeric tethers attached to specific locations on the nanoparticle surface and observed formation of structures beyond what may be predicted using concepts from block copolymer microphase separation and liquid-crystal phase ordering [51, 142]. Systematic study of the design of simple patchy sphere models that reversibly self-assemble into monodisperse icosahedral clusters revealed that the optimal patch width is a compromise between structural specificity and kinetic accessibility [133] and that patchiness has a strong effect on the phase diagram [9]. Such patchy particles then self-assemble into a diamond structure from an initially disordered state [141], which is of particular interest in nanotechnology due to

its similarity with a photonics crystal structure. However, component design based solely on structural considerations rarely results in successful assembly. The picture emerging from these studies is one of intense competition between thermodynamics and kinetics: even for simple patchy nanoparticles, stochastic growth dynamics and the difficulties of defect annealing condition successful assembly on the fine-tuning of nanoparticle attraction strengths and binding geometry specificities.

TEM imaging of single *S. shibatae* cells stained with uranyl acetate showed some cells in which in lightly stained areas filamentous structures were seen at both low and high magnifications. The distribution and interweaving of these filaments suggest that they extend throughout the cytoplasm and stereomicrographs confirmed that they were distributed throughout the cell volume. The width of the intracellular filaments was typically $\approx 11 \pm 1$ nm for individual filaments, which was very similar to the width of the chaperonin filaments formed in vitro ($\approx 10.7 \pm 0.6$ nm) [124]. Additional experimental self assembly of rosettasomes revealed that beta homo-oligomeric rosettasomes and all hetero-oligomeric rosettasomes typically associate into filaments and *in vivo*, the rosettasome structure is determined by the relative abundance of subunits and not by a fixed geometry [59].

We want to draw attention to the striking qualitative observation from experiments, which suggests rosettasome aggregation to be a binary rather than a hybrid process: structures intermediate in character between bundled strings and few-layered sheets are not seen in micrographs. Together with the observation that planar and filamentous structures sometimes appear in the same sample, this fact suggests a precipitous transition in aggregate geometry, from string-like to sheet-like, as control parameters (temperature, concentration, protein amino acid sequence, etc.) are changed. By contrast, our simulations of collections of patchy nanoparticles give rise to products that display a gradual transformation with model parameters from string-like to sheet-like character. We resolve this discrepancy by considering in addition a conformational flexibility of assembling units. The kinetic effects of this simple local detail are global in both space and time. Although our examination of this cooperative dynamical phenomenon is inspired by the behavior of a specific protein complex, its implications are quite general: endowing nanoscopic self-assembling components with a simple degree of conformational flexibility permits modes of assembly unattainable to components of fixed conformation [31].

2.3 Methods

In this paper we use theoretical models based on these physical perspectives as well as computer simulations to explain the dual nature of rosettasome assembly.

Theory

The precise interactions between two rosettasome units, each comprising about 10^5 atoms, are enormously complicated. Detailed molecular dynamics simulations might accurately re-

produce these solvent-mediated forces, but such an approach could scarcely be applied over time scales of milliseconds or longer that give rise to the higher-order structures considered here. Therefore, we treat free energies of interaction in a schematic way, considering for each pair of units only the dependence on a handful of geometric parameters, as sketched in Figure 2.1. Our effective potentials are constructed with attention to a few key physical facts and ideas: (1) each unit excludes a roughly spherical volume of radius $a = 8.5$ nm; (2) attractive interactions extend over distances small compared to a and vary weakly with rotation of a unit about its axis. While rosettasomes are 9-fold symmetric, they are observed to pack in hexagonal lattices incommensurate with such symmetry so we assume that equatorial interactions vary weakly with longitude and (3) potentials of mean force stabilize bindings pole-to-pole and equator-to-equator, as indicated by high-resolution microscopy [90].

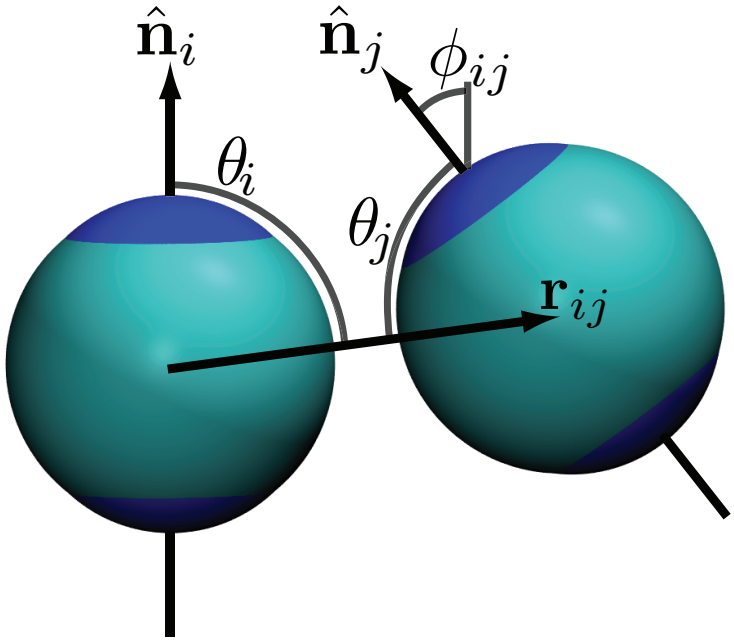


Figure 2.1: Geometry for our model of associating rosettasomes. Each rosettasome is represented here as a spherical object. Angular variables are related to orientation vectors by $\hat{\mathbf{n}}_i \cdot \hat{\mathbf{n}}_j = \cos\phi_{ij}$ and $\hat{\mathbf{n}}_i \cdot \frac{\vec{r}_{ij}}{|\vec{r}_{ij}|} = \cos\theta_i$.

In detail, our model consists of N non-overlapping spheres of radius a, each representing a single rosettasome unit. The attraction u_{pair} between units i and j possesses polar and equatorial contributions, both depending on the inter-unit separation vector $\mathbf{r}_{ij}$ and unit orientation vectors $\hat{\mathbf{n}}_i$ and $\hat{\mathbf{n}}_j$:

$$u_{\text{pair}}(\mathbf{r}_{ij}, \hat{\mathbf{n}}_i, \hat{\mathbf{n}}_j) = -\epsilon_{\text{eq}} v_{\text{eq}}(\mathbf{r}_{ij}, \hat{\mathbf{n}}_i, \hat{\mathbf{n}}_j) \\ -\epsilon_{\text{pol}} v_{\text{pol}}(\mathbf{r}_{ij}, \hat{\mathbf{n}}_i, \hat{n}_j). \tag{2.1}$$

The strengths ϵ_{pol} and ϵ_{eq} of these two interactions, relative to that of typical thermal excitations $k_{\text{B}}T$, constitute the most important control parameters of our study. In simulations

that permit units to change conformation, these parameters become in effect fluctuating variables.

The short range of rosettasomes' interactions relative to their size motivates our simplified choice of the potential's distance dependence: the model functions v_{pol} and v_{eq} vanish for separations greater than $2a + a/4$, yielding an attraction range $a/4 \approx 2$ nm consistent with Debye screening at experimental ionic strengths. Below this cutoff distance these functions vary only with particle orientations, and do so according to the equations

$$v_{eq} = \hat{C}_1(\phi_{ij}; \sigma_{align}) C_0(\theta_i; \sigma_{eq}) C_0(\theta_j; \sigma_{eq}); \tag{2.2}$$

$$v_{pol} = \hat{C}_1(\phi_{ij}; \sigma_{align}) \hat{C}_1(\theta_i; \sigma_{pol}) \hat{C}_1(\theta_j; \sigma_{pol}). \tag{2.3}$$

Equations (1.2) and (1.3) describe in mathematical terms the patchiness sketched in Figure 2.1. The effective size of polar and equatorial patches is determined by inverse specificities σ_{align}, σ_{eq} and σ_{pol}; these parameterize a function $C_\alpha(\psi; \sigma) \equiv e^{-(\cos\psi - \alpha)^2/\sigma^2}$ that rewards the alignment of angles ψ and $\cos^{-1}\alpha$ within a tolerance σ. The function $\hat{C}_\alpha(\psi; \sigma) \equiv C_\alpha(\psi; \sigma) + C_{-\alpha}(\psi; \sigma)$ is made symmetric to preserve the up-down equivalence of rosettasome units suggested by crystal structures of homologous proteins. We set $\sigma_{align} = \sigma_{eq} = 0.3$ and $\sigma_p = 0.12$.

Computational Implementation

We study long trajectories of this model system, corresponding to roughly tens of milliseconds of evolution for an experimental system. Dynamics were propagated using a virtual-move Monte Carlo procedure [130] designed to mimic the aggregation kinetics of strongly associating nanoparticles. We drew translation magnitudes from a uniform distribution with maximum $0.5a$, and drew rotation angles from a uniform distribution with maximum 22.5°. We scaled collective translation acceptance rates by the reciprocal of an approximate hydrodynamic radius of the moving cluster [130], and scaled acceptance rates for rotations by the cube of the same factor. We initiated each simulation by selecting particle orientations and positions at random (while respecting steric constraints) within a cubic box of length L. Periodic boundary conditions were imposed in imitation of bulk surroundings.

Model Parameters

We set an upper bound on the angular tolerance (i.e. inverse specificity) of the polar interaction in our simulations, σ_{pol}, by noting that bundled chaperonin strings in experiments appear to grow in register, via the addition of monomers to each string tip, rather than via the recruitment of monomers shared between neighboring strings. In simulations, the angular component of the polar energy of interaction between two aligned model chaperonins [i.e., $\phi = 0$, see Figure 2.1] is (Equation (1.4))

$$v_{pol}(\theta) = -\epsilon_{pol} \exp\left(-2\sigma_{pol}^{-2}[\cos\theta - 1]^2\right), \tag{2.4}$$

where θ is the (small) angle between the orientation directors of the units and the vector joining their centers. To set an upper bound on σ_{pol} we require that in simulations a bundle of two perfectly-aligned strings finds it energetically favorable to recruit monomers preferentially to the tips of each of its two filaments, rather than by sharing monomers between filaments. We therefore require that $v_{pol}(0) < 2v_{pol}(\pi/6)$, or

$$\sigma_{pol}^2 < \frac{2}{\ln 2}\left(\cos(\pi/6) - 1\right)^2, \tag{2.5}$$

giving $\sigma_{pol}<0.23$. We choose $\sigma_{pol} = 0.12$ to strongly disfavor off-register growth of bundled strings.

Given the specificity of the polar interaction we can put an approximate lower bound on its strength ϵ_{pol} by requiring that in thermal equilibrium we observe chains of (at least) hundreds of units (assuming a sufficiently large system). Chaperonin strings of many hundreds of units are observed in experiment [124]. Strings are quasi-one-dimensional structures (each particle in the bulk binds to only two neighbors) and so a numerically large binding energy is required to offset the unfavorable entropy of polymerization that results from the growth of long strings. We can estimate how large this binding energy must be using the Wertheim thermodynamic perturbation theory [129] outlined in Ref. [112]. This theory provides an expression for the free energy of associating particles possessing two identical binding sites (such as an identically sticky north and south pole, as here), and so may be used to calculate the extent of string formation for a collection of model chaperonins. From the results of Ref. [112] we find that chaperonin string-formers of number density $\rho = N/V$ form strings whose equilibrium length distribution $\rho(l)$ is

$$\rho(l) = \rho\ell_{eq}^{-2}\left(1 - \ell_{eq}^{-1}\right)^{l-1}. \tag{2.6}$$

Here ℓ_{eq} is the average string length in equilibrium, and satisfies

$$2\ell_{eq} = 1 + \sqrt{1 + 8\rho\Delta}. \tag{2.7}$$

The parameter Δ is related to the second virial coefficient of our model system. It is calculated by integrating the Mayer f function for two chaperonin units over all angles and distances for which they interact:

$$\Delta = 4\pi \int d\boldsymbol{\omega} dr\, r^2\, g_{HS}(r)\langle e^{-\beta U(\boldsymbol{r})} - 1\rangle_{\boldsymbol{\omega}}. \tag{2.8}$$

In this expression $\boldsymbol{r}$ is the vector joining the two units, and r is its magnitude; $\boldsymbol{\omega}$ represents two sets of angles (we assume one chaperonin to be fixed at the origin with orientation along $\hat{\boldsymbol{z}}$); $g_{HS}(r)$ is a (reference) pair correlation function for hard spheres (for this we use the approximation of Ref. [112]); $\langle\cdot\rangle_{\boldsymbol{\omega}}$ represents an average over the angles $\boldsymbol{\omega}$; and U is the energy of interaction of the two units. We plot in Figure 2.2 the average string length in equilibrium (for two chaperonin concentrations) and the string length distribution (for two

attraction strengths), assuming $\sigma_{pol} = 0.12$. We find that ϵ_{pol} must exceed roughly 15 k_BT (to the nearest 5 k_BT) in order that chains formed in simulations may approach hundreds of units in length. Therefore, we take $\epsilon_{pol} = 16\,k_BT$ for the polymerization interaction of string-forming units.

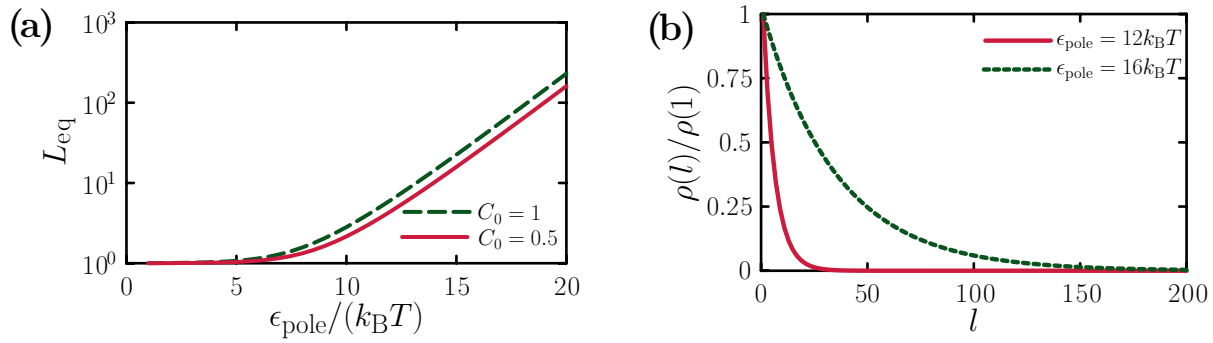


Figure 2.2: Predictions of Wertheim thermodynamic perturbation theory [129, 112] applied to a collection of our model string-forming units (with inverse specificity parameter $\sigma_{pol} = 0.12$ and no equatorial coupling). (a) Average string length ℓ_{eq} in equilibrium as a function of ϵ_{pol} for two concentrations; C_0 is the concentration corresponding to 1500 units in a cubic box of length $64a$, as considered in the main text. (b) String length l distribution $\rho(l)$ for two binding strengths. Our conclusion from these estimates is that ϵ_{pol} must exceed roughly 15 k_BT in order that the longer strings in simulations approach hundreds of units in length (as is seen in experiment).

With polar strength and specificity determined by equilibrium considerations, we find that nucleation and growth of filaments proceeds rapidly. Sheet formation, by contrast, is relatively sluggish: the requirement that self-assembly of sheets proceed without the development of substantial kinetic traps imposes stringent limits on the strength of equatorial interactions (given their specificity) [130]. Based on our previous work we choose for sheet-forming units the parameters $\sigma_{eq} = 0.3$ and $\epsilon_{eq} = 6.5k_BT$ to ensure that planons self-assemble into sheets with a low density of defects even at the largest concentrations we consider.

2.4 Results

We report below on long trajectories of this model system, corresponding to roughly tens of milliseconds of evolution for an experimental system to study both kinetic and thermodynamic aspects of assembly. As mentioned before we encounter discrepancy between initial theoretical results and experimental observation which we resolve by considering in addition a conformational flexibility of assembling units. The kinetic effects of this simple local detail are global in both space and time as shown below.

Kinetic Phase Diagram

We have explored the assembly dynamics of units with fixed conformation as a function of ϵ_{eq} and ϵ_{pol}. For many sets of values for these parameters, we ran trajectories of $N = 800$ units in a box of length $60a$. Units therefore occupied 1.5% of the box volume, corresponding to a protein concentration of 10.9 mg/ml, typical of experimental samples (1−20 mg/ml). Results of these numerous trajectories are summarized by the kinetic phase diagram in Fig. 2.3, which reports on both the occurrence of assembly (judged by the presence of clusters comprising 20 or more units; the 'no assembly' cutoff of size 20 means that we ignore assembly in regions of the phase diagram where short filaments form, but this neglect does not change in an important way the trends we identify) and on resulting geometries.

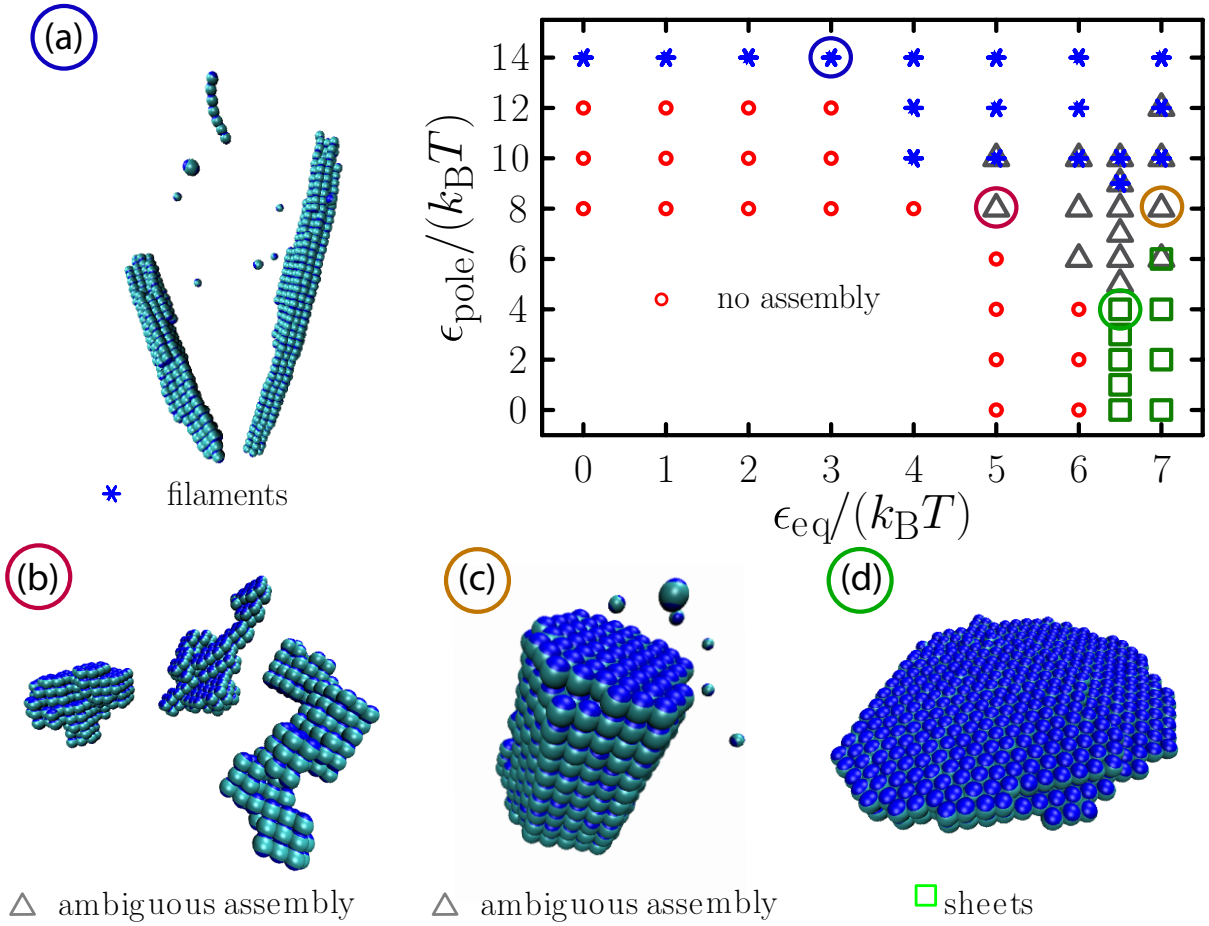


Figure 2.3: Kinetic phase diagram for units of fixed conformation. We identify regions in which assembled structures are string-like, e.g. (a), sheet-like, e.g. (d), or of ambiguous character, e.g. (c). Kinetically-frustrated structures, e.g. (b), are found with increasing abundance as we increase the potency of the equatorial coupling.

In order to quantify the level of assembly and the morphology of each aggregate we defined an order parameter:

$$S \equiv \frac{1}{6}(6 - n_{eq}) - \frac{1}{2}(2 - n_{pol}) \tag{2.9}$$

. For a given cluster above the cutoff size of 20 units, n_{eq} estimates the number of equatorial neighbors per particle (computed by summing v_{eq} over all interacting pairs and dividing by the number of particles); n_{pol} is calculated in an analogous manner. S takes the value $+1$ for an infinitely long filament, $S = -1$ for an infinitely extended close-packed sheet, and intermediate values for structures of finite size and for geometric character intermediate between sheets and strings. Any aggregate having $S > 1/4$ is identified as a 'string', while any aggregate for which $S < -1/4$ is classed as a 'sheet'. We consider aggregates satisfying $-1/4 \leq S \leq 1/4$ to be of ambiguous character.

Kinetic accessibility is a crucial issue in the assembly of sheet-like structures. Although sheets are thermodynamically stable under many of the conditions we have considered, the range of ϵ_{eq} in which planar aggregates grow in an ordered fashion is small. The region $\epsilon_{eq} \geq 7k_BT$ is dominated by kinetic frustration, while the region $\epsilon_{eq} = 6k_BT$, from $\epsilon_{pol} = 0$ to $\epsilon_{pol} = 4k_BT$ is supercooled: sheets are thermodynamically stable but nucleate too slowly to appear in the simulation times considered.

The kinetic phase diagram and illustrated configurations indicate a propensity for hybrid assembly when both equatorial and polar attractions are appreciable. Under no conditions do structures of distinct geometry coexist at the same thermodynamic state. Instead, we observe a continuous variation with model parameters from sheet- to string-like assembly. At some values of (ϵ_{eq},ϵ_{pol}), the two modes of assembly clearly hinder one another, either by sequestering material or by leading the ordered growth of nascent clusters astray. At other parameter values these modes cooperate, generating well-ordered hybrid structures that interpolate smoothly between layered sheets and bundled strings.

Kinetics versus thermodynamics for fixed-conformation self-assembly

Since large regions of the phase diagram of Figure 2.3 are dominated by nonequilibrium effects we investigated the contrast between the products of aggregation *dynamics* seen in Figure 2.3 with configurations typical of true thermal equilibrium (which might be reached dynamically on time scales too long to witness even in the laboratory) Figure 2.4. Equilibrium structures were obtained using a Monte Carlo algorithm that samples the Boltzmann distribution much more efficiently than does the system's natural dynamics [27]. In these simulations a nonlocal move set dramatically accelerated the annealing of defects [58, 104] by shuttling particles between the interaction regions of other particles. Those nonlocal moves [27] reveal that close-packed sheets are the thermodynamically preferred state and that kinetically frustrated aggregates at state $(\epsilon_{eq}, \epsilon_{pol}) = (7, 8)\, k_BT$ observed previously are replaced by structures with few defects such as highly ordered sheets These examples indicate the considerable

difference between what is kinetically accessible and what is thermodynamically preferred in this example of self-assembly.

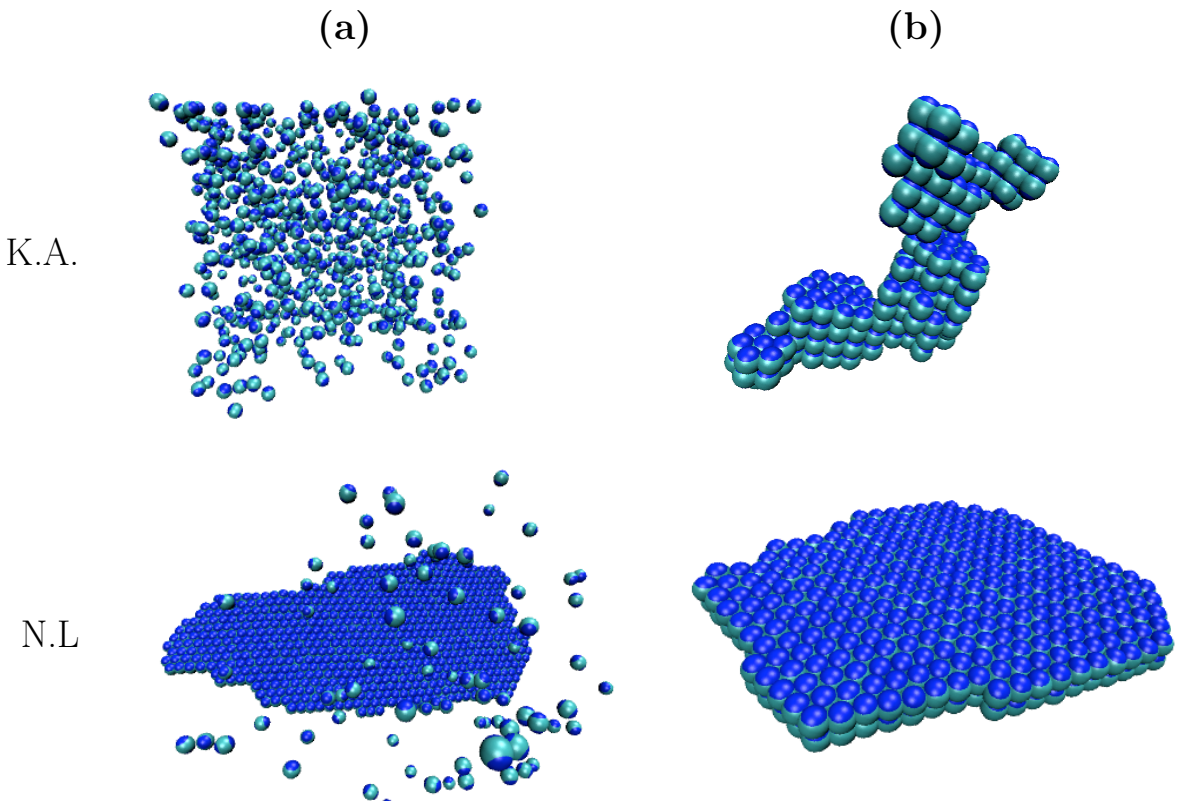


Figure 2.4: Large regions of the phase diagram of Figure 2 (main text) are dominated by nonequilibrium effects. (a) At state $(\epsilon_{eq}, \epsilon_{pol}) = (6k_BT, 0k_BT)$, kinetically accessible (K.A.) configurations, i.e., those generated by an algorithm designed to mimic a natural dynamics over experimental time scales, primarily comprise unassociated units. However, nonlocal (N.L.) moves [27] reveal that close-packed sheets are the thermodynamically preferred state. (b) Kinetically frustrated aggregates at state $(\epsilon_{eq}, \epsilon_{pol}) = (7, 8)\, k_BT$ are replaced by structures with few defects and with a very different form under nonlocal moves.

The Importance of The Conformational Flexibility of the Rosettasome Units

This smooth evolution is expressly *not* what is observed in experiments. Instead, laboratory measurements suggest a very sharp crossover from sheets to strings with changing environmental conditions. The absence of an intermediate regime of hybrid assembly is highlighted by the observation of behavior akin to phase coexistence. Accord between experiments and

our patchy nanoparticle model is unlikely to be achieved by revising our choices of interaction strengths and specificities: we have performed many additional simulations in which we varied component concentrations, system sizes, binding tolerances and energy scales, and failed to observe sheet-string coexistence in all cases.

The key question arises: why does the model we have presented fail to produce coexisting superstructure types? We cannot rule out the possibility that fine details of intermolecular forces are key to capturing this emergent behavior. We argue, however, that the qualitative contrast between binary assembly reminiscent of a phase transition and hybrid aggregation that varies smoothly with model parameters points to the neglect of a more fundamental feature. We propose that this missing ingredient is the conformational flexibility of each rosettasome unit.

Previous work has established as many as three different conformations for the *S.* Shibatae rosettasome [99, 111], which are summarized in the table below.

Simulation of Patchy Particles with Different Conformations

Motivated by these studies and by our findings above, we have augmented our model of patchy nanoparticles by allowing each particle the freedom to change between two different conformations. One of these conformations is biased towards forming sheet-like geometries. When particles adopt this state, we refer to them as *p*lanons. Particles assuming the alternate conformation, which we dub *s*tringons, have a preference for filamentous self-assembly. We will see that the ability of particles to fluctuate between these identities indeed greatly sharpens the transition between assembled structures and can effect coexistence of sheets and filaments.

These preferences conferred by internal degrees of freedom appear in our model through variations in binding energies. In all the calculations we will describe, the equatorial coupling strength ϵ_{eq} between two adjacent planons is set to $6.5\,k_{\mathrm{B}}T$, sufficient to ensure that a collection of particles constrained to their planon states can form large-scale ordered sheets. To ensure the possibility of filamentous growth we choose the strength ϵ_{pol} of polar interactions between two adjacent stringons to be $16\,k_{\mathrm{B}}T$. We also assign *minor* couplings of varying strengths: these consist of polar interactions between planons; equatorial interactions between stringons; and inter-species interactions. These are summarised in Table 1.

In the course of model assembly dynamics we permit monomer conformations to fluctuate at random, subject to biases from interaction energies and from an intrinsic free energy difference between planon and stringon states. Each unit possesses an internal energy $u_{internal}$, which we set to $-k_{\mathrm{B}}T\ln\delta$ for stringons and 0 for planons. The parameter δ thus represents the ratio of stringons to planons in a sample of noninteracting rosettasome units. Just as internal states bias propensity for structure formation, however, so aggregation can bias conformational populations significantly away from ideal statistics.

Algorithmically, we attempted a change in each unit's conformation on average once

interacting species	ϵ_{eq}	ϵ_{pol}
planon-planon	$6.5\,k_{\mathrm{B}}T$	$\epsilon_{\mathrm{pol}}^{(min)}$
stringon-planon	$\epsilon_{\mathrm{eq}}^{(min)}$	$\epsilon_{\mathrm{pol}}^{(min)}$
stringon-stringon	$\epsilon_{\mathrm{eq}}^{(min)}$	$16\,k_{\mathrm{B}}T$

Table 2.1: Stringon and planon interaction strengths. We consider fixed major couplings and different sets of minor couplings. The specific case $\epsilon_{\mathrm{eq}}^{(min)} = \epsilon_{\mathrm{pol}}^{(min)} = 0$ corresponds to planons and stringons interacting only through hard core constraints. We have considered minor equatorial attractions as strong as $\epsilon_{\mathrm{eq}}^{(min)} = 4k_BT$ and minor polar attractions as strong as $5k_BT$. In the limit $\epsilon_{\mathrm{pol}}^{(min)} = 16k_BT$, $\epsilon_{\mathrm{eq}}^{(min)} = 6.5k_BT$, the distinction between planons and stringons vanishes.

every $M_{change} = 1000$ Monte Carlo sweeps. By accepting such changes with probability

$$p_{change} = \min\left[1, \exp\left(-\beta\Delta E - \beta\Delta u_{internal}\right)\right], \tag{2.10}$$

we established a basic rate $\omega_{change} = \omega_0(1+\delta)/M_{change} \approx 10^5\,\mathrm{s}^{-1}$ for conformational fluctuations of isolated units that is consistent with estimates for similar large-scale rearrangements in related proteins [57]. The results we present are qualitatively unchanged by varying M_{change} from 1 to 1000. This robustness suggests that assembly behavior is not strongly sensitive to changes in the rate of conformational fluctuations, provided it outpaces the rate of basic structure formation. Here ω_0^{-1} is the duration of a single MC sweep (involving on average one proposed move per particle). Based on the diffusivity expected from Stokes Law for a sphere of radius 8.5 nm in liquid water, we estimate that $\omega_0 \approx 10^8\,\mathrm{s}^{-1}$. In Eq. (2.10), ΔE and $\Delta u_{internal}$ are respectively the changes in interaction energy and internal energy resulting from a proposed conformation change.

We computed assembly trajectories for $N = 1500$ conformationally flexible particles in a box of length $64a$. By examining several values of δ and of the minor couplings, we varied the strength of interactions between planons and stringons, as well as the degree of distinction between them. Units' initial identities were randomly assigned as stringons with probability $\delta/(1+\delta)$, or alternatively as planons with probability $1/(1+\delta)$.

Trajectories with Vanishing Minor Couplings

We focus first on the case of vanishing minor couplings. Here, stringons and planons are orientationally ambivalent toward one another, interacting only through sterics but intimately linked through interconversion. In Figure 2.5 we show a time-ordered series of snapshots taken from a single trajectory featuring this extreme example of conformational change. Allowing unit identities to fluctuate can indeed yield at late times a coexistence of structures possessing sheet- and string-like symmetry. Interestingly, in order to observe such coexistence we must bias the internal energy of units strongly in favor of planons; here $\delta = 0.05$, corresponding to an internal bias of $3k_{\mathrm{B}}T$. The necessity for this imbalance originates in an

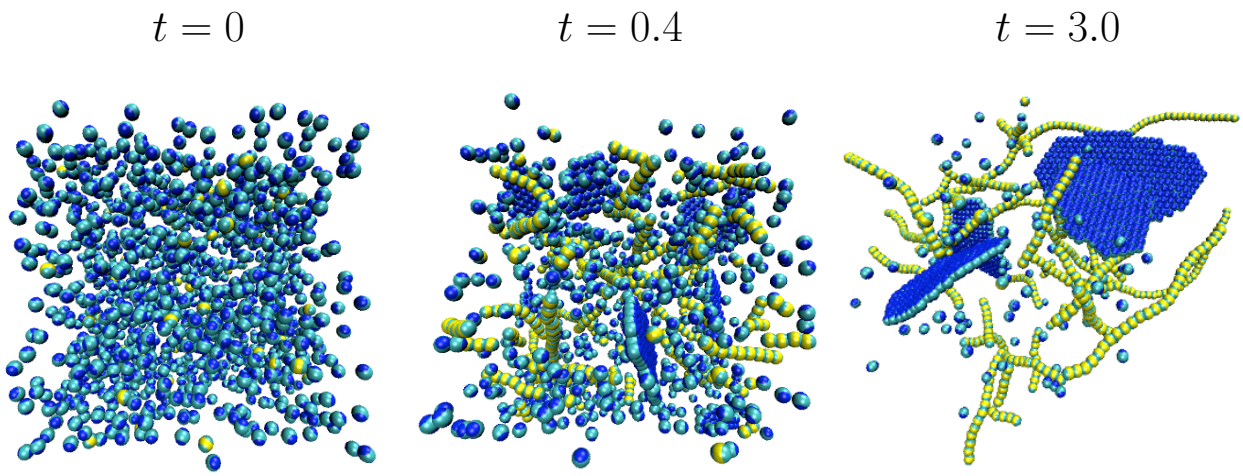


Figure 2.5: Time sequence of the assembly of conformation-changing units. Here, units of different types interact only through interconversion and sterics $\left(\epsilon_{eq}^{(min)} = 0 = \epsilon_{pol}^{(min)}\right)$. The internal energy bias parameter is $\delta = 0.05$, so that isolated planons (blue) are strongly favored thermodynamically over stringons (yellow). However, string formation is facile, leading to an increase in the number of stringon units with time. Parameters are such that we observe coexistence between large-scale assembled structures of the two distinct symmetries. Time is measured in millions of Monte Carlo moves per particle.

unambiguous physical mechanism: As one-dimensional structures, strings lack a nucleation barrier and thus form much more rapidly than sheets, which must await the appearance of a critical nucleus.

Thermodynamics of sheet formation

We have quantified the free energy barrier to sheet nucleation at various planon concentrations using umbrella sampling simulations. In Figure 2.6(a) we plot the free energy $G(n)$ of a small sheet as a function of the number n of its constituent planon units, for several concentrations of free planons with no polar couplings. Each of these curves, calculated using umbrella sampling techniques, features a barrier $\Delta G^{\ddagger}$ to nucleation. In Figure 2.6(b) we show the decay of this barrier height with increasing planon concentration.

At the highest density, corresponding to the initial state of an assembly trajectory, this barrier is little larger than energies typical of thermal fluctuations, so that sheet initiation is only weakly impeded. However, string formation quickly sequesters units, in effect diluting the pool of monomers available for sheet formation. Our umbrella sampling results show that removing 60% of the available units raises $\Delta G^{\ddagger}$ by several $k_{\mathrm{B}}T$, slowing nucleation more than 100-fold. This dynamic elevation of the nucleation barrier necessitates a large

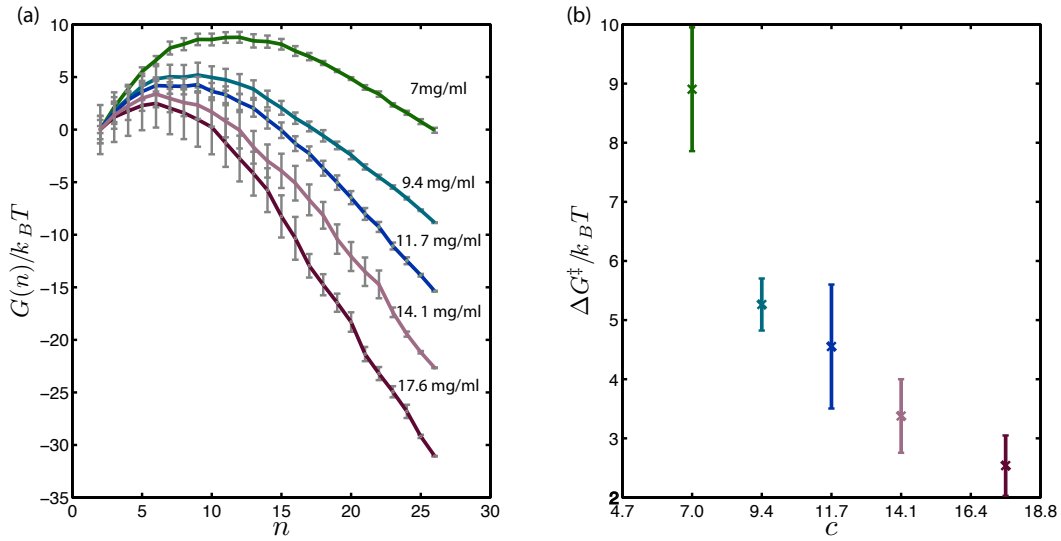


Figure 2.6: (a) Free energy of a system of planons that comprise a single sheet-like cluster of n units and $N-n$ unassociated units. Results are shown for several concentrations $c = \frac{N}{V}$, with volume $V = (64a)^3$: $c = 7$ mg/ml (red), $c = 9.4$ mg/ml (blue), $c = 11.7$ mg/ml (green), $c = 14.1$ mg/ml (black), and $c = 17.6$ mg/ml (pink). (b) Barrier height as a function of planon concentration c in mg/ml.

bias in internal energy in favor of planons (small δ) in order to obtain coexistence. These considerations underscore the fundamentally nonequilibrium nature of assembly within our model: when fully relaxed (at inaccessibly long times), this system would possess almost no strings or stringons.

Time evolution of schizophrenic assembly

In Figure 2.8 (a) we quantify the kinetics of dichotomous assembly by plotting as a function of time the fraction of planon units for a trajectory possessing the same parameter values as that shown in Figure 2.5. We plot also the number of polar and equatorial contacts per particle. The celerity of string growth with respect to sheet nucleation in general causes f to be noticeably diminished from its ideal value of $(1+\delta)^{-1}$. Steady increases in both n_{eq} and n_{pol} at intermediate times reflect the simultaneous growth of corresponding structures. In Figure 2.8 (b) we show the time dependence of f for trajectories obtained with several different values of the bias δ. String formation suppresses sheet nucleation almost completely for even modest δ.

Nonzero minor couplings facilitate the formation of structures of a composite character

Equipping units with nonzero minor couplings introduces inter-species attractions, which, as one might expect, facilitate the formation of structures of a composite character, namely stacked sheets and bundled filaments. The minor couplings also control inter-species energetic interactions, which influence the ease of unit interconversion. We find that the kinetic bias toward string formation identified previously is qualitatively unchanged, but interactions between species modify quantitatively the rate at which structures of the two species proliferate.

First we qualitatively investigate the effect of minor coupling on assembly. We show snapshots from long trajectories generated using nine different sets of parameters (representing a variety of thermodynamic states in the real system) in Figure 2.7. These states differ in the internal energies of conformational states (set by δ), and in their minor energetic couplings.

To quantitatively study the differences between the assembly kinetics of interconverting units with vanishing minor couplings and considerable minor couplings we compare (1) the time dependence of the fraction of planon units and the number of polar contacts per particle for the trajectory in Figure 2.7 and (2) time dependence of the fraction of planon units for single trajectories with different values of internal bias.

When units have vanishing minor couplings rapid string formation lowers the planon-to-stringon ratio so that both strings and sheets proliferate. In Figure 2.9 we present the analog of Figure 2.8 for units with considerable minor couplings. While the rapidity of planon to stringon conversion is influenced by the inter-species interactions, the fundamental kinetic bias in favor of one-dimensional stringons is not.

2.5 Composite structures can arise from planon-stringon interactions

Units' conformational flexibilities can induce interactions between superstructures of different types. The interconversion of units at the surfaces of aggregates can fuse different structures together, or promote nucleation of one kind of structure at the surface of another. We show an example product of sheet-string fusion in Figure 2.10. In contrast to hybrid structures that proliferate when unit conformations are fixed, different elements of these assemblies tend to maintain their essential shapes even when fused. When units are free to undergo conformational change in our simulations, the superstructures they form are qualitatively similar to those seen in experiments.

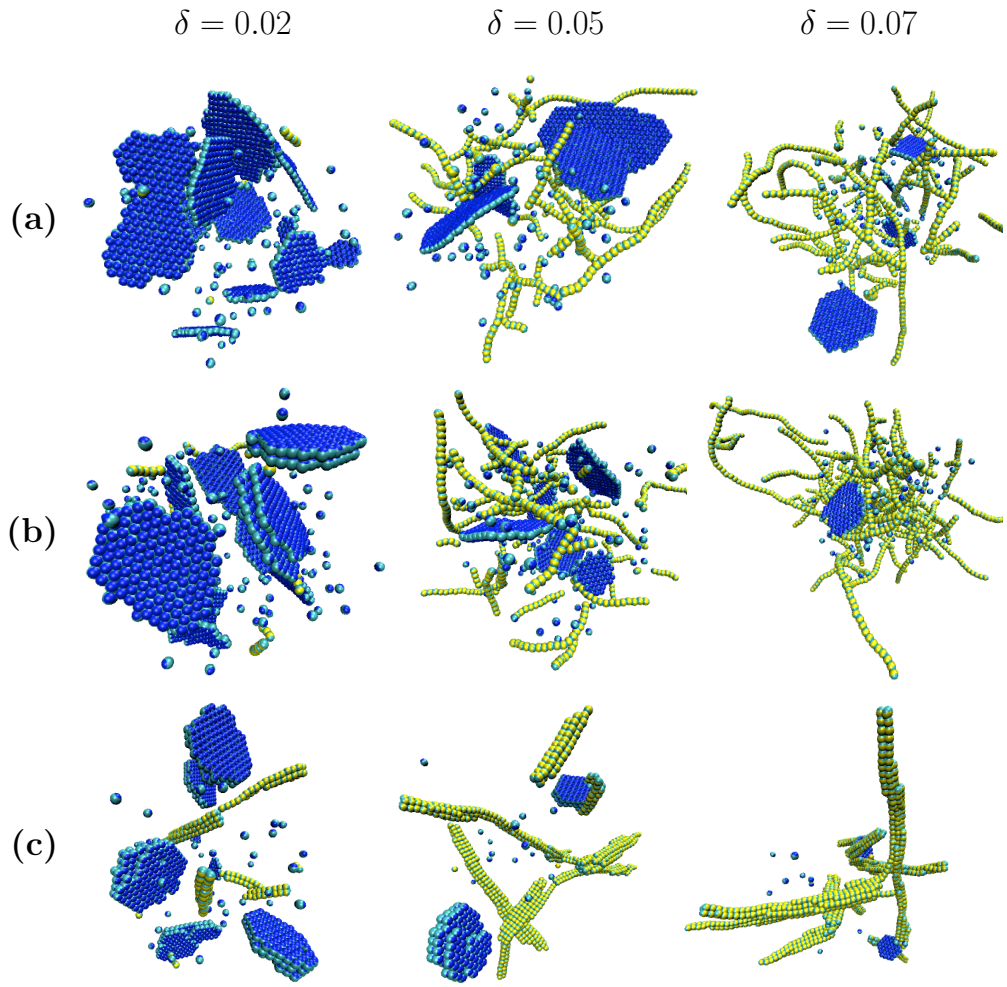


Figure 2.7: Long-time conformations for self-assembly of conformation-changing units at nine different thermodynamic states. Columns contain states at fixed values of δ; rows contain states at fixed values of the minor couplings. (a) Non-attracting species ($\epsilon_{eq}^{(min)} = 0 = \epsilon_{pol}^{(min)}$); (b) conditions under which planons have a tendency to stack ($\epsilon_{eq}^{(min)} = 0, \epsilon_{pol}^{(min)} = 3k_BT$); and (c) strongly-attracting species ($\epsilon_{eq}^{(min)} = 4k_BT, \epsilon_{pol}^{(min)} = 5k_BT$). Assemblies with hybrid character appear as minor couplings are increased, but the fundamental kinetic bias in favor of string formation remains.

2.6 Discussion

Cooperative binding as a mechanism for precipitous response is a textbook notion of biochemistry: interactions with one ligand heighten affinity for additional binding events, establishing positive feedback and enabling sharp switching behavior. Our results illustrate the dramatic *kinetic* implications of this kind of cooperativity. The molecular dynamics of self-assembly enriches traditional cooperativity scenarios in several interesting ways. The

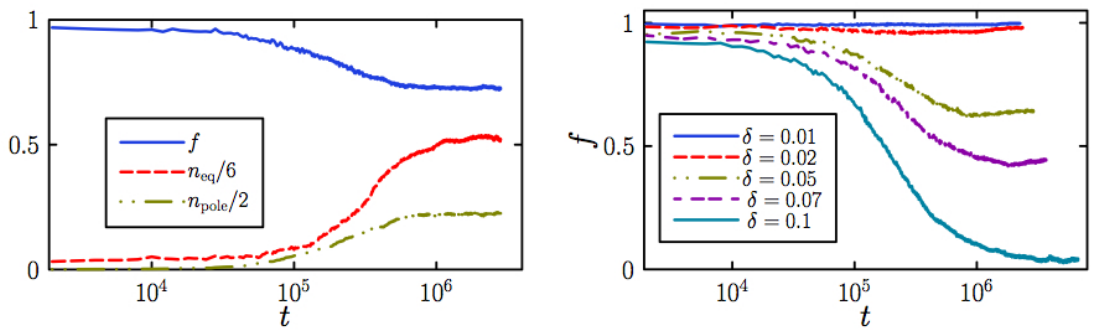


Figure 2.8: Assembly kinetics for interconverting units with vanishing minor couplings ($\epsilon_{eq}^{(min)} = 0 = \epsilon_{pol}^{(min)}$). (a) Time-dependence of the fraction of planon units f and the number of optimal equatorial (polar) contacts per particle n_{eq} (n_{pol}) for the trajectory shown in Figure 4 (with $\delta = 0.05$, main text). (b) Time dependence of the fraction of planon units f for single trajectories with different values of the internal bias δ.

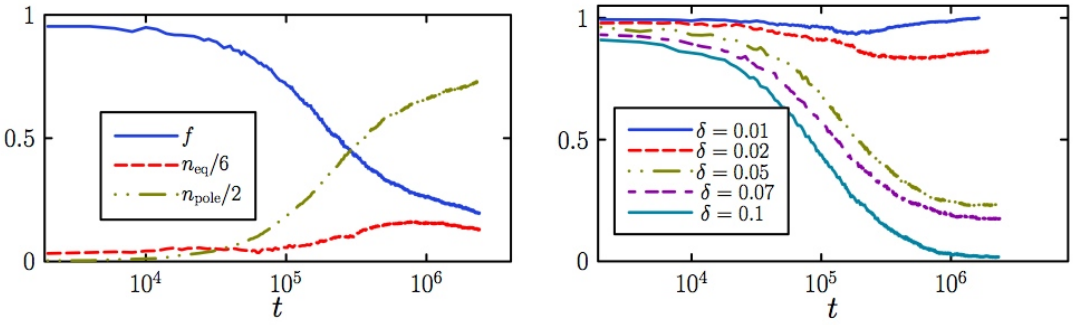


Figure 2.9: Analog of Figure 2.8 with considerable minor couplings ($\epsilon_{eq}^{(min)} = 4k_BT, \epsilon_{pol}^{(min)} = 5k_BT$). Time-dependence of the fraction of (a) planon units and the number of equatorial contacts (b) planon units f for single trajectories with different values δ.

growth of extended structures, for example, introduces spatially heterogeneous effects of binding and leads to the emergence of 'microphases' of distinct response types. In addition, kinetic barriers to the formation of distinct structures can dramatically alter the threshold conditions at which switching occurs, imbuing the resulting behavior with a fundamentally nonequilibrium character. Further, cooperative binding can drive the formation of composite assemblies through fusion of well-defined superstructures.

In the context of rosettasome self-assembly our results support a picture in which internal conformational fluctuations bias superstructure nucleation, thereby inducing the simultaneous growth of structures with distinct symmetries. Rosettasome allostery is a known phenomenon, while superstructure coexistence is a striking feature of recent experiments involving the wild-type protein and its genetically-engineered variants. While the function

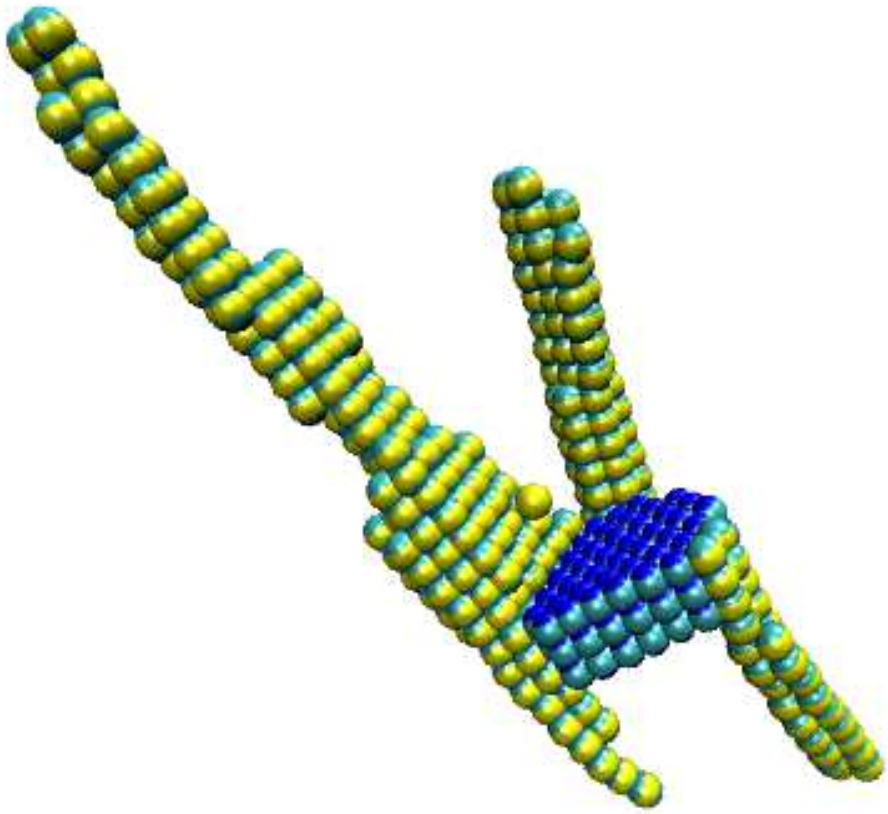


Figure 2.10: Structure obtained from a long trajectory of units possessing considerable minor interactions ($\epsilon_{eq}^{(min)} = 4k_BT, \epsilon_{pol}^{(min)} = 5k_BT$). The interconversion of units near the periphery of sheets and strings permits the fusing together of distinct superstructure types.

of many proteins depends on self-organization at the nanometer scale, the emergence of long-range order with a *variable* outcome is exceptional. Such a capability might be of considerable use to a hyperthermophilic organism seeking to respond to extremes of heat by promoting, for instance, the rapid assembly of membrane-stabilizing scaffolding (biases for particular rosettasome conformations might be mediated by changes in their protein composition, known to vary in response to environmental cues [59]). From the perspective of nanoscience, mimicking such cooperativity offers a plausible route to adaptive assembly, whereby subtle changes in environment can engender a dramatic change in the assembly products of a system.

Chapter 3

Cyclic Peptide Self Assembly

3.1 Polymer Nanocomposites

Polymer nanocomposites comprise a broad class of novel materials in which nanoparticles are embedded in a polymer matrix. The types of nanoparticulate matter and polymers that can be used to create polymer nanocomposites are quite general. The first systems studied experimentally occurred in the early 1990s, when scientists at Toyota Corp. realized that the addition of a small amounts of nanoparticulate clay to homopolymer matrices could have a profound impact on the mechanical strength, temperature resistance and durability of the plastics now widely used in automobile manufacturing [56]. Since then, much experimental research has sought discover new functional and mechanical properties of nanocomposites, and to elucidate their microscopic origins. The field has advanced to the point where coarse control over the distribution and orientation is possible [114][61][60][143][102], but improved, nanometer-level control over feature spacing, domain shape and alignment and function chemistry, which could have profound impacts on our ability to engineer novel functional materials, remains elusive.

One particular application of interest is the creation of polymeric thin films with vertically aligned pores. These films have potential uses in a variety of fields, ranging from photonic band gap materials [33], battery applications [117], nanoelectrode arrays [106][46], gas separation [19], but high fidelity assembly over macroscopic length scales remains a challenge. While crosslinked homopolymer networks may generate semi-porous networks suitable for thin film applications, the inherently unpredictable dispersity of pore sizes makes them unfit for the applications that require finely tuned control over pore size, distribution, and interior chemistry.

Meanwhile, block copolymers are low cost, easily manipulated systems well known to undergo an order-disorder transition (ODT) below a critical temperature into a variety of phases, including a cylindrical, hexagonal packing [70]. Using plasma etching techniques, it is possible to produce nanoporous thin films from cylindrical block copolymers with nearly monodisperse pores as small as a few nanometers [107]. However, even this scale is too

large by a factor of 10 to selectively discriminate among small molecules demanded in the preceding areas.

Modern chemical synthesis techniques allow for the size and shape of nanoparticles to be precisely tailored, and the synthesis of many types of functional nanoparticles, such as inorganic nanoparticles, carbon nanotubes, fullerenes, dendrimers, and biological materials, is possible. Thus, thin-film spanning pores comprised of sub-nanometer scale nanoparticles inserted into the polymer thin film matrix is seen as a possible mechanism for enabling the above technologies. In particular, the alignment of preformed carbon nanotubes inside the polymer matrix is one possible strategy that has been considered [137], but orienting carbon nanotubes has so far proven challenging [62][43].

A promising mechanism for subnanometer pore alignment is to allow the pore building blocks to co-assemble with the polymer matrix, rather than being inserted into the matrix pre-assembled. Such an approach leverages the power of modern chemical synthesis, with many known self-assembling motifs that offer precise control over pore size, shape and chemistry. Unlike preformed carbon nanotubes, where high aspect ratio units make the system prone to aggregation and kinetically arrested assemblies, the ability for individual units to cooperatively assemble with the polymer matrix suggests that appropriate coupling between nanoparticle units and the cylindrical polymer phase, as well as specific intermolecular interactions such as hydrogen bonding or other electrostratic interactions to promote nanoparticle assembly, may be sufficient for proper nanotube localization and assembly. Such a strategy also obviates the need to prepare nanotubes monodisperse in length, since such assembled tubes should, at least in principle, span the thickness of the film.

In this thesis, we draw inspiration from a particular organic "nanoparticle" known as cyclic peptides [52], which are comprised of an even number of D and L amino acids, cyclized to form a ring structure of well defined size and shape (see Figure 3.1).

Assembly of multiple cyclic peptide units is facilitated by the same hydrogen bonding interactions responsible for α and β secondary structure formation in true biological systems. The interactions responsible for biologal secondary structure are generated from the atoms on the peptide backbone, meaning the specific choice of sidechain is quite flexible for cyclic peptides, leaving the exterior of the self assembled cyclic peptide tube open to extensive modification, via choice in specific amino acid sequence, as well as "post-translational" modifications to amino acid structure found in biology, allowing for a myriad of approaches for passivating and coupling cyclic peptides to each block copolymer segment. Further, the inward facing C_α hydrogen is available for covalent substitution, and many such artificial amino acids are known which effectively modifiy the chemical properties of the interior of the assembled cyclic peptide tube. Further, using solid phase peptide synthesis provides extensive control over the number and location of modifications in the peptide sequence [25].

Beyond the obvious importance for understanding the specific behavior of the cyclic peptide - block copolymer system, this work is representative of an emerging, yet not very well understood, field of self assembly, in which multiple, interacting motifs must assemble into a final coordinated structure. Does assembly of each component proceed independently, cooperatively, or hierarchically? What are the possibilities for such assembly pathways? The

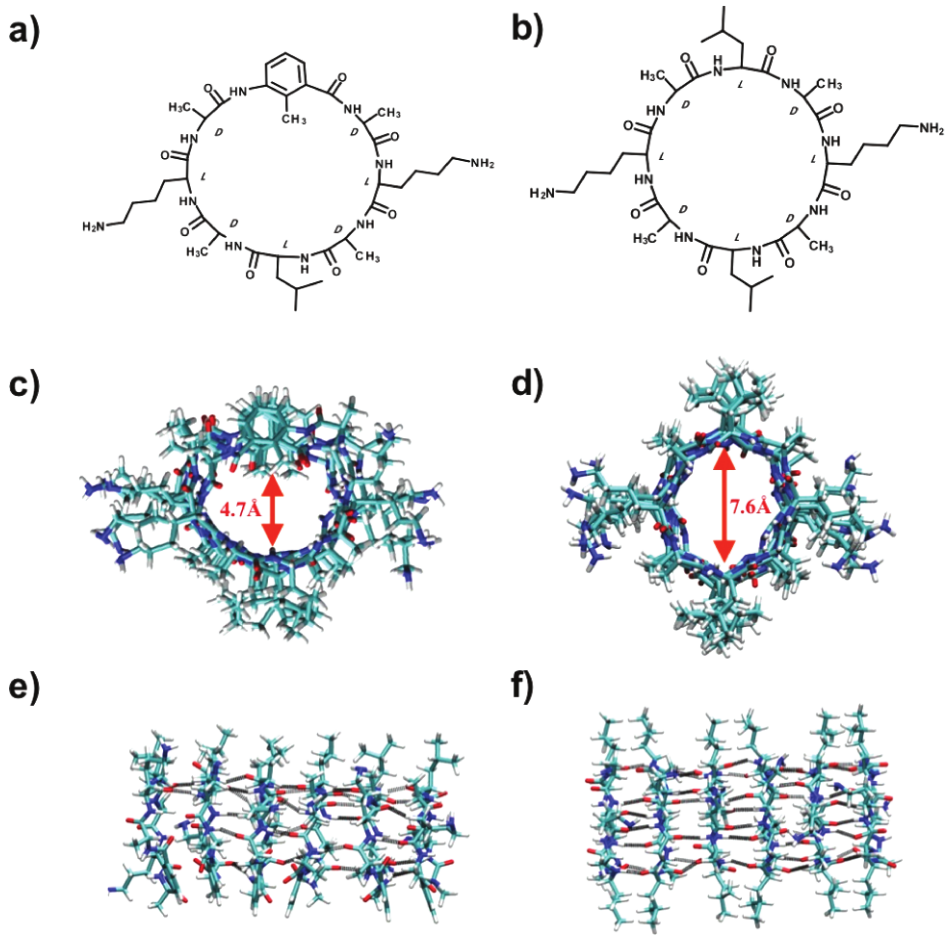


Figure 3.1: Figure taken from [52]. Chemical structures of a substituted 8-membered cyclic peptide (a) and (b) its conventional analogue. Detailed assembled structure of a cyclic peptide oligomer, computed from Molecular Dynamics simulations for each respective structure are shown in (c-f).

cyclic peptide system serves as an example of such a complicated system, while still being sufficiently simple compared to biological analogues, for instance.

3.2 Model

In this work we consider the cyclic peptide - block copolymer system from [137] as motivation to study the self assembly of nanoparticles within a self-assembling block copolymer matrix. Our efforts are directed at developing a model that incorporates the relevant physics of both constituent systems, and then investigating the possibilities for self assembly and characterizing the driving forces behind different modes of assembly. To acheive this we will need to develop a model for block copolymer assembly as well as a model for cyclic pep-

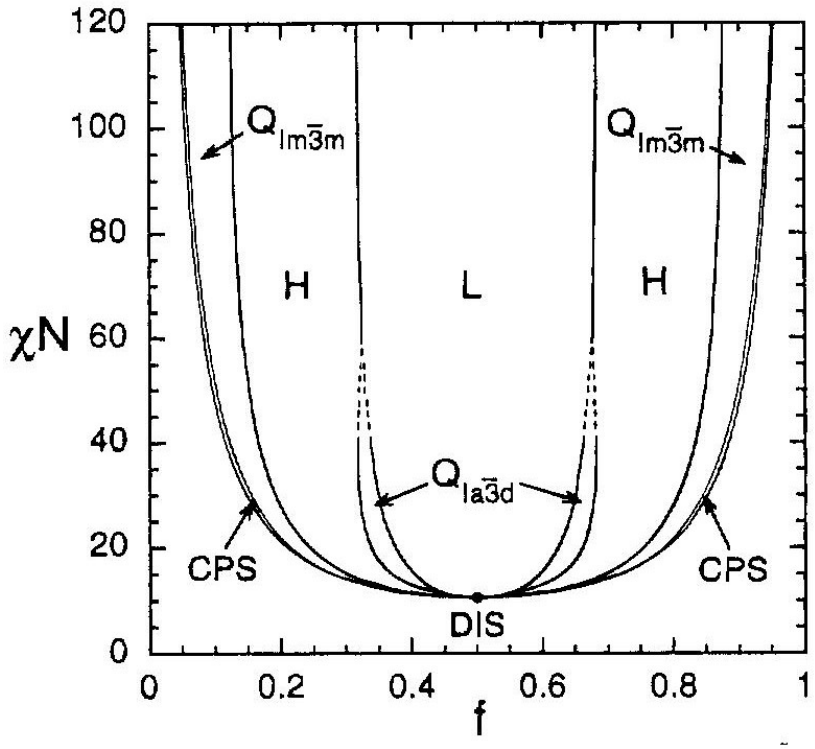


Figure 3.2: Phase Diagram for an AB diblock copolymer of length N in terms of the fraction f of A monomer and the AB Flory interaction paramter, χ. Below a critical value X_sN no ordering occurs. Above χ_sN, the lamellar phase (L) occupies most of the diagram and is the only available phase for a symmetric ($f = 0.5$ copolymer. Other phases include a hexagonal packing of cylinders (H), close packed spheres (CPS), gyroid (Q) and iambic (Q). Taken from [80]

tide/nanoparticle assembly. Then, we will develop a coupling between these two systems. For the remainder of the thesis, we will assume a diblock copolymer to be composed of two contiguous regions of distinct monomer types, A and B, which are in some sense chemically incompatible, which loosely means there is a tendency for the two types to separate, due to chemical differences. The extent to which the segment types are incompatible is captured by the Flory χ parameter, and together with the fraction f of the polymer composed of A monomers, the self assembly phase diagram is well known, see 3.2.

For the block copolymer system, we use a field-theoretic Hamiltonian originally dervied by Leibler [70], and subsequently reduced to a well known model for systems undergoing a fluctuation-induced first order phase transition [36], the Brazovskii model [18]. Leibler's model treats the individual polymers as ideal Gaussian polymers, meaning an energy func-

tional quadratic in local stretching,

$$\beta H_G[r] = \frac{3}{2b^2}\int_0^N ds \left|\frac{dr(s)}{ds}\right|^2 \tag{3.1}$$

together with a Flory-Huggins style interaction between A and B monomers, which depends on the local densities of each monomer type at each point in space. For a system of n block copolymers, the interaction is given by

$$V\left[\hat{\rho}_A, \hat{\rho}_B\right] = \chi \int dr\ \hat{\rho}_A(r)\hat{\rho}_B(r), \tag{3.2}$$

where $\hat{\rho}_A(r)$ and $\hat{\rho}_B(r)$ are given by,

$$\hat{\rho}_A(r) = \sum_{j=1}^{n}\int_0^f ds\ \delta(r - r_j(s)) \tag{3.3}$$

$$\hat{\rho}_B(r) = \sum_{j=1}^{n}\int_f^N ds\ \delta(r - r_j(s)). \tag{3.4}$$

Making the assumption that the block copolymer melt is incompressible, (i.e. $\hat{\rho}_A(r)+\hat{\rho}_B(r) = \rho_0$), the partition function is given by

$$Z = \int \mathcal{D}r^n e^{-\beta H_G[r^n]-\beta V[\hat{\rho}_A,\hat{\rho}_B]}\ \delta\left[\hat{\rho}_A + \hat{\rho}_B - \rho_0\right] \tag{3.5}$$

By invoking functional analogues of well known δ-function identities,

$$1 = \int \mathcal{D}\rho_A \delta\left[\hat{\rho}_A - \rho_A\right] = \int \mathcal{D}\psi_A \int \mathcal{D}\rho_A\ e^{-i\int dr \psi_A(r)\rho_A(r)} \tag{3.6}$$

Eq. 3.5 may be converted into an integration over the density fields ρ_A, ρ_B,

$$Z = \int \mathcal{D}\rho_A \mathcal{D}\rho_B e^{-H[\rho_A,\rho_B]} \tag{3.7}$$

Eq. 3.7 is an exact expression for the partition function, but the Hamiltonian is highly anharmonic in $\rho_A(r)$ and $\rho_B(r)$, making analytical evaluation as hopelessly complicated as evaluation of Eq. 3.5. Leibler's result is obtained by applying the random phase approximation (RPA) to evaluate the otherwise intractable nonlinear field theory 3.7. Due to the incompressibility constraint in Eq. 3.5, only one of the density fields fluctuates independently. Therefore it is convenient to express the Hamiltonian in terms of the order parameter field

$$\phi(r) = \rho_A(r) - \rho_B(r) \tag{3.8}$$

This representation has the advantage that $\phi = 0$ signifies the interface between A-rich and B-rich regions. Further, in the disordered phase, $\langle\phi(r)\rangle = 0$. The Hamiltonian is then given by

$$\beta H\left[\phi\right] = \rho_0 \int dr \frac{\lambda}{4!}\phi(r)^4 + \frac{\alpha}{3!}\phi(r)^3 + \frac{1}{2}\phi(r)\left[\tau + \epsilon q_0^{-2}(\nabla^2 + q_0^2)^2\right]\phi(r) \tag{3.9}$$

For a symmetric block copolymer ($f = 0.5$), α vanishes, leaving a symmetric $\phi^4 - \phi^2$ potential. Further, the fourth order gradient can be integrated by parts, $\int dr\phi\nabla^4\phi = \int dr\left(\nabla^2\phi\right)^2$, and it is instructive to rewrite the Hamiltonian as

$$\beta H\left[\phi\right] = \rho_0 \int dr \frac{\lambda}{4!}\left(\phi^2 + \frac{6\tau}{\lambda}\right)^2 + \frac{\epsilon}{2q_0^2}\left(\nabla^2\phi + q_0^2\phi\right)^2 + const \tag{3.10}$$

Thus, the Hamiltonian is comprised of two separate, positive quadratic terms. Since $\lambda > 0$ and $\epsilon > 0$, the lowest energy configuration will must strike a balance between minimizing each term, since the constraint that $\int dr\phi(r) = 0$ implies Eq. 3.10 can never reach 0. For $\tau \geq 0$, no phase separation can be expected since the first term is a single well potential centered around $\phi(r) = 0$. For $\tau < 0$, the first term drives $\phi(r)$ towards ± 1 which can physically be interpreted as double-well, local free energy density that drives macroscopic phase separation between unlike monomer types, A and B. Since $\int dr\phi(r) = 0$, the system must phase separate into equal volumes of A rich and B-rich regions. The second term ascribes ordering to the phase separation. It is easily seen that cosine waves $\cos\ (\vec{q}_0 \cdot \vec{r})$ with wavevector $|\vec{q}_0|^2 = q_0$ minimize the quadratic gradient term while satisfying the constraint $\int dr\phi(r) = 0$. Thus, the model's preference for ordering along any wavevector of magnitude q_0 is manifest.

Since the random phase approximation takes a functional Taylor expansion to quadratic order about a homogenous state [37] of the nonlinear portion of the Hamiltonian in 3.7, 3.10 is useful for studying systems only weakly perturbed past the order-disorder transition (ODT). The advantage of the model is that simulations of the model are significantly less computationally demanding, and significantly easier to implement, than field-theoretic simulations of the full non-Gaussian Hamiltonian without the RPA, and can access time scales and system sizes much greater than more detailed molecular simulation techniques, such as atomistic or even bead-spring molecular dynamics and Monte Carlo techniques. The field theoretic approach may be expected to capture the same long wavelength phenomena of more detailed molecular approaches, while sacrificing the shorter wavelength structure lost in the field-theoretic coarse-graining. From a self-assembly perspective, such a system strikes a good balance between faithful representation of the block copolymer structure, while still acheiving the many characteristic wavelengths of order required to properly study structure formation in a nanocomposite.

Although some parameters in the model can be estimated given certain assumptions and approximations of the physical system, we are not concerned with model behavior at

particular parameter values, but rather, the types of assembled structures possible at a variety of reasonable parameter values. For our purposes it is sufficient to know that

$$\tau = 2\left(\chi_s N - \chi N\right) \tag{3.11}$$

is the only term in the model dependent upon temperature, and χ_s denotes the value for the A-B monomer interaction at the spinodal limit of stability and N is the index of polymerization.

Since the cyclic peptide system serves as motivation for studying interacting nanocomposite systems, we are interested only in the essential physics inherent in the system. Since the individual cyclic peptide units are able to interact via hydrogen bonds, we consider a patchy model in which spherical particles interact via head-to-tail interactions that vary rapidly over distance and weakly under angular rotations of the constituent particles. Thus, we choose to use the same geometric factors from the string-forming ("stringon") interactions of the archaeal chaperonin model. The justification for such a coarse short-ranged potential is the same as in the chaperonin model. When running Monte Carlo simulations we also adopt hard-core volume exclusions between particles,

$$u_{\text{rep}}(|r_i - r_j|) = \begin{cases} \infty, & |r_i - r_j| < \sigma \\ 0, & \text{else} \end{cases} \tag{3.12}$$

whereas in Brownian Dynamics simulations we use repulsive WCA cores,

$$u_{\text{rep}}(|r_i - r_j|) = \begin{cases} 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] + \epsilon, & |r_i - r_j| < 2^{1/6}\sigma \\ 0, & \text{else} \end{cases} \tag{3.13}$$

In sum, the total Hamiltonian for M cyclic peptides is

$$\mathcal{H}_{\text{np}} = \sum_{i \neq j} u_{\text{arch}}(|r_i - r_j|) + u_{\text{rep}}(|r_i - r_j|) \tag{3.14}$$

It is possible to formulate an exact description of the coupling between nanoparticle and polymer within the field-theoretic framework [115] by introducing a new field for nanoparticle density into the functional integration for the partition function 3.7. In this case, a Flory Huggins style interaction between nanoparticle density and the polymer composition field ϕ can be introduced to capture their interaction, and field theoretic sampling strategies may be applied [71]. Since the most prominent effect of the nanoparticles within the polymer melt is to exclude volume, large variations in the order parameter $\phi(r)$ would be expected in systems of practical interest. For a nanoparticle favorably coupled to and embedded within the A phase, ϕ would need to vary from $\phi(r) = 0$ within the nanoparticle, to $\phi(r) = 1$ just out of the particle. It is unclear whether the RPA, and hence our model, is suitable for such situations, since it relies on a system characterized by weak density inhomogeneities. We will delay a detailed investigation of this issue until Chapter 5. For now, we assume that

we can capture excluded volume effects by setting $\phi(r_{in}) = 0$ at every r_{in} located inside a nanoparticle, and the interaction between cyclic peptide and the polymer melt may be written in terms of the total nanoparticle density field and ϕ,

$$\mathcal{H}_{\text{cpl}} = \gamma \int dr \phi(r) \rho_p(r) \tag{3.15}$$

To define ρ_p, we place a density profile for each individual cyclic peptide at the particle's position, and then define the total density as the sum of these individual densities. The form of the density profile is not too important, provided it can be normalized to 1 and vanishes outside of a bounded interval, such as a truncated and shifted Gaussian. For computational techniques we use in this thesis, it is convenient to have a density profile that is infinitely differentiable so that the Fourier representation of the distribution converges quickly. Since a truncated and shifted Gaussian has a kink at the truncation point, we use the following instead

$$\rho_p(r) = \sum_{j=1}^{M} e^{\frac{1}{1-r \cdot r}} \tag{3.16}$$

3.3 Methods

We would like to simulate M nanoparticles of diameter 1 in an AB diblock copolymer melt in 3 dimensions, where the order parameter $\phi(r, t)$ has been discretized on a square lattice with N_x grid points in x, y and z. We assume the spatial extent of each dimension is L, which indicates $\Delta x = \frac{L}{N_x}$. We use two different dynamical rules to generate a canonical ensemble of fluctuations in the order parameter ϕ consistent with the Hamiltonian 3.10. The most straightfoward method is a Metropolis Monte Carlo algorithm, which has the advantage that a variety of proposal moves can be constructed, including those which preserve the volume excluded interior of embedded nanoparticles. A pseudospectral method to simulate a conserved Langevin dynamics of the conserved order parameter ϕ is also developed [15]. Such an approach has the advantage that at each time step, every degree of freedom in the system is updated. Further, as we will show below, using spectral techniques to evaluate the derivatives in 3.17 allows for an exponential gain in accuracy for a fixed system discretization, allowing for a much larger time step Δt than if a simple finite differences scheme was used to propagate the dynamics. Unfortunately, discontinuities in the Hamiltonian introduce singularities into the propagation rule that are not easily handled via the Langevin approach. This fact becomes especially important when we consider a block copolymer - nanoparticle system with volume exclusion. Although it should be possible to exclude the polymer from the interior of a nanoparticle by introducing a Lagranian constraint into the dynamical equations [44], we have instead used Metropolis Monte Carlo to generate trajectories with quasi-dynamical trial moves. Such an approach easily allows us to enforce the constraint that polymer density vanishes in the interior of nanoparticles.

Pseudospectral Langevin simulations

Since the order parameter $\phi(r)$ is conserved over space, an appropriate Langevin dynamics that preserves this constraint is the so-called Model B dynamics [50],

$$\frac{\partial \phi(r)}{\partial t} = \nabla^2 \frac{\delta \mathcal{H}[\phi]}{\delta \phi(r)} + \eta(r,t) \tag{3.17}$$

where $\eta(r,t)$ is a conserved Gaussian random field that satisfies

$$\langle \eta(r,t) \rangle = 0 \tag{3.18}$$

$$\langle \eta(r,t)\eta(r',t') \rangle = -2\beta^{-1}\nabla^2 \delta(r-r')\delta(t-t') \tag{3.19}$$

In order to calculate $\eta(r,t)$ in 3.17, we follow Puri and Oono [97] and first consider a non-conserved 3-dimensional Gaussian random field $\vec{\gamma}(r,t)$ such that

$$\langle \vec{\gamma}(r,t) \rangle = 0 \tag{3.20}$$

$$\langle \gamma_a(r,t)\gamma_b(r',t') \rangle = 2\beta^{-1}\delta_{a,b}\delta(r-r')\delta(t-t') \tag{3.21}$$

$\eta(r,t)$ is then given by

$$\eta(r,t) = \nabla \cdot \vec{\gamma}(r,t) = \sum_{j=1}^{3} \frac{\partial \gamma_j(r,t)}{\partial r_j} \tag{3.22}$$

Defined in this way, η satisfies 3.18 and 3.19.

Since the term $\nabla^2 \frac{\delta \mathcal{H}[\phi]}{\delta \phi(r)}$ in 3.17 will contain gradient terms of at least second order, discretization will introduce errors associated with a discrete approximation to the spatial derivatives. A centered finite difference approximation to the partial derivatives of ϕ is $O(\Delta x^2)$,

$$\frac{\partial \phi(r)}{\partial r_x} = [\phi(r_x + h) - \phi(r_x - h)]/(2\Delta x) + O(\Delta x^2) \tag{3.23}$$

In general, the error associated with higher order centered finite differences scales with the number of lattice sites used to compute the estimate of the derivative. For example, for a fourth order centered finite difference, the error is $O(\Delta x^4)$. However, since we are interested in a system of finite volume V with periodic boundary conditions, the Fourier expansion of a function $f(r)$ defined on the system is

$$f(r) = \frac{1}{V} \sum_k c_k e^{-ik\cdot r} \tag{3.24}$$

$$c_k = \int dr f(r) e^{ik\cdot r} \tag{3.25}$$

Since differentiation is exact in Fourier space,

$$\frac{\partial \phi(r)}{\partial r_x} = -i\frac{1}{V}\sum_k k_x \; c_k e^{-ik\cdot r} \tag{3.26}$$

When only N modes are retained in the expansion, the approximation to the derivative is an N-point formula, and one would hope that, in analogy with the centered finite differences, the error scales as $O(\Delta x^N) = O(N^{-N})$. Fortunately, for a cubic system with periodic boundaries, the error in this approximation *can* scale as $O(N^{-N})$, depending on the location of singularities in $\phi(r)$ [16]. To see this, note that with periodic boundaries differentiation is a self-adjoint operator, and if f has m continuous derivatives, then integrating 3.25 by parts m times gives

$$c_k = (-ik)^{-m}\int dx f^{(m)}(x)e^{ikx} \leq |f^{(m)}|/k^m \tag{3.27}$$

For infinitely differentiable functions, 3.27 implies that the Fourier coefficients c_k decay exponentially fast for sufficiently large k. However, if ϕ only has m continuous derivatives on the system domain, then convergence of the Fourier series will be algebraically limited to $\mathcal{O}(N^{-m})$ [16]. In terms of overall complexity associated with an update for time Δt, the finite difference scheme requires $\mathcal{O}(N)$ work, and is $\mathcal{O}(N^{-2})$ accurate. Meanwhile, a pseudospectral update requires a forward and reverse FFT pair, requiring $\mathcal{O}(N \log N)$ operations but is $\mathcal{O}(e^{-N})$ accurate if ϕ is sufficiently smooth. Thus, a pseudospectral method is marginally more expensive per degree of freedom, but asymptotically, exponentially more accurate.

For the bilinear portion of $\mathcal{H}$, the pseudospectral technique can easily be applied, since

$$\frac{\delta \mathcal{H}[\phi]}{\delta \phi(r)} = \frac{\delta}{\delta \phi(r)}\int dr'' \int dr' \phi(r')\chi(r''-r')\phi(r'') \tag{3.28}$$

$$= \int dr' \phi(r')\chi(r-r') = \frac{1}{V}\sum_{k'} \hat{\chi}(k')\hat{\phi}(k')e^{-ikr} \tag{3.29}$$

Upon Fourier transforming,

$$\frac{\delta \mathcal{H}[\hat{\phi}]}{\delta \hat{\phi}(k)} = \hat{\chi}(k)\hat{\phi}(k) \tag{3.30}$$

$$\frac{\partial \hat{\phi}(k)}{\partial t} = -k^2\hat{\chi}(k)\hat{\phi}(k) + \hat{\eta}(k,t) \tag{3.31}$$

However, for the quartic term, the Fourier space representation has nonlocal couplings which are costly to evaluate,

$$H[\phi] = \frac{1}{4}\int dr\phi(r)^4 \tag{3.32}$$

$$\frac{\delta H[\phi]}{\delta\phi(r)} = \phi(r)^3 = \int dk_1 \int dk_2 \int dk_3 \; \hat{\phi}(k_1)\hat{\phi}(k_2)\hat{\phi}(k_3)e^{-ir(k_1+k_2+k_3)} \tag{3.33}$$

$$\frac{\delta H[\hat{\phi}]}{\delta\hat{\phi}(k)} = \int dk_1 \int dk_2 \int dk_3 \; \hat{\phi}(k_1)\hat{\phi}(k_2)\hat{\phi}(k_3)\delta(k_1+k_2+k_3-k) \tag{3.34}$$

The solution is to use a semi-implicit algorithm where the bilinear term is treated by an implicit integration step in Fourier space, while the nonlinear terms are treated with explicit steps in real space, which is much more stable than a fully explicit Euler integration step. For the Hamiltonian 3.10, the continuous and discretized equations are (absorbing ρ_0 into the units of time),

$$\frac{\partial\hat{\phi}(k)}{\partial t} = -k^2\frac{\lambda}{3!}\widehat{\phi^3}(k) - k^2\frac{\alpha}{2!}\widehat{\phi^2}(k) - k^2\left[\tau + \frac{\epsilon}{q_0^2}(q_0^2-k^2)^2\right]\hat{\phi}(k) + \hat{\eta}(k,t) \tag{3.35}$$

$$\frac{\phi_k^{n+1}-\phi_k^n}{\Delta t} = -k^2\frac{\lambda}{3!}\left[\phi^3\right]_k^n - k^2\frac{\alpha}{2!}\left[\phi^2\right]_k^n - k^2\left[\tau + \frac{\epsilon}{q_0^2}(q_0^2-k^2)^2\right]\phi_k^{n+1} + \eta_k^{n+1} \tag{3.36}$$

Solving for ϕ_k^{n+1} and inserting $\eta_k^{n+1} = \sum_{j=1}^3 ik_j\left[\sigma_j\right]_k^{n+1}$ (from 3.22) gives

$$\phi_k^{n+1} = \frac{-k^2\Delta t\left[\frac{\lambda}{3!}\left[\phi^3\right]_k^n + \frac{\alpha}{2!}\left[\phi^2\right]_k^n\right] + \sum_{j=1}^3 ik_j\left[\sigma_j\right]_k^{n+1}}{1 + k^2\Delta t\left[\tau + \frac{\epsilon}{q_0^2}\left(q_0^2-k^2\right)^2\right]} \tag{3.37}$$

After randomly initializing the system, the algorithm consists of the following steps for each update:

1. Compute $\left[\phi^4\right]_r^n$, $\left[\phi^2\right]_r^n$
2. Generate new Gaussian random fields σ_j^{n+1}
3. Compute ϕ_k^n, $\left[\phi^4\right]_k^n$, $\left[\phi^2\right]_k^n$, $\left[\sigma_j\right]_k^{n+1}$ via Fast Fourier Transform (FFT)
4. Compute ϕ_k^{n+1} via 3.37
5. Inverse FFT to obtain ϕ_r^{n+1}

Metropolis Monte Carlo

To approximate the kinetics associated with diffusive order parameter dynamics, we use local proposal moves that mimic the flow of density fluctuation between neighboring lattice

cells. Order parameter conservation is therefore ensured on a local scale. A "seed" lattice site is chosen with uniform probability and then a neighboring lattice site is chosen with uniform probability from among the neighbors. Finally, a displacement $\Delta \in [-\Delta_{max}, \Delta_{max}]$ of density from the seed lattice site to its neighbor is chosen. Such a scheme has symmetric generation probabilities and the Metropolis acceptance probability is therefore

$$P_{acc} = \min\left[1, e^{-\beta \Delta H}\right] \tag{3.38}$$

3.4 Results and Conclusion

Bare Model

In Figures 3.7 and 3.8 we display self assembly phase diagrams in two different regimes of the diblock fraction f. Further, in Figures 3.3 to 3.6 we show representative trajectories for self-assembling block copolymer systems at a variety of parameters. At different parameter values, we are able to reproduce lamellar 3.3, cylindrical 3.4 and spherical 3.5 phases predicted by the theory, 3.2. Although for some parameter choices the model consistently assembles into well ordered strucutres, slow and frustrated dynamics are more typically encountered 3.6 at most parameter choices. In particular, large values for the quench depth, τ encourage rapid spinodal decomposition, whereupon local orderings around different $\vec{q_0}$ develop and defects between different regions do not anneal, allowing for long range frustration to set in. Since q_0 in 3.10 only specifies the magnitude of preferred wavevector ordering in the system, the inability to achieve long range order is inescapable, as all such wavevectors on the $d-$dimensional sphere are equally preferred.

Pseudo Volume Exclusion

In Figures 3.9 - 3.11 we show assembly trajectories in which nanoparticles have been inserted and the Hamiltonian 3.15 has been used to couple the nanoparticles to the block copolymer field ϕ. For physically plausible coupling values (γ = 1-20 kT), nanoparticles may be driven to particular phases of the block copolymer and both nanoparticle and polymer structure formation may be seen. However, these results are not surprising, since 3.15 couples linearly to the block copolymer density and simply acts as an externally applied chemical potential field. At larger coupling values, block copolymer is immediately recruited to the vicinity of each cylcic peptide, and overall tube structure formation is frustrated.

Corrected Volume Exclusion

As we show in Chapter 5, a first step towards introducing proper nanoparticle volume exclusion in the model is to enforce the constraint that $\phi(r) = \rho_A(r) - \rho_B(r) = 0$ inside each nanoparticle. To this end, we modify the Monte Carlo and Langevin methods from the

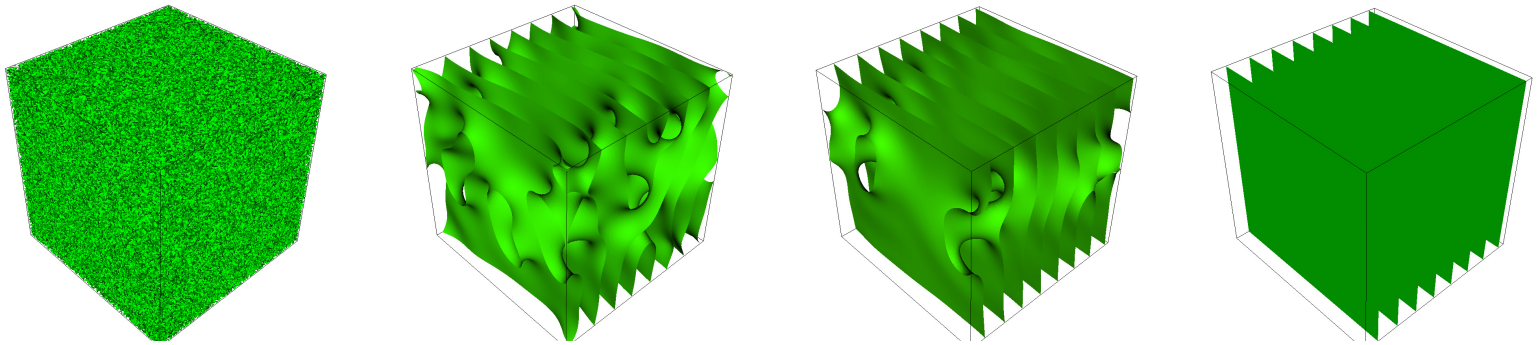


Figure 3.3: Self assembly trajectory with $f = 0.5,\ q_0 = 4.0,\ \tau = 1.0,\ \epsilon = 2.0,\ \lambda = 1.0$

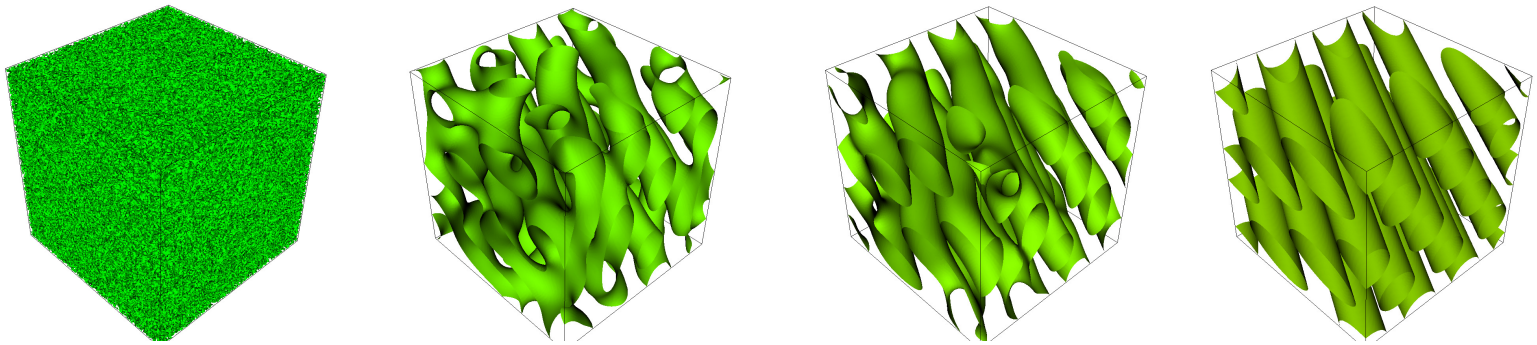


Figure 3.4: Self assembly trajectory with $f = 0.1,\ q_0 = 4.0,\ \tau = 1.0,\ \epsilon = 2.0,\ \lambda = 1.0$

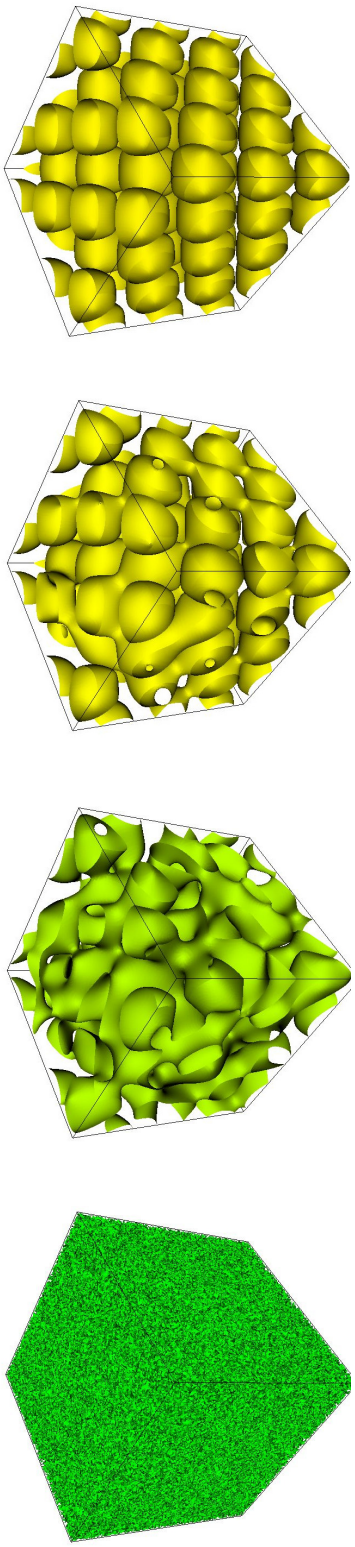


Figure 3.5: Self assembly trajectory with $f = 0.01,\ q_0 = 4.0,\ \tau = 1.0,\ \epsilon = 2.0,\ \lambda = 1.0$

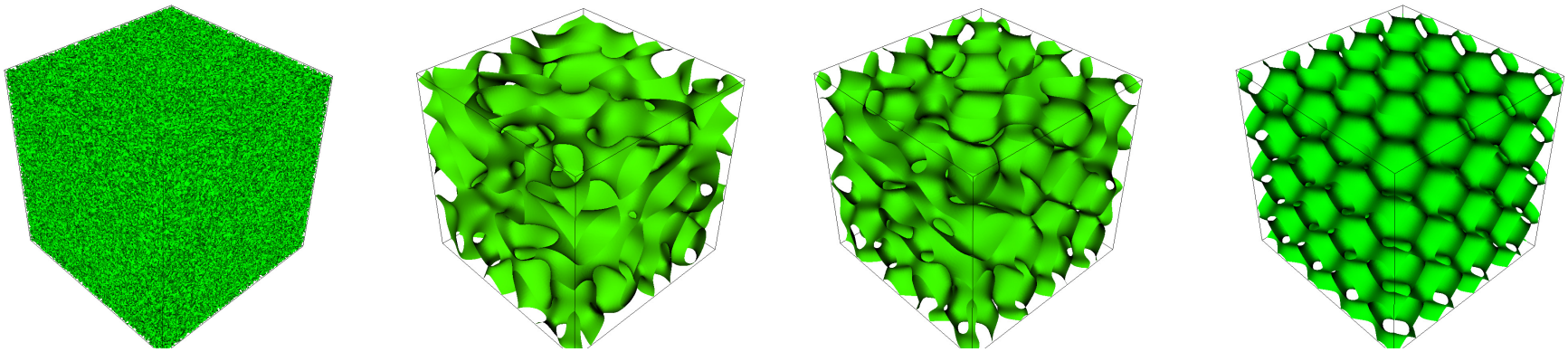


Figure 3.6: Self assembly trajectory with $f = 0.4,\ q_0 = 4.0,\ \tau = 1.0,\ \epsilon = 2.0,\ \lambda = 1.0$

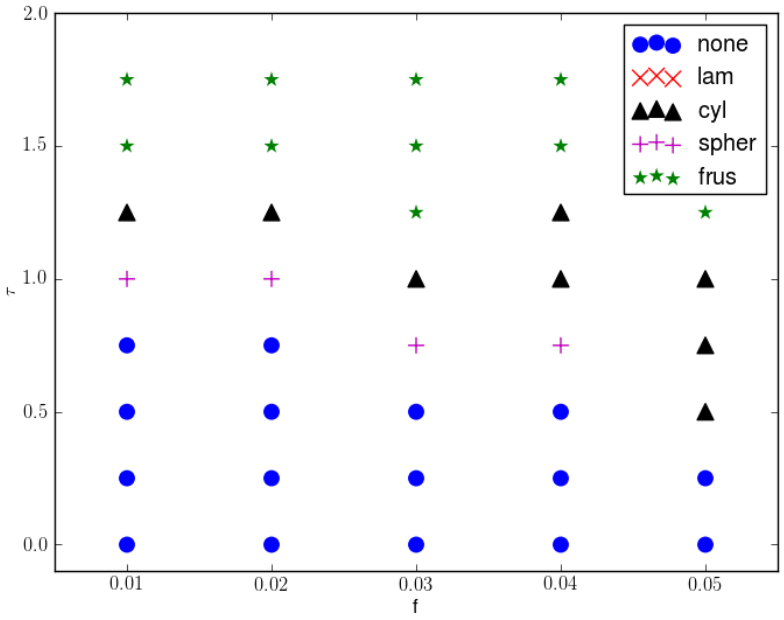


Figure 3.7: Self assembly phase diagram with $q_0 = 4.0, \epsilon = 2.0,\ \lambda = 1.0$, in the small f regime in which close-packed spherical (CPS) structures are formed.

previous section to enforce $\phi(r_{in}) = 0$ for all appropriate r_{in} inside a nanoparticle. Monte Carlo has the advantage that this constraint may be explicitly enforced when proposing trial configurations. Given that the current configurations satisfies the constraint, proposal schemes are constructed, with symmetric generation probabilities, to always propose a new configuration that also satisfies the constraint, ensuring detailed balance is satisfied.

Representing cyclic peptides as volume excluding spheres on a square lattice makes calculating the overlap between each sphere and square lattice site analytically infeasible and computationally onerous, requiring some flavor of numerical integration to calculate volume overlap exactly (ignoring numerical integration error). To ameliorate this problem, rather than computing partial overlaps, we fully exclude volume from the lattice sites whose centers lie within a nanoparticle sphere, and leave those with centers outside of the sphere untouched. This approximation is reasonably accurate for a sufficiently fine lattice. For a nanoparticle of diameter σ located at r_i, this amounts to

$$\phi_r = 0, \quad |r - r_i| < \sigma/2 \tag{3.39}$$

Enforcing the volume exclusion constraint when proposing changes in lattice values ϕ_r may be done by only proposing configurations that respect the volume constraint. Likewise, changes in the orientations of the nanoparticles require no modification, since these

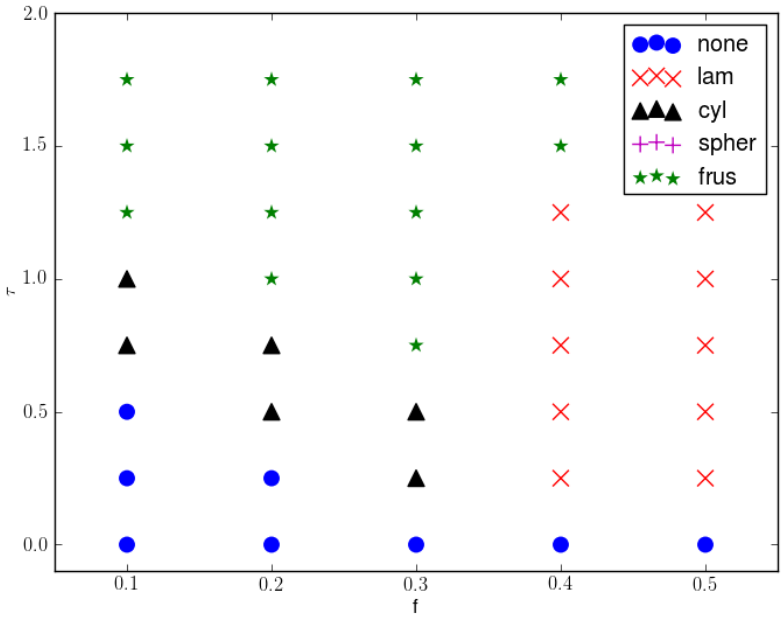


Figure 3.8: Self assembly phase diagram with $q_0 = 4.0, \epsilon = 2.0,\ \lambda = 1.0$, in the $f = 0.5$ regime in which predominantly lamellar structures are formed.

moves keep the positions of the nanoparticles intact and do not interfere with the volume exclusion constraint. Proposing changes in the positions of nanoparticles is complicated, however, since a change in the position of a nanoparticle necessitates a change in the block copolymer field to maintain volume exclusion at the new nanoparticle position. However, since the nanoparticles are constrained to occupy an integral number of lattice sites, after any move the number of newly vacated lattice sites will equal the number of newly occluded sites. Therefore, one possible way to propose a nanoparticle update is to make a one-to-one mapping between newly vacated and and newly occluded sites and copy the values from the newly occluded sites directly to the newly vacated sites. The newly occluded lattice values are then set to 0. Such a procedure is easy to implement and has symmetric Monte Carlo generation probabilities. Further, when a move occurs, it should be more beneficial to make this mapping in such a way that mimics physical dynamics by "pushing" polymer out of the way (Figure 3.12b), rather than reflecting about the particle's axis (Figure 3.12a).

A significant problem with this approach arises when considering the discrete finite differences used in evaluating the Hamiltonian's gradients. For simplicity, consider a cubic lattice with equal lattice spacing in 2 dimensions. The discrete form of the Hamiltonian for the

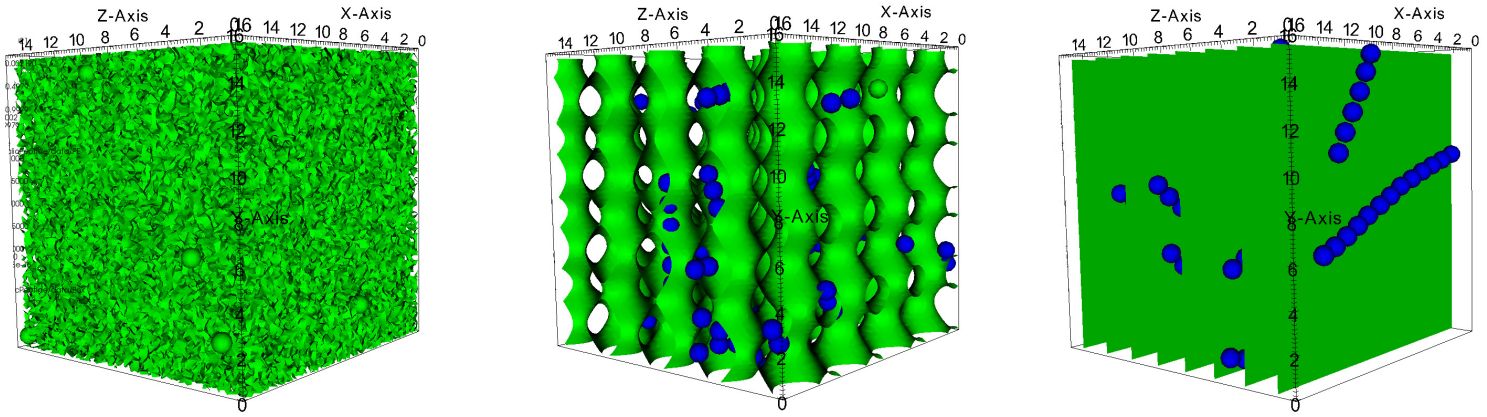


Figure 3.9: Self assembly trajectory for a no-exclusion nanocomposite, with $f = 0.1,\ q_0 = 4.0,\ \tau = 1.0,\ \epsilon = 2.0,\ \lambda = 1.0,\ \gamma = 10kT$

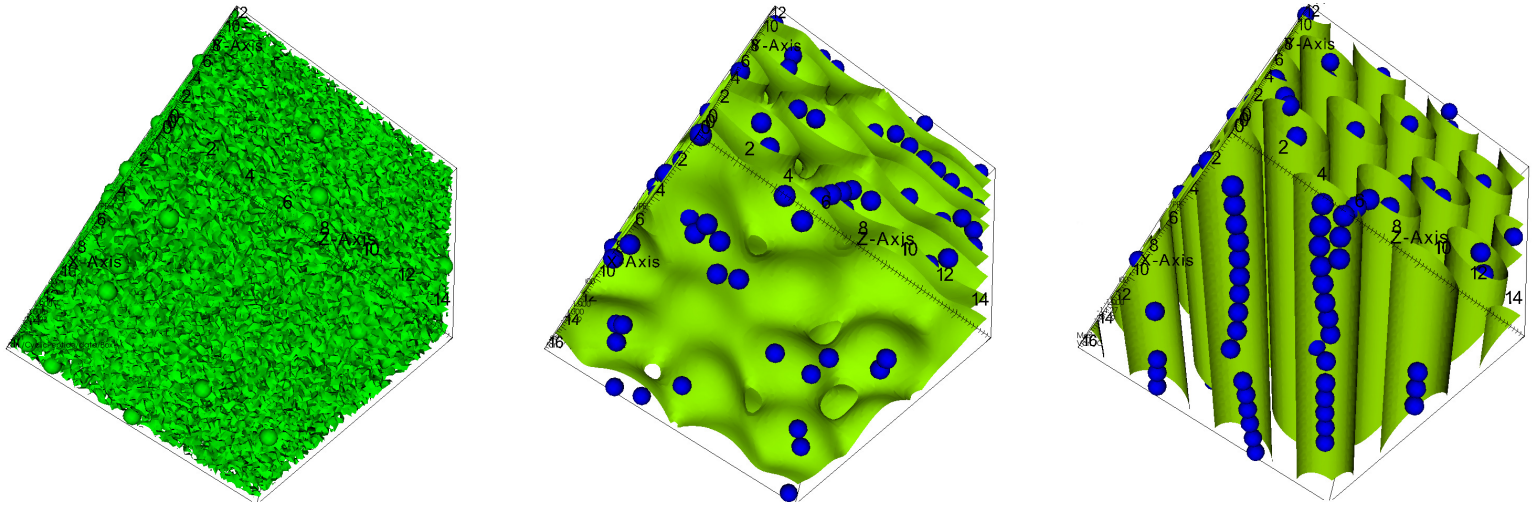


Figure 3.10: Self assembly trajectory for a no-exclusion nanocomposite, with $f = 0.1,\ q_0 = 4.0,\ \tau = 1.0,\ \epsilon = 2.0,\ \lambda = 1.0,\ \gamma = 10kT$

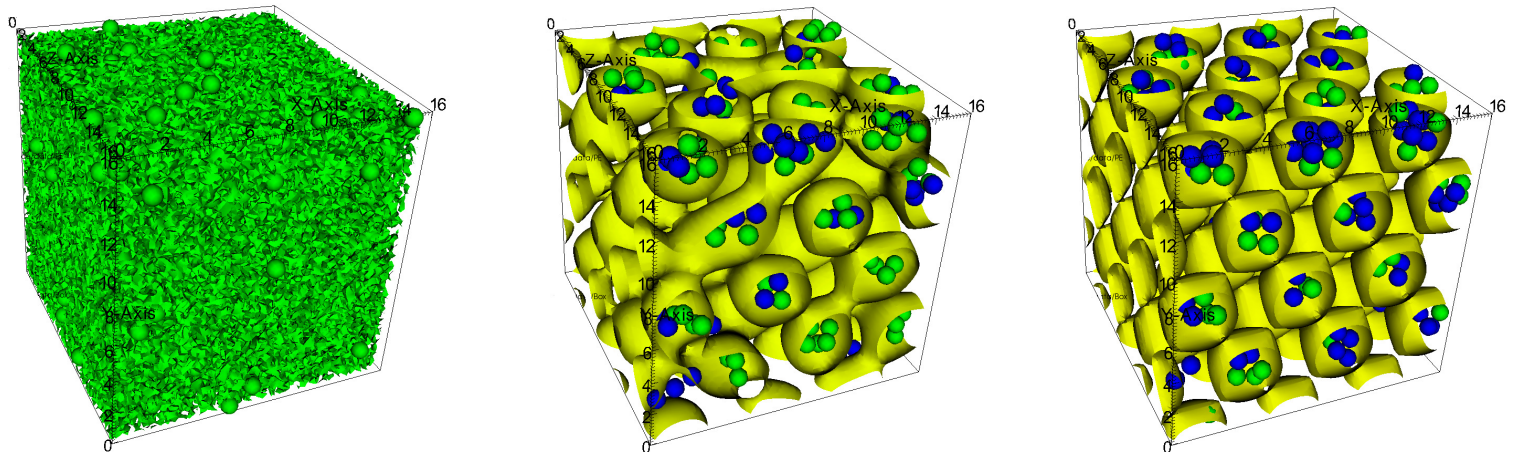
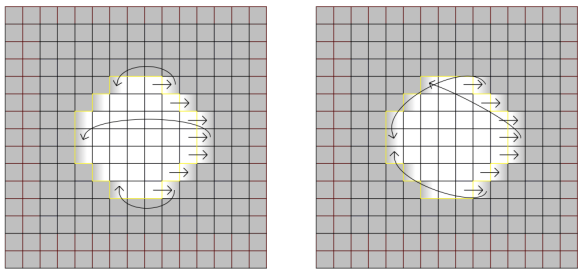


Figure 3.11: Self assembly trajectory for a no-exclusion nanocomposite, with $f = 0.1,\ q_0 = 4.0,\ \tau = 1.0,\ \epsilon = 2.0,\ \lambda = 1.0,\ \gamma = 10kT$



(a) (b)

Figure 3.12: Nanoparticle MC move. A 1-1 mapping is made to transfer lattice values from newly occluded sites to newly vacated sites. This mapping is illustrated by arrows showing direction of transfer. (a) The mapping is made by reflecting about the particle axis. (b) Lattice values are "pushed" to the side as the polymer moves.

symmetric caes (i.e. $\alpha = 0$ in 3.9) is

$$\beta H\left[\phi\right] = \rho_0 \sum_{i,j} \Delta x^2 \left(\frac{\lambda}{4!}\phi_{i,j}^4 + \frac{1}{2}\phi_{i,j}\left[\tau + \epsilon q_0^{-2}(\mathcal{L}^2 + q_0^2)^2\right]\phi_{i,j} \right) \tag{3.40}$$

Here, $\mathcal{L}$ is a choice of discrete Laplacian. Expanding the Laplacian term and inserting a centered finite differences for derivatives, 3.23, gives

$$\beta H\left[\phi\right] = \rho_0 \sum_{i,j} \Delta x^2 \left(\frac{\lambda}{4!}\phi_{i,j}^4 + \frac{1}{2}(\tau + \epsilon q_0^2)\phi_{i,j}^2 + \frac{\epsilon}{\Delta x^2}\phi_{i,j}\psi_{i,j} + \frac{\epsilon}{2\Delta x^4}\psi_{i,j}^2 \right) \tag{3.41}$$

$$\psi_{i,j} = \phi_{i+1,j} + \phi_{i-1,j} + \phi_{i,j+1} + \phi_{i,j-1} - 4\phi_{i,j} \tag{3.42}$$

From 3.41, it is clear that when the order of the derivative is greater than the dimension of the system, the gradient-squared term in 3.41 will have a prefactor inversely proportional to the lattice spacing. Since we have already assumed that the lattice spacing must be small to adequately represent the circular shape of the nanoparticle, it follows that the coupling strength for the difference between adjacent lattice sites $\phi_{i+1,j} - \phi_{i,j}$ is very large. When excluding volume, we require a discontinuity in the order parameter ϕ. We expect that for a nanoparticle embedded in a polymer, the field will jump from 0 inside the particle to a non-zero value representative of an A-rich or B-rich region of the polymer (this value is given by the zeros of the $\phi^4 - \phi^2$ polynomial, $\pm\sqrt{\frac{6|\tau|}{\lambda}}$). The enormously high energetic contributions to the Hamiltonian of the discrete Laplacians is representative of the fact that, as the lattice spacing approaches a continuum, these "derivatives" should diverge at a point of discontinuity in ϕ.

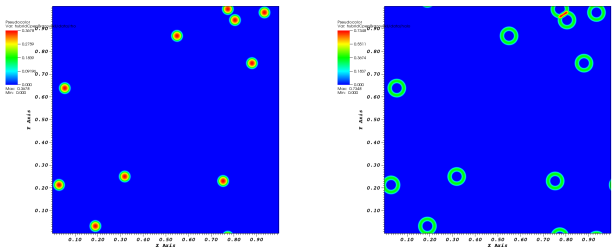
In practice we have determined the following rules of thumb. For a 2-dimensional system with parameters for the model to be $\rho_0 = 1000, q_0 = 4$, $\tau = 1$, $\epsilon = 1$, $f = 0.5$, each

lamellar ordering consists of an A-rich and B-rich region, for $4 \times 2 = 8$ total regions. The lattice spacing necessary to adequately resolve each region can be set by considering that the nanoparticles under consideration are on the scale of $1 - 5$nm and occlude at least 1 lattice site, which puts a maximum lattice size around 1nm. Given that typical feature scales are 10-100nm per ordering, we need $40 - 400$ lattice sites in each dimension. For the parameters selected, the double-well driving phase separation, $(\phi^2 - \frac{\tau}{\lambda})^2$ has minima at $\phi_{min} = \pm\sqrt{6}$, indicating that the boundary between nanoparticle and block copolymer should transition from 0 inside the nanoparticle to ≈ 2.44 just outside. The interfacial penalty associated with a single set of adjacent interfacial cells is $6\rho_0 \left(\frac{40}{2\pi}\right)^2 = 240000kT$. For realistic systems, we would need to represent the nanoparticle as more than a single lattice cell, 16 to 20 would be more reasonable. The result is that differences in the polymer composition between the two faces of the nanoparticle are amplified by a factor easily on the order of thousands of kT for system sizes barely large enough to be considered sufficiently resolved, and this amplification scales quadratically with increases in the system size. For small system sizes, the differences in neighboring lattice values is significant. But by increasing the number of lattice sites, and assuming the final ordered state will be a cosine wave, we decrease the difference only linearly with N, since $cos \approx 1 - \frac{x}{N}$, whereas the penalty is growing $\propto N^2$. What these coarse calculations suggest is that to accept a nanoparticle move with appreciable probability, we must simultaneously move the nanoparticle and redistribute the surrounding copolymer composition in a way that minimizes interfacial energy penalties. Such a cooperative move is difficult to predict and to program in such a way that satisfies detailed balance. At the present, the construction of such a move is the foremost impediment to developing this simulation technique further.

The advantage of the Monte Carlo approach is that the volume exclusion is easily enforced by the appropriate choice of Monte Carlo moves. Its disadvantage is that creating a move with reasonable acceptance probability is difficult. An alternative approach is to assign a continuous density to each nanoparticle and introduce a coupling between polymer and nanoparticle into the Hamiltonian that repels polymer from the interior of each nanoparticle. Additionally, a halo region (denoted $h(r)$) surrounding each nanoparticle that couples linearly to either the A-rich or B-rich region of block copolymer, as in 3.15.

$$\mathcal{H}_{cpl} = \gamma \int dr\phi(r)^2\rho(r) + \nu \int dr\phi(r)h(r) \tag{3.43}$$

$\rho(r)$ is chosen to be a symmetric, infinitely differentiable function that vanishes outside a finite interval from the center of the nanoparticle, such as 3.16. The differentiability condition ensures exponential convergence of the Fourier series, making it suitable for the pseudospectral techniques from the previous section. The $\phi(r)^2$ term is minimized by setting $\phi(r) = 0$ wherever $\rho(r)$ is nonzero, while the linear coupling between $\phi(r)$ and the halo region $h(r)$ is attractive for $\phi(r) > 0$, corresponding to an A-rich region attraction, and repulsive for B-rich regions, where $\phi(r) < 0$. In Figure 3.13 we show a snapshot of the density fields from a representative simulation to illustrate the newly defined fields.



(a) (b)

Figure 3.13: Continuous (a) nanoparticle density, $\rho(r)$, and (b) halo attraction, $h(r)$, fields introduced into Langevin simulations to couple nanoparticles to polymer field ϕ.

This strategy is inspired by assuming that each particle imparts a small but finite compressibility in its interior, and then considering the limit as this compressibility vanishes. Hence, the $\delta[]$-constraint in the partition function 3.5, inside each nanoparticle, is replaced by a finite, quadratic compressibility, $\gamma \int dr\phi(r)^2\rho(r)$. Since this term is also a delta sequence when viewed as a sequence as $\gamma \to \infty$, the term will converge towards the correct $\delta[]$-constraint in the original model formulation.

$$e^{-\gamma \int dr\phi(r)^2\rho(r)} \to \prod_{r_{in}} \delta\left(\phi(r_{in})\right) \qquad \gamma \to \infty,\ \rho(r_{in}) > 0 \tag{3.44}$$

Thus, such an approach is theoretically well-behaved. Since the compressibility term favors $\phi(r_{in}) = 0$, the most energetically favorable way to accommodate this condition in the absence of a halo region is for the particle to localize to the interface between A-rich and B-rich regions, where $\phi(r)$ naturally tends towards 0. Without a halo, this behavior is easily reproduced in simulations on this model. Representative configurations are given in Figure 3.14.

When a halo region is present, the favorable energy gained from A-halo interactions may be sufficient to drive the nanoparticle into the interior of an A-rich region. In this case, two things may happen. Either the block copolymer will vacate the interior of the nanoparticle, paying an interfacial penalty for excluding a volume inside the A-rich region, or the block copolymer density will remain non-zero inside the nanoparticle, avoiding the interfacial penalty from the gradient terms in the Hamiltonian, but paying an energetic penalty associated with a nonzero ϕ value inside the nanoparticle due to the compressibility term. For the same reasons as discussed for Monte Carlo simulations, the interfacial penalty associated with creating an excluded volume in the interior of an A-rich region is enormous. As a result, for volume to spontaneously be excluded, the energetic penalty from compressibility must be greater than from interfacial terms, mean compressibility parameter γ must be on the same scale as the interfacial penalty (thousands of kT). As a result, the strength of the halo

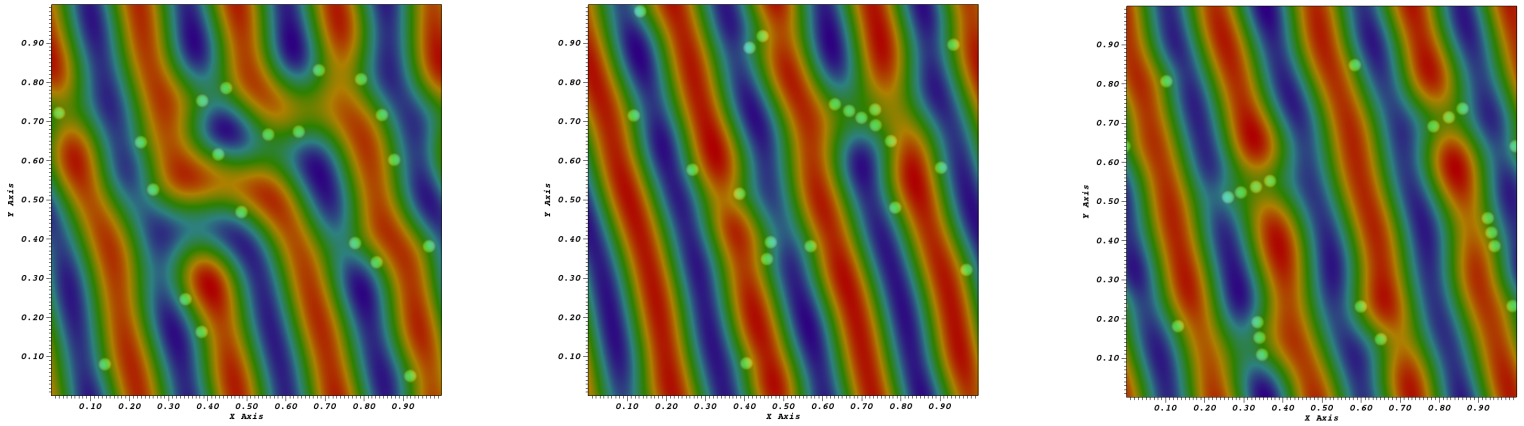


Figure 3.14: Late-time assembly trajectory for a corrected volume exclusion nanocomposite with $f = 0.5,\ q_0 = 4.0,\ \tau = 1.0,\ \epsilon = 2.0,\ \lambda = 1.0,\ \gamma = 10.0$

region must also be roughly this large in order for the particle to localize into the interior of a phase, instead of sitting at the A-B interface.
Due to the extreme energetic scales necessary to actually acheive volume exclusion, an extremely small time step Δt is necessary for the pseudospectral Langevin simulations, effectively putting meaningful assembly simulations out of reach. In the parameter regimes used in this thesis, the timestep Δt must be decreased by a factor of $10,000$ in order for the scheme to be stable. Representative snapshots obtained using this technique are shown in Figure 3.15.

There are two possible approaches to remedy this situation that warrant further investigation. The first is to use a higher order integrator for time stepping. The integrator used in this thesis is an $\mathcal{O}(\Delta t)$ accurate Euler time stepping algorithm. Higher order algorithms are available [21], which are essentially semi-implicit analogues of the familiar (explicit) Runge-Kutta integrators. In general, these algorithms require a linear increase in computation at each time step for a geometric increase in accuracy. For instance, the $4-5$ Runge-Kutta scheme requires evaluation of $\frac{\partial \phi}{\partial t}$ at 5 separate times, and yields an update $\phi(t+\Delta t)$ that is accurate up to $\mathcal{O}(\Delta t^4)$, allowing for a much bigger time step. A second possible solution is to enforce the volume constraint in the Langevin equations by using a Lagrange multiplier, similar to rigid body constraints in molecular simulations [1]. This approach would amount to adding a Lagrange multiplier to the equation of motion at each lattice site neighboring a volume-excluded nanoparticle to ensure that the BCP density gradient is normal to the nanoparticle surface, so that no polymer flows into the excluded volume. A similar approach has been successfully applied in the context of Cell Dynamical Simulations for a similar nanoparticle-polymer system [44], however it is unknown whether this is compatible with the semi-implicit, pseudospectral time-marching scheme used in this study.

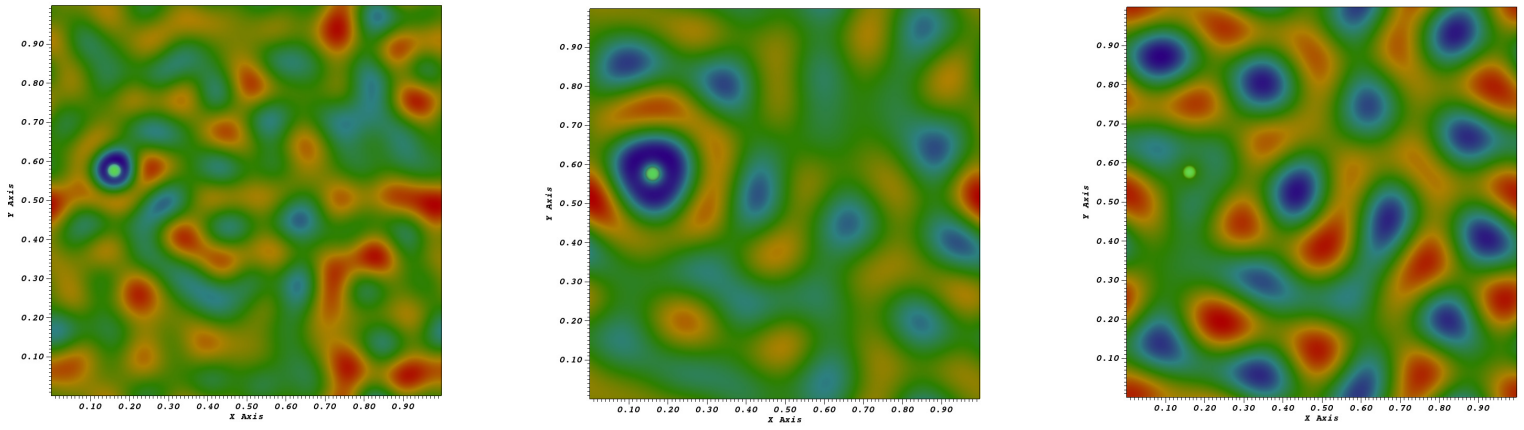


Figure 3.15: Assembly trajectory for a corrected volume exclusion nanocomposite with $f = 0.5$, $q_0 = 4.0$, $\tau = 1.0$, $\epsilon = 2.0$, $\lambda = 1.0$, $\gamma = 10000$, $\nu = 10000$. Inside the nanoparticle, the polymer field vanishes as required for volume exclusion.

Chapter 4

Algorithms for GPU Accelerated Simulations of Mesoscale Systems

Moore's Law famously predicts that transistor counts will follow an exponential increase, roughly doubling every two years, with chip performance doubling every 1.5 years (Figure 4.1). While this prediction has proven remarkably accurate since its inception in the early 1970s, within the past ten years these gains have more frequently come not from increased clock speeds or more efficient chip architectures, but from the production of processors that combine multiple execution units together onto a single die. This trend is expected to continue, since the limits of thermal power density seem to have been reached with current processor designs, leaving the addition of more independent cores as the only reliable way to increase performance. As a result, there has been a renewed interest in the development of algorithms in statistical physics that exploit massively parallel design principles via shared memory (multicore processors) and distributed memory paradigms (commodity clusters) due to the increased availability of computing resources at low cost. One particular option, general purpose computing on graphics processing units (GPGPU) is especially attractive due to theoretical GFLOP/s rates frequently 10-100 fold greater than traditional state of the art CPUs.

Here we present the development of an algorithm for Metropolis Monte Carlo lattice simulations on the GPU, and its application to the Ising Model and more general $\phi^4 - \phi^2$ Hamiltonians, such as 3.10. We demonstrate that the algorithm is able to successfully capture the physics of those systems and we quantify computational advantages of them.

GPU Architecture and Its Implication for Algorithm Development

The large demand for graphics processing power in consumer industries (video gaming, video processing and rendering, etc) has allowed the research and development of increasingly powerful graphics processors. Since many graphics processing tasks can be formulated as simple operations (i.e. translation and rotation) on arrays of vertices, a highly parallel architectural paradigm has evolved for graphics processors, in which a relatively simple (by

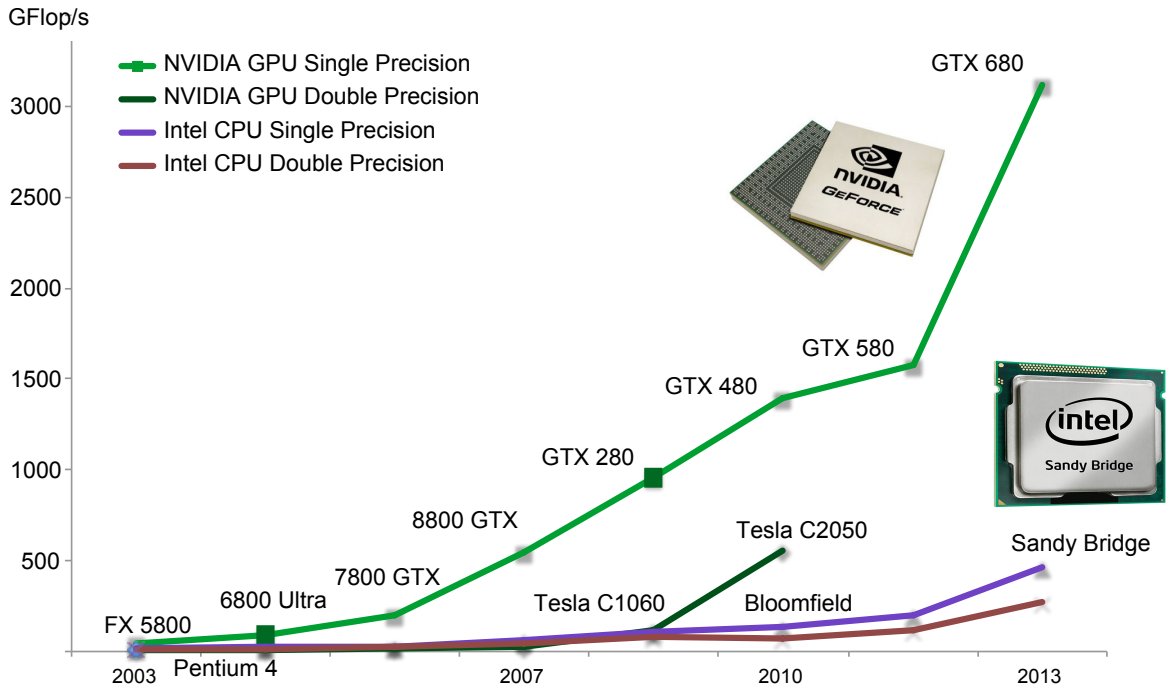


Figure 4.1: Comparison of Nvidia GPU and Intel CPU architecure timeline.

CPU standards) memory layout and instruction set provide enough memory bandwidth to fully saturate a vast array of independent compute elements with data for manipulation. As a result, graphics processors excel at transforming extremely large data sets, provided the transformation can be structured in such a way that exploits the computation-intensive and data-independent design of the processor.

Previously, to employ a graphics processor for scientific purposes, it was necessary to disguise one's computation in the language of graphics manipulations using a graphics API (OpenGL, for example) [123]. In 2006, Nvidia introduced CUDA (Compute Unified Device Architecture) [84], which included a set of extensions to the C programming language which expose Nvidia graphics hardware via a thread-based programming paradigm. A specialized compiler allows code to be compiled with these extensions and eventually run on Nvidia hardware. Reported speedups of 10-fold to many hundred-fold in the statistical physics community are common (although the appropriate metric for measuring speedup is somewhat murky, and discussion of Amdahl's rule is frequently avoided).

The code in this thesis was developed on a GeForce GTX 275. Since the particular architecture of the GeForce GTX 275 is already outdated by the time of this writing (by two major hardware revisions, Fermi and Kepler, no less), the extended architectural details of the 275's hardware will be omitted. Instead, it will suffice throughout the discussion to refer to Figure 4.3, which gives a brief overview of the Nvidia GPU architecture at a level that has remained relatively consistent throughout hardware revisions.

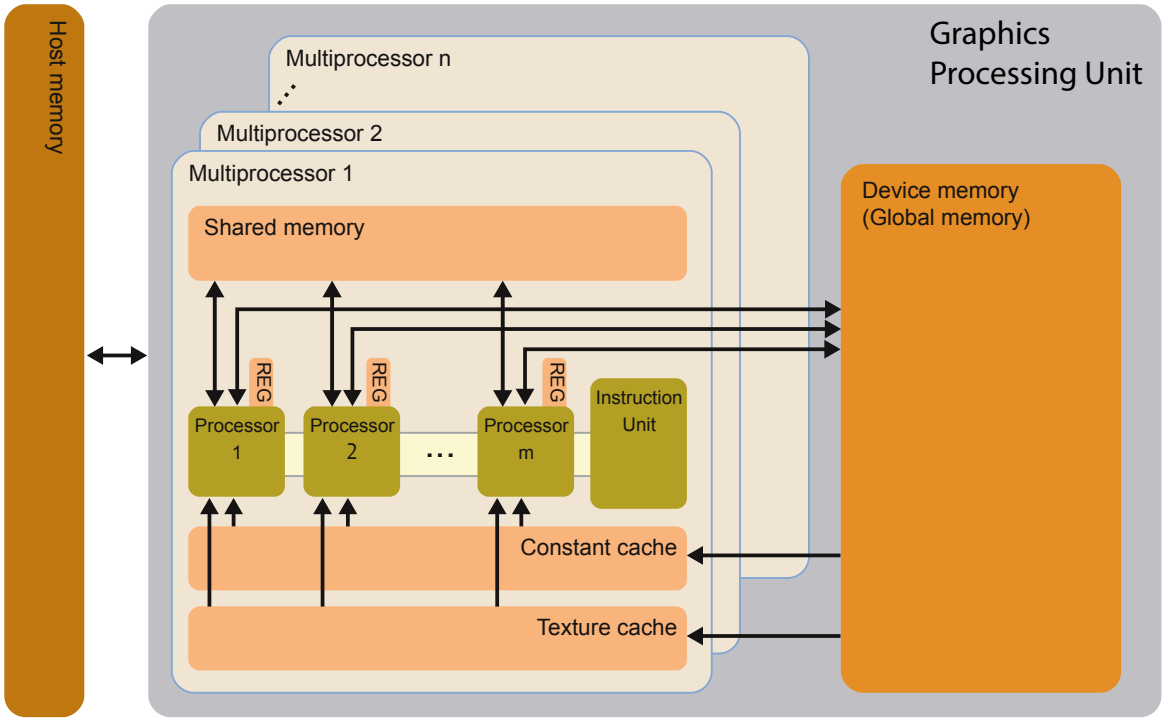


Figure 4.2: Schematic for the Nvidia GPU architecture. The GPU consists of many independent multiprocessors that communicate via global memory and caches. On each multiprocessor, individual processors share an instruction set and communicate via smaller, faster, on-chip shared memory.

The GPU is made up of a collection of multiprocessors, each one itself composed of a number of invidivdual compute elements capable of concurrent execution. Figure 4.3 gives a more detailed view of one multiprocessor. Each compute element has its own private register memory space, and can communicate with a global memory pool either directly or through constant and texture caches. Additionally, the compute elements in the same multiprocessor are able to communicate with one another via a small (tens of kBs in size), fast, shared memory space. Different multiprocessors execute independently from one another, meaning the compute elements in one multiprocessor cannot rely on, nor can they effectively synchronize with, other multiprocessors. Shared memory allows a meager amount of coordinated computation within a single multiprocessor, and threads of computation running one a multiprocessor can synchronize. Due to this design, increased parallelism and compute power can always be acheived by adding more multiprocessors.

To further abstract hardware details such as multiprocessor count and compute power, the CUDA programming model defines an abstract collection of threads, each representing a virtual compute element. Threads are collected into grids, with a grid representing a virtual

multiprocessor, and grids are arranged in a grid block so that each grid can access its relative position in the overal pool of grids. Thus, the needed parallelism may be tailored to the needs of the application, with highly efficient codes being those that spawn massive grid blocks with a high amount of computation per thread (and per global memory access).

4.1 Ising Model

The Ising model is one of the most studied models in statistical physics, due to its ability to describe a diverse array of phenomena (of which ferromagnetism was the first), and since its exact solution is known in 1 and 2 dimensions [11]. A well known fact of the Ising model is that it manifests a continuous phase transition in 2 dimensions and higher, but not in 1 dimension [11]. As the critical point associated with this phase transition is approached, long range correlations encourage a critical slowing down in traditional single spin flip Monte Carlo techniques [5]. To circumvent this, a large number of alternative numerical techniques have been developed to study the Ising model, making the Ising model an excellent starting point for prototyping new numerical techniques due to the wealth of information that exists on this model. For this reason, we chose to implement the Ising model as a first proof of concept for our GPU algorithm.

Model and Implementation

Tomov and coworkers were one of the first groups to implement a Monte Carlo Ising model algorithm on the GPU [123]. Their work utilized NVIDIA's Cg library, which was state of the art at the time. Due to the difficulty of programming with Cg, it was not until Pries et al [94] in 2009 reported a similar algorithm using NVIDIA's more accessible CUDA interface that widespread interest grew in GPUs for simulation. Following [94], we model the simple two dimensional ferromagnetic square lattice Ising model, which consists of an arrangement of ± 1 spins arrange on a square lattice with periodic boundary conditions to mimic a macroscopic system. These spins interact with their nearest neighbors on the lattice with a Hamiltonian given by

$$\mathcal{H} = -J \sum_{<i,j>} S_i S_j - H \sum_i S_i \tag{4.1}$$

where $S_i = \pm 1$ represents a spin at site i and H denotes the external magnetic field. We use the standard Metropolis criterion for single spin flips[82]

$$p_{acc}(S_i \rightarrow -S_i) = \min[1, e^{-\beta \Delta E}] \tag{4.2}$$

The advantage of single spin flip dynamics is that neighboring states on the Monte Carlo chain differ at most by a single spin, so that $\Delta\mathcal{H}$ can be calculated by examining the proposed

spin to flip and its neighboring spins. Since this depends only on the spin on the given lattice site as well as its 4 immediate neighbors, the necessary calculations can be made in a highly parallel fashion by using the well known checkerboard lattice decomposition. For an $N \times N$ lattices of $n = N^2$ total spins, the lattice is divided into two sublattices with the property that no lattice site in a sublattice is a neighbor of any other site in that sublattice. If all of the sites in one sublattice are updated with a single spin flip exactly once, then the probability of observing the system in state $\mathbf{s}^1$ after updating one sublattice, decomposes according to

$$P(\mathbf{s}^1|\mathbf{s}^0) = P(\mathbf{s}_0)\prod_{j=1}^{N/2} P(\mathbf{s}^j|\mathbf{s}^{j-1}) = \tag{4.3}$$

$$= P(\mathbf{s}^0)\prod_{j=1}^{N/2} P(\mathbf{s}^j|s^{j-1}_{\sigma(j-1)}) = \tag{4.4}$$

$$= P(\mathbf{s}^0)\prod_{j=1}^{N/2} P(\mathbf{s}^j|s^{0}_{\sigma(j-1)}) = \tag{4.5}$$

$$= P(\mathbf{s}^0)\prod_{j=1}^{N/2} P_{gen}(s^j_{\sigma(j-1)})P_{acc}(s^j_{\sigma(j-1)}) \tag{4.6}$$

Since the final probability 4.6 is invariant under the permutation σ chosen, the order in which the moves are executed is irrelevant, and we can make the moves in any order, and even in parallel, with one processor updating one lattice site. We can synchronize the processors after 1 Monte Carlo step ($n/2$ total), update the second sublattice of $n/2$ sites, and repeat until equilibrium is reached. One caveat is that due to the deterministic way in which lattice sites are selected, the Markov chain is periodic at $T = \infty$ and thus not ergodic, since each sublattice's spins are simply flipped in parallel.

The checkerboard decomposition is a reasonably efficient and relatively simple way to exploit parallelisn in the Ising model, but it does not fully exploit the idiosyncracies of the Nvidia memory hierarchy, in which individuals 'processors' on the GPU are grouped into 32-wide multiprocessors that share a local, cached memory space. Directly implementing the checkerboard decomposition does not utilize this memory space, resulting in poor memory bandwidth, since all data reads and writes are directly to/from the global memory space, which has much lower latency. Further, each sublattice update requires a separate kernel invocation, introducing even more overhead.

To remedy this problem, the entire lattice is first divided into a larger array of blocks, with each block separated from its neighbors by a boundary region, as shown in Figure4.3. As a consequence of this decomposition each block can be updated independently from the others for an arbitrary number of Monte Carlo steps, provided the boundary regions are never updated. The size of the blocks is chosen so that multiple blocks can be fit into the shared memory space of a multiprocessor, and then each block may be collectively loaded

into shared memory by the threads in the thread block. Each block is further decomposed into a standard checkerboard decomposition, which allows for all spins inside the block to be iteratively updated in the typical half-half checkerboard fashion for a number of times before synchronization with the other blocks occurs.

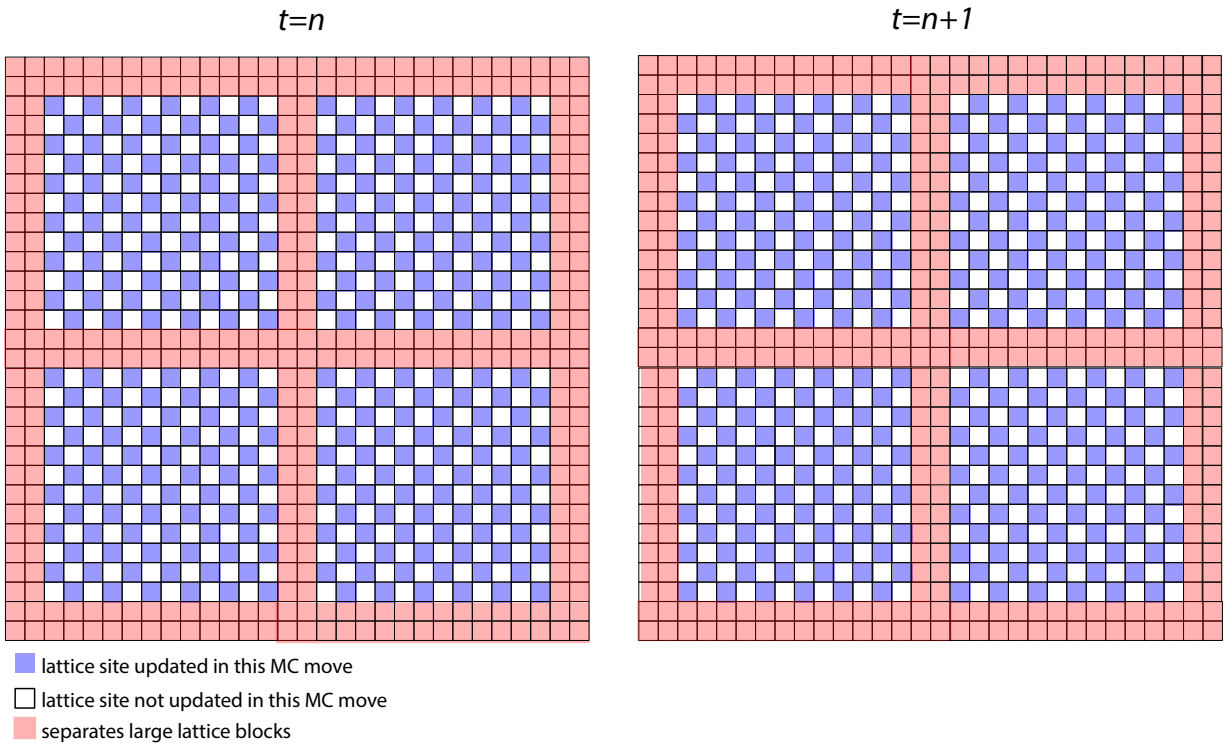


Figure 4.3: Ising model decomposition. The lattice is broken up into distinct regions separated by boundary regions that are never updated. Within each region, a checkerboard decomposition is made. A block of threads then iterates through a single checkerboard sublattice. In the case of the Ising model, the checkerboards have 2 sublattices, and a block alternates MC proposals on each sublattice in every step. After a user-chosen number of updates, the new values for each region's spins is written to global memory, the boundary regions are shifted and the process repeats.

In practice we first launch a kernel which assigns half (ever other one) of the large tiles to a separate thread block. All of the thread blocks will be executed in parallel. Then the threads of each thread block cooperatively load the spin configuration of their tile plus a boundary layer, necessary to evaluate acceptance criteria, into shared memory. subsequently the threads of each block perform a Metropolis update of each lattice site type (e. g blue) in their tile in parallel and then all the threads of each block are synchronized, ensuring that all of them have completed the previous step. In the next step the tiles that have not been previously updated (in this case white) are updated and the threads of each block perform

a Metropolis update on each site in their tile in parallel and again the threads of each block are synchronized. Finally another kernel is launched to evaluate the remaining large tiles of the checkerboard in the same fashion.

4.2 Block Copolymer

In this section we modify the parallel algorithm described for the Ising model to work with the block copolymer Hamiltonian, 3.10. The principal difference between 3.10 and the Ising Hamiltonian is the presence of gradient terms that couple further than nearest neighbor lattice sites together, requiring a bigger footprint for each Monte Carlo move. Additionally, the order parameter field $\phi(r)$ is continuous and locally conserved, requiring an even bigger footprint, since a local transfer of order parameter density from neighboring lattice sites must be made. Straightforward application of a checkerboard decomposition to the block copolymer model would require 8 distinct sublattices and require a heroic programming effort to even implement. Further, the extremely low efficiency associated with 1 lattice site for 8 memory transfers would make checkerboard decomposition a very unappealing solution. This algorithm avoids the problems inherent in the traditional checkerboard algorithm by efficiently exploiting the GPU memory hierarchy to maximize memory transfers per lattice update.

In this section we approximate gradients by finite differences $\nabla^4\phi \to \delta_k^4[\phi]$. The stencil used for $\delta_k^4[\phi]$ will later determine the sublattices for the parallel checkerboard.

Briefly, we define a notional superlattice of subsystems separated by sufficient distance that they do not directly interact i.e we first decompose the entire $N \times N$ lattice into $(N/L) \times (N/L)$ lattice of $L \times L$ blocks. Each of these regions is assigned to a separate multiprocessor (block of threads), which performs many MC sweeps, advancing moves locally with the intervening boundary regions held fixed. Each block is decomposed into noninteracting sublattices and a parallel checkerboard algorithm is used. Density transfer between adjacent cells is proposed for cells on the current (blue, Figure 4.4) sublattice and accepted according to standard Metropolis criteria. Different sublattices are sequentially chosen, until the entire block is equilibrated. The edges are not updated in this process. After the entire block is equilibrated the the gridlines defining the $L \times L$ blocks (red in Figure 4.4) are shifted by $L/2$ in the x and/or y direction, ensuring all edges eventually are equilibrated.

The advantage of this approach lies in the ability to choose a size of subsystems large enough to exploit efficient memory access (so-called coalesced memory transactions) yet small enough to fit inside of fast memory pools on the GPU. Such strategies should scale well between different hardware types and even when new generations of hardware comes out.

Results

To test the new algorithm, we first compare the lamellar and cylindrical morphologies of symmetric diblock copolymers obtained from simulations with those observed in experiments. As shown in Figure 4.5, simulated morphologies match qualitatively the experimental results including both the two-dimensional lamellae (Figure4.5A) and cylindrical structures (Figure4.5B).

To quantify the speedup of the GPU code relative to CPU code, we approximate it as the ratio of the MC step rate on GPU/CPU and then measure the speedup as a function of lattice size. As shown on Figure4.6A the speedup increases as the lattice size increases but tends to plateau for very large lattices (around $N = 2000$). Another, perhaps more functionally relevant, way to measure the performance of GPU code is to track the number of Monte Carlo steps necessary to reduce the spin-spin autocorrelation, $C(t) = \langle s_i(t)s_i(0)\rangle$ to 1% of its initial value at a single parameter set above the transition temperature. The ratio of CPU/GPU values gives an idea of how efficiently the GPU scheme can generate decorrelated microstates. As shown in Figure4.6B, the decorrelation ratio decreases almost exponentially with increasing lattice size as the GPU code requires fewer MC steps to decorrelate the density field. This demonstrates clearly the benefits of the local GPU algorithm for the block copolymer model.

Discussion

We have quantified the speedup by looking at the ratio of the MC step rate on GPU/CPU of the the number of Monte Carlo steps necessary to reduce $C(t)$ to 1% of its initial value at one parameter set above the transition temperature. While those measures reflect the speedup closely an even better quantification of speedup would measure the system's decorrelation rate, in real seconds.

The speedup afforded by the GPU code depends on how many MC steps are taken before switching grid boundaries. Nevertheless, it robustly shows that the bigger the lattice, the slower the CPU code but faster the GPU code, highlighting the ability of GPU algorithms to extend MC simulations of polymeric systems to larger and more complex systems. The algorithm performs all computations in shared memory, a fast, but small (16kB) memory space on each MP, which is the most likely reason behind the extraordinary speedup, since the GPU experiences significant memory latency only when shifting gridlines (after millions of MC steps have been completed).

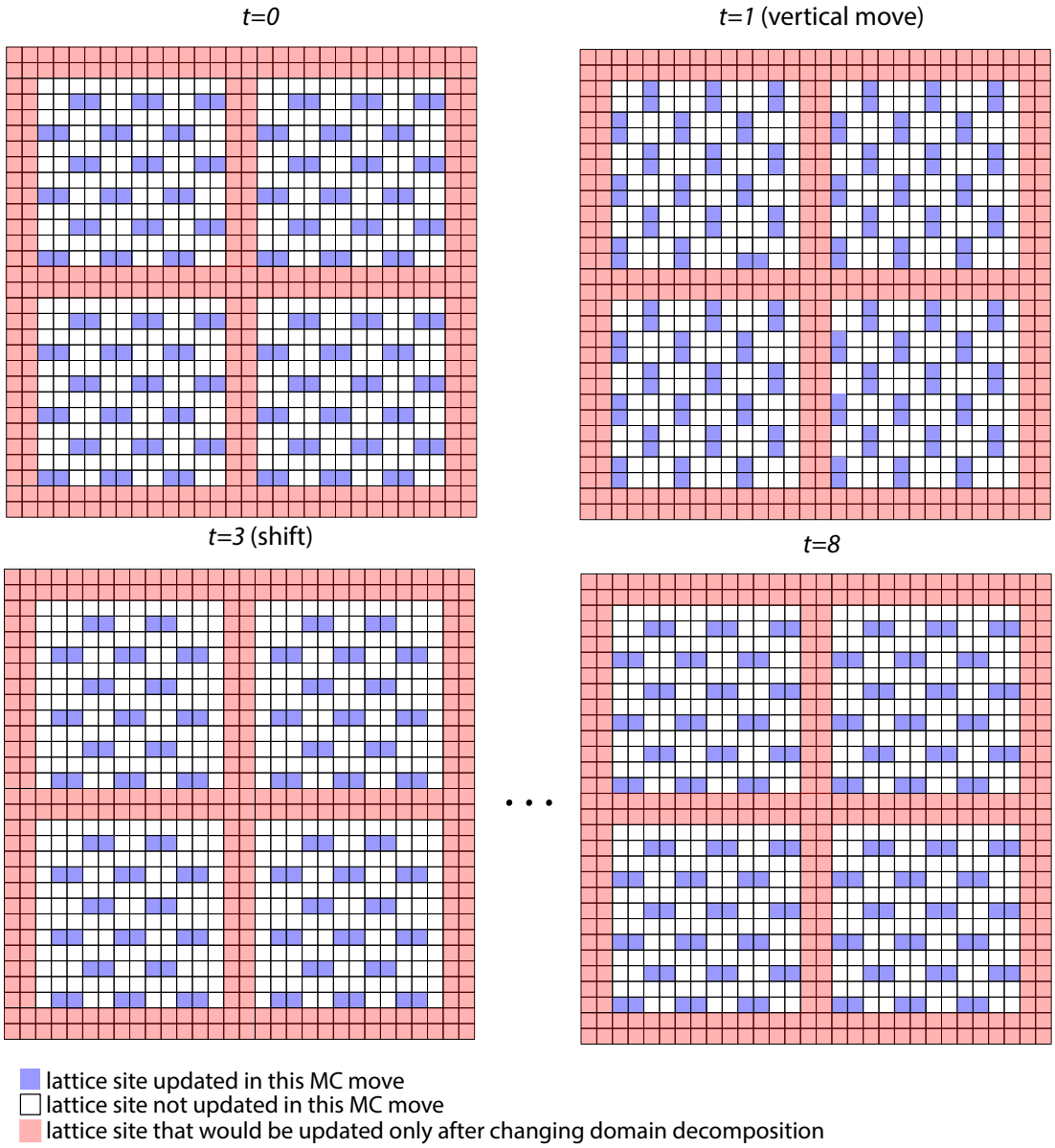


Figure 4.4: Similar to the Ising model procedure in Fig. 4.3, except that block copolymer updates require the local transfer of density between neighboring lattice sites to maintain local conservation of density. This necessitates the more complicated lattice decomposition shown.

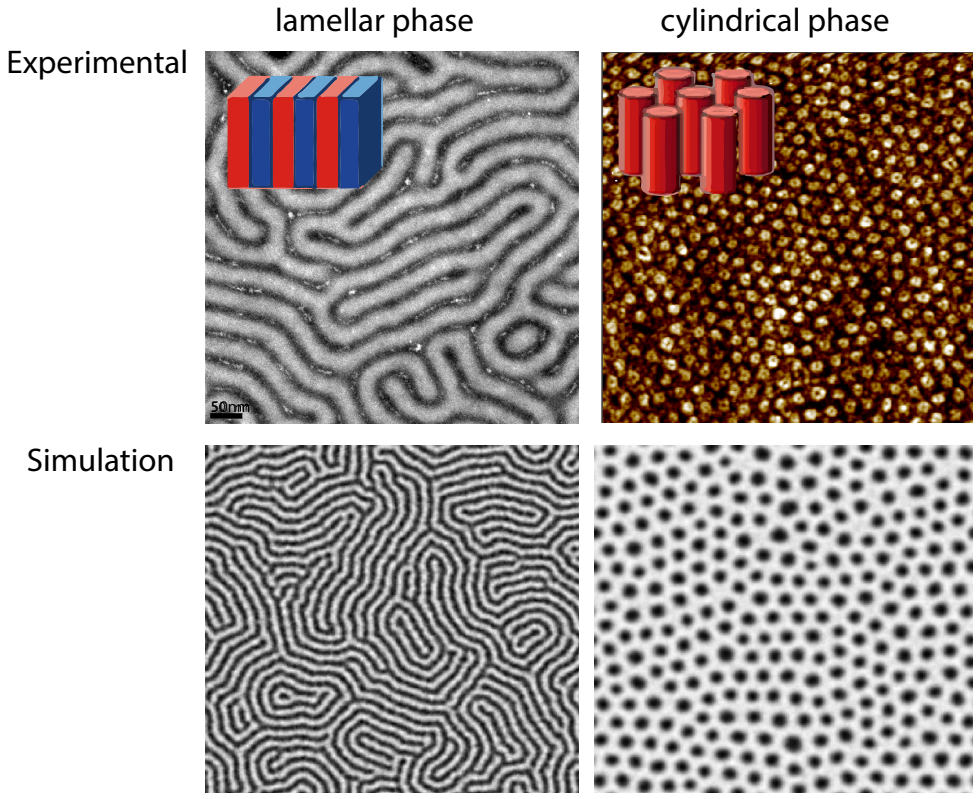


Figure 4.5: Self assembly results for GPU block copolymer code.

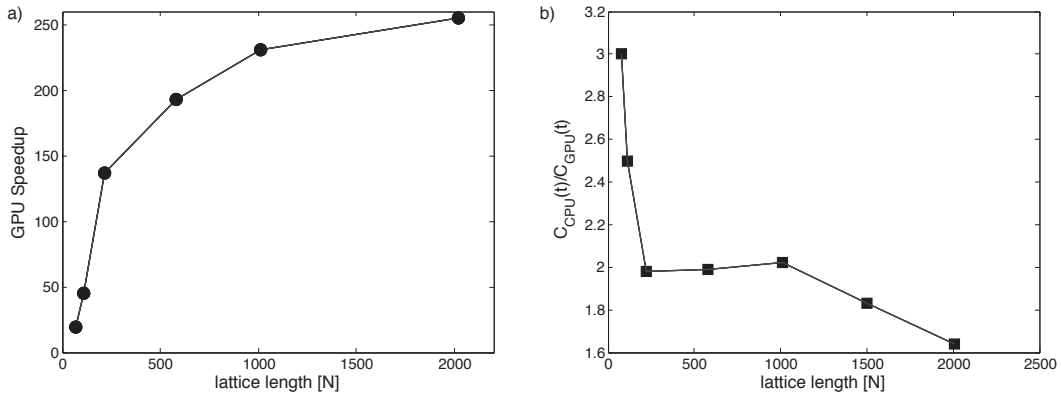


Figure 4.6: Benchmark of GPU code vs CPU code. (a) The amount of GPU speedup rapidly increases for larger lattices. At some point, the multiprocessors become saturated and this increase tapers off. (b) Due to the uniform updates of lattices sites across the system (rather than at randomly selected sites), the local lattice regions decorrelate much faster on the GPU than on the CPU.

Chapter 5

Excluding volume in field theories

The central goal of Chapter 3 was to develop a model for a block copolymer nanocomposite and then to implement an efficient numerical scheme to investigate the self assembly behavior of the model over many wavelengths of ordering. At the time, we assumed that volume exclusion leaves the $\phi^4 - \phi^2$ model intact and only modifies the model insofar as it sets $\phi(r) = 0$ inside regions containing a nanoparticle. In this chapter we investigate the validity of this assumption. In addition, we exploit the fact that the ϕ^4 term is an approximate term that restores stability to an unstable Gaussian model, and we introduce a hard-constraint approximation that fulfills the same role, while making much more transparent the connection between our theory and Gaussian theories for liquids, the density statistics of which have been extensively studied in the context of volume exclusion.

5.1 The Gaussian case

For a simple liquid of a single particle type with an intermolecular potential $u(|r_i - r_j|)$, we can apply the inverse Hubbard-Stratonovich transformation to convert a particle theory into a field theory in which we integrate over all density fluctuations, similar to the procedure leading to 3.7 [35]. If we are only interested in density fluctuations in the liquid far away from a critical point, then we may make a harmonic approximation to the statistics of the fluid. In this regime, density fluctations in the liquid are observed according to a Gaussian weight, with partition function

$$Z_G = \int \mathcal{D}\rho e^{-\frac{1}{2}\int \rho(r)\chi^{-1}(r-r')\rho(r')} \tag{5.1}$$

As always, the functional integral in (5.1) implies a limiting process from a discrete set of coupled Gaussian random variables to a continuum. For the purposes of excluding volume, it will be easier to work directly within a discrete framework, where the partition function is approximated as the normalization of a discrete (albeit quite large), multidimensional Gaussian integral. We therefore discretize space onto a Euclidian grid, with one fluctuating

variable for each discrete lattice site. In general, we will denote fluctuating variables with $\mathbf{x}'s$, means with $\boldsymbol{\mu}'s$, covariance matrices with $\boldsymbol{\Sigma}'s$, and precision matrices (inverses of covariance matrices) with $\boldsymbol{\Lambda}'s$. Note that $\boldsymbol{\Sigma}^{-1} = \boldsymbol{\Lambda}$. For a simple liquid, when we exclude volume, we need to set a portion of the fluctuating variables to 0. It will be convenient to block partition the variables into freely fluctuating variables as $\mathbf{x}_f$, and constrained variables, $\mathbf{x}_c$.

In terms of $\mathbf{x}$, the probability of a configuration is simply

$$p(\mathbf{x}) = \mathcal{N}(\mathbf{x}|\mu, \Sigma) = (2\pi)^{-N/2} \left|\Sigma\right|^{-1/2} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Lambda} (\mathbf{x}-\boldsymbol{\mu})} \tag{5.2}$$

This probability can also be written in terms of the partitioned variables as

$$p(\mathbf{x}) = p(\mathbf{x}_f, \mathbf{x}_c) = (2\pi)^{-N/2} \left|\Sigma\right|^{-1/2} e^{-\frac{1}{2}\left(\delta\mathbf{x}_f^T \Lambda_{ff} \delta\mathbf{x}_f + \delta\mathbf{x}_c^T \Lambda_{cc} \delta\mathbf{x}_c + 2\delta\mathbf{x}_f^T \Lambda_{fc} \delta\mathbf{x}_c\right)} \tag{5.3}$$

In 5.3 we have used $\delta\mathbf{x} = \mathbf{x} - \boldsymbol{\mu}$ and the block matrix

$$\boldsymbol{\Lambda} = \begin{pmatrix} \boldsymbol{\Lambda}_{ff} & \boldsymbol{\Lambda}_{fc} \\ \boldsymbol{\Lambda}_{cf} & \boldsymbol{\Lambda}_{cc} \end{pmatrix} \tag{5.4}$$

We know that the density vanishes in excluded volume regions, which necessitates $\mathbf{x}_c = 0$. Therefore, we seek an expression for the conditional probability of the remaining unconstrained variables $\mathbf{x}_f$, given that $\mathbf{x}_c = 0$. Bayes's Theorem and a little algebra show that

$$p(\mathbf{x}_f|\mathbf{x}_c) = \frac{p(\mathbf{x}_f, \mathbf{x}_c)}{p(\mathbf{x}_c)} = \frac{p(\mathbf{x}_f, \mathbf{x}_c)}{\int d\mathbf{x}_f \; p(\mathbf{x}_f, \mathbf{x}_c)} \tag{5.5}$$

$$= \frac{e^{-\frac{1}{2}(\delta\mathbf{x}_f - \boldsymbol{\Lambda}_{ff}^{-1}\boldsymbol{\Lambda}_{fc}\delta\mathbf{x}_c)^T \boldsymbol{\Lambda}_{ff} (\delta\mathbf{x}_f - \boldsymbol{\Lambda}_{ff}^{-1}\boldsymbol{\Lambda}_{fc}\delta\mathbf{x}_c)}}{\int d\mathbf{x}_f \; e^{-\frac{1}{2}(\delta\mathbf{x}_f - \boldsymbol{\Lambda}_{ff}^{-1}\boldsymbol{\Lambda}_{fc}\delta\mathbf{x}_c)^T \boldsymbol{\Lambda}_{ff} (\delta\mathbf{x}_f - \boldsymbol{\Lambda}_{ff}^{-1}\boldsymbol{\Lambda}_{fc}\delta\mathbf{x}_c)}} \tag{5.6}$$

The denominator in 5.6 is simply the normalization for the Gaussian exponential in the numberator, meaning the unconstrained variables $\mathbf{x}_f$ retain their Gaussian character, with a new mean and variance determined in part by the constrained modes,

$$p(\mathbf{x}_f|\mathbf{x}_c) = \mathcal{N}(\mathbf{x}_f|\boldsymbol{\mu}_{f|c}, \boldsymbol{\Sigma}_{f|c}) \tag{5.7}$$

$$\boldsymbol{\mu}_{f|c} = \boldsymbol{\mu}_f - \boldsymbol{\Lambda}_{ff}^{-1}\boldsymbol{\Lambda}_{fc}(\mathbf{x}_c - \boldsymbol{\mu}_c) \tag{5.8}$$

$$\boldsymbol{\Sigma}_{f|c} = \boldsymbol{\Lambda}_{ff}^{-1} \tag{5.9}$$

Note that $\boldsymbol{\Sigma}_{f|c} \neq \boldsymbol{\Sigma}_{ff}$, the original covariance among $\mathbf{x}_f$ when $\mathbf{x}_c$ is unconstrained. 5.7-5.9 are general results for any choice of $\mathbf{x}_c$, but for excluded volume we know that $\mathbf{x}_c = 0$. Further, both $\boldsymbol{\mu}_f$ and $\boldsymbol{\mu}_c$ are equal to the bulk density $\bar{\rho}$. Using these facts, and assuming there are N_f free variables and N_c constrained variables, gives

$$\boldsymbol{\mu}_{f|c} = \boldsymbol{\mu}_f + \bar{\rho} \sum_{j=1}^{N_c} \mathbf{a}_j \tag{5.10}$$

$$\mathbf{A} = [\mathbf{a}_1 \dots \mathbf{a}_{N_c}] = \boldsymbol{\Lambda}_{ff}^{-1} \boldsymbol{\Lambda}_{fc} \tag{5.11}$$

$\mathbf{a}_j$ are the colums of the $N_f \times N_c$ matrix $\mathbf{A}$. Using the following matrix identity,

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{M} & -\mathbf{MBD}^{-1} \\ -\mathbf{D}^{-1}\mathbf{CM} & \mathbf{D}^{-1} + \mathbf{D}^{-1}\mathbf{CBD}^{-1} \end{pmatrix} \tag{5.12}$$

where $\mathbf{M}^{-1}$ is the Schur complement of the left hand side, along with a fair amount of algebra, we can show that

$$\boldsymbol{\Sigma}_{f|c} = \boldsymbol{\Sigma}_{ff} - \boldsymbol{\Sigma}_{fc} \boldsymbol{\Sigma}_{cc}^{-1} \boldsymbol{\Sigma}_{cf} \tag{5.13}$$

5.10 and 5.13 are the discrete analogues of the continuum results presented in [42] and [24]. In particular, the structure in 5.10 and 5.13 are highly similar to Equation 7.8 and Equation 7.7 in [42], where the partitioning of varibles into unconstrained and constrained variables is analogous to restricting the modified correlation function $\boldsymbol{\chi}^{(m)}$ to the subspace spanned outside of the hard solute via projection operators.

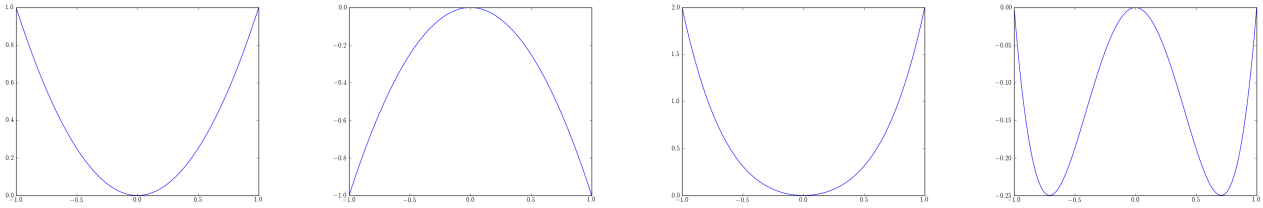
5.2 Excluding volume in ϕ^4-ϕ^2 field theories

Extending the results from the previous section to a $\phi^4 - \phi^2$ theory such as the block copolymer Hamiltonian 3.10 is relatively straightforward. Since the structure factor is computed by the random phase approximation, $\boldsymbol{\Lambda}$ is known, and $\boldsymbol{\Lambda}$ captures the bilinear terms in 3.10, essentially the quadratic and gradient portions, leaving the quartic term as the only addition to 3.10. Therefore, the probability for a configuration becomes

$$p(\mathbf{x}) = Z^{-1} \; e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Lambda} (\mathbf{x}-\boldsymbol{\mu}) + \lambda \sum_i^N x_i^4} \tag{5.14}$$

$$Z = \int d\mathbf{x} \; e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Lambda} (\mathbf{x}-\boldsymbol{\mu}) + \lambda \sum_i^N x_i^4} \tag{5.15}$$

In contrast to a Gaussian probability, the normalization Z cannot be evaluated in this case. But this is not a problem, since upon splitting $\mathbf{x}$ into $\mathbf{x}_f$ and $\mathbf{x}_c$, and an application of Bayes's rule, the normalizations cancel,



(a) (b) (c) (d)

Figure 5.1: Illustration of the breakdown of stability in a Gaussian field theory. (a) Above the transition temperature, the quadratic form is positive definite, and finite fluctuations are observed. (b) When $T < T_c$, the quadratic form becomes negative definite, leading to unbounded energy associated with larger density fluctuations. (c) For $T > T_c$ in the quartic model, stability is also observed. (d) For $T < T_c$ in the quartic model, stability is still observed due to the quartic restoring force, as seen by the energy minima at finite values.

$$p(\mathbf{x}_f|\mathbf{x}_c) = \frac{p(\mathbf{x}_f, \mathbf{x}_c)}{p(\mathbf{x}_c)} = \frac{p(\mathbf{x}_f, \mathbf{x}_c)}{\int d\mathbf{x}_f \; p(\mathbf{x}_f, \mathbf{x}_c)} \tag{5.16}$$

$$= \frac{e^{-\frac{1}{2}(\delta\mathbf{x}_f - \mathbf{\Lambda}_{ff}^{-1}\mathbf{\Lambda}_{fc}\delta\mathbf{x}_c)^T\mathbf{\Lambda}_{ff}(\delta\mathbf{x}_f - \mathbf{\Lambda}_{ff}^{-1}\mathbf{\Lambda}_{fc}\delta\mathbf{x}_c) + \lambda\sum_i^{N_f} x_i^4}}{\int d\mathbf{x}_f \; e^{-\frac{1}{2}(\delta\mathbf{x}_f - \mathbf{\Lambda}_{ff}^{-1}\mathbf{\Lambda}_{fc}\delta\mathbf{x}_c)^T\mathbf{\Lambda}_{ff}(\delta\mathbf{x}_f - \mathbf{\Lambda}_{ff}^{-1}\mathbf{\Lambda}_{fc}\delta\mathbf{x}_c) + \lambda\sum_i^{N_f} x_i^4}} \tag{5.17}$$

What is remarkable about 5.17 is that, similar to the Gaussian field theory case, the process of excluding volume does not fundamentally change the structure of the field theory. The probability distribution remains a $\phi^4 - \phi^2$ theory. Since the ϕ^4 term is entirely local, the constrained lattice sites $\mathbf{x}_c$ only influence the remaining lattice sites $\mathbf{x}_f$ through the Gaussian quadratic form, in exactly the same way as in the GFT case. What this means is that, to study the effects of volume exclusion in a $\phi^4 - \phi^2$ theory, it suffices to study the effect of volume exclusion independently from the effect of the quartic restoring force, when added to the Gaussian case, 5.3.

We now focus attention on the effect of the quartic term in 5.17. Since the $\phi^4 - \phi^2$ Hamiltonian is used as a model for phase separated systems, it is assumed that the quadratic form is negative definite. Physically, a negative definite quadratic form would reward evermore extreme density fluctuations and lead to unphysical configurations. To correct this, a quartic term is added to the Hamiltonian to provide a restoring force against extreme fluctuations. This view is illustrated in Fig. 5.1.

The local free energy density for a single lattice site rewards unbounded density fluctuations in the absence of a restoring force. When a quartic term is added, the resulting double-well potential has well-defined energy minimums. However, the choice of a quartic term to provide stability is not special. Formally, it represents the next term in a functional Taylor expansion of the field-theoretic Hamiltonian in powers of the density fluctuation field

(with odd powers assumed to vanish due to symmetry). In practice, however, even the ϕ^4 term is an approximation, since in an exact treatment the term would be of the form

$$\frac{1}{2!}\int_q \gamma_2(q,-q)\phi(q)\phi(-q)+\frac{1}{4!}\int_{q1}\int_{q2}\int_{q3}\gamma_4(q_1,q_2,q_3,-q_1-q_2-q_3)\phi(q_1)\phi(q_2)\phi(q_3)\phi(-q_1-q_2-q_3) \tag{5.18}$$

Despite this, it is standard to approximate the Hamiltonian's fourth order vertex function, γ_4 as a constant, $\lambda = \gamma_4(0,0,0,0)$. Upon inverse Fourier transform, this term becomes the familiar

$$\frac{\lambda}{4!}\int_{q1}\int_{q2}\int_{q3}\phi(q_1)\phi(q_2)\phi(q_3)\phi(-q_1-q_2-q_3) = \frac{\lambda}{4!}\int dr\phi(r)^4 \tag{5.19}$$

The quartic term $\int dr\phi(r)^4$ is necessary to restore stability to the negative definite quadratic form, but the particular choice of a quartic is merely a matter of convenience. However, as shown in Chapter 3, this quartic term introduces extreme headaches when put into a numerical scheme, necessitating expensive and less-accurate explicit updates for the pseudospectral technique, reducing its computational efficiency considerably.

In Fig. 5.2, we plot a family of polynomials of increasing order, to illustrate the basic fact that

$$\lim_{n\to\infty} e^{-\left(\frac{x}{x_{max}}\right)^n} = \theta(x_{max}-x)\theta(x_{max}+x), \tag{5.20}$$

with the exception of $x = \pm 1$. Given this fact, for a sequence of random variables X_n with probability measures

$$p_n(x) \propto e^{-\frac{1}{2}\mathbf{x}^T\mathbf{\Lambda}\mathbf{x}+\sum_i\left(\frac{x_i}{x_{max}}\right)^n}, \tag{5.21}$$

we know that X_n converges to a random variable X with probability measure

$$p_\infty(x) \propto e^{-\frac{1}{2}\mathbf{x}^T\mathbf{\Lambda}\mathbf{x}}\prod_i \theta(x_{max}-x_i)\theta(x_{max}+x_i) \tag{5.22}$$

Thus, the limitng case of an increasingly stiff restoring force against unphysical fluctuations is simply a hard constraint on the fluctuation that can occur. This is simply a truncated Gaussian distribution. Whereas the $\phi^4-\phi^2$ approach is motivated by reasons of simplicity, i.e. simply terminating the infinite order power series as soon as stability is restored to the Gaussian portion, this approach is motivated by the realization that there *must* be an infinitely stiff restoring force against physically unrealistic fluctuations that place more polymer density at a point in space than excluded volume arguments from a particle-based description would allow. In other words, the $\phi^4-\phi^2$ approach says unphysical density profiles are unlikely, but possible, whereas we know excluded volume necessitates that these density distributions can simply never be observed.

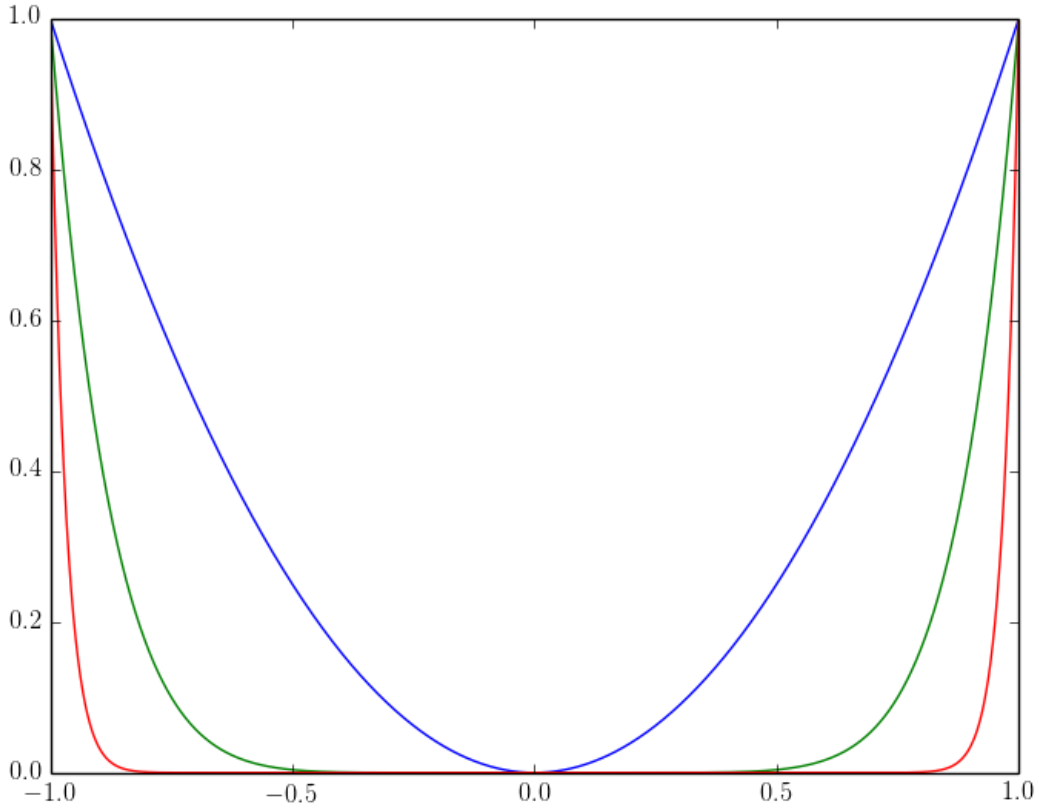


Figure 5.2: Family of polynomials of increasing order, showing convergence towards hard constraints, 5.1. (blue) x^2, (green) x^8, (red) x^{32}.

The advantage of the infinitely stiff restoring force approach is that the distribution that results is essentially a Gaussian distribution, restricted to a $d-$dimensional cube. Alternatively, by diagonalizing the exponent through a variable transformation, $\mathbf{y} = \Lambda^{1/2}(\mathbf{x}-\boldsymbol{\mu})$, the partition function can be expressed as uncoupled Gaussian integrals, restricted to a convex, $d-$dimensional polytope V_d,

$$Z = \int_{V_d} d\mathbf{y}\ e^{-\frac{1}{2}\sum_i \lambda_i y_i^2} \tag{5.23}$$

Whereas the integral for the $\phi^4 - \phi^2$ partition function is unsuitable for Gaussian quadrature, the partition function 5.23 and associated ensemble averages can, at least in principle, be iteratively solved via a computer algebra system for a closed form solution in terms of error functions.

5.3 Normal modes of the system

Due to the translational invariance of the physical system, the resulting correlation matrix, Σ is circulant. As a result, its eigenvectors are the complex Fourier exponentials,

$$\mathbf{\Sigma} = \mathbf{U}\mathbf{D}\mathbf{U}^T \tag{5.24}$$

$$\mathbf{U} = [\mathbf{v}_1 \dots \mathbf{v}_N] \tag{5.25}$$

$$\mathbf{v}_k = \left(1, \omega_k, \dots, \omega_k^{N-1}\right) \tag{5.26}$$

$$\omega_k = e^{\frac{2\pi i k}{N}} \tag{5.27}$$

Therefore, the Fourier modes, $\tilde{\mathbf{x}} = \mathbf{U}\mathbf{x}$, are also the normal modes of the system. Whereas box constraints on real-space density fluctuations independently limit each component of $\mathbf{x}$ to $-x_{max} < x_i < x_{max}$, the effects of these constraints on $\tilde{\mathbf{x}}$ are more complicated. Although the coupling between components of $\mathbf{x}$ is removed after diagonalizing, the support for $\tilde{\mathbf{x}}$ is now a convex polytope, meaning the feasibility of a component $\tilde{x}_k$ taking a value is coupled to the values of the other components in $\tilde{\mathbf{x}}$. As x_{max} is increased, the probability for $\tilde{\mathbf{x}}$ becomes Gaussian, as do all of the marginal probabilities $p(\tilde{x}_k)$. Therefore, it is interesting to consider the extent to which the marginal probability, 5.29, for the the real component (or imaginary, since they have identical statistics) of wavevector k^*, which describes the eventual ordering periodicity, is Gaussian at relevant values for x_{max}.

$$p(\tilde{\mathbf{x}}) = \int_{V_d} d\tilde{\mathbf{x}} e^{-\frac{1}{2}\sum_k |\tilde{x}_k|^2} \tag{5.28}$$

$$p(\tilde{x}^R_{k^*}) = \int_{V_d} d\tilde{x}^R_{k\neq k^*} d\tilde{\mathbf{x}}^I p(\tilde{\mathbf{x}}) \tag{5.29}$$

5.4 Monte Carlo simulations

For high-dimensional vectors $\tilde{\mathbf{x}}$, which correspond to a field theory with very fine lattice resolution, computing the marginal distribution 5.29 is non-trivial. We have chosen to compute this distribution through histogramming values of $x^R_{k^*}$ generated through Metropolis Monte Carlo sampling of the full distribution 5.28. In order to run these simulations, a suitable value for the matrix $\mathbf{\Lambda}$ is needed. Since the approximated $S_{RPA}(\mathbf{k})$ is directly given by Leibler's calculations under the random phase approximation, $\mathbf{\Lambda}$ is easily obtained since it is the discretized inverse structure factor, $S(\mathbf{k})^{-1}$ for a melt of non-interacting ideal Gaussian chains. $S_{RPA}(\mathbf{k})$ from Leibler's calculations, evaluated along the x-axis (ie $k_z = k_y = 0$) is shown in 5.3.

After choosing a starting configuration for $\tilde{\mathbf{x}}$, new samples are proposed by choosing a random component $\tilde{x}_k$ and perturbing it by a random amount, $\tilde{x}_k \to \tilde{x}_k + \delta$. In order to ensure that the inverse FFT of $\tilde{\mathbf{x}}$ remains real, we also update wavevector $-k$ by the complex

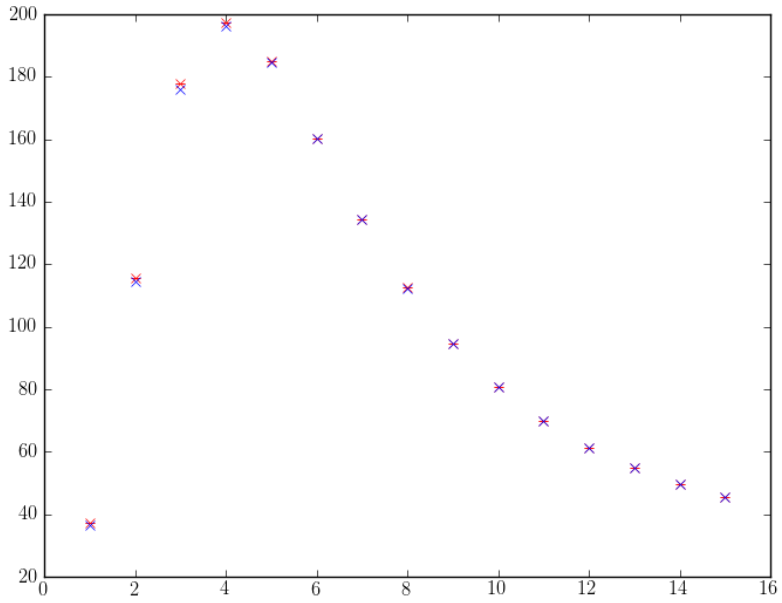


Figure 5.3: Structure factor $S(\mathbf{k})$, varied along the k_x axis, with $k_y = k_z = 0$, for a 64-chain system. Theoretical prediction based on direct calculation of $S(k)$ for a bead-spring system (blue) is compared against averages of $S(k)$ computed from Monte Carlo simulations of an explicit chain system (red).

conjugate of δ, δ^*. After proposing the update, it is also necessary to verify that $\tilde{\mathbf{x}}_{new}$ does not cause its real-space FFT pair, $\mathbf{x}_{new}$, to violate its constraints $-x_{max} < \mathbf{x}_{new} < x_{max}$. Whereas a discrete Fourier transform is $\mathcal{O}(N \log N)$, we can instead check this condition in $\mathcal{O}(N)$ time by keeping track of $\mathbf{x}$ throughout the simulation and noting that

$$\mathbf{x}_{new} = \mathbf{x}_{old} + \delta \mathbf{u} \tag{5.30}$$

$$\mathbf{u} = \left[e^{\frac{i\mathbf{r}_0 \cdot \mathbf{k}}{N}} \dots e^{\frac{i\mathbf{r}_{dN} \cdot \mathbf{k}}{N}} \right] \tag{5.31}$$

Starting from a random configuration, simulations were typically conducted for thousands of time steps to allow equilibration of the Markov MC chain before recording the value of $\tilde{x}_k^R$ after each Monte Carlo sweep (1 sweep = N MC steps). To increase the sampling of extremely unlikely values of $x_{k^*}^R$, quadratic potentials biasing $x_{k^*}^R$ were placed according to

$$U_{bias}(x_{k^*}^R) = \alpha (x_{k^*}^R - x_0)^2 \tag{5.32}$$

Tens to hundreds of simulations were run in parallel, with a different bias in each simulation, and the value of $x_{k^*}^R$ was recorded at every sweep for thousands of time steps in each sim-

ulation. A final estimate for $p(x^R_{k^*})$ was then computed by reweighting the histograms of each simulation and combining them using the weighted histogram analysis method (WHAM) [67].

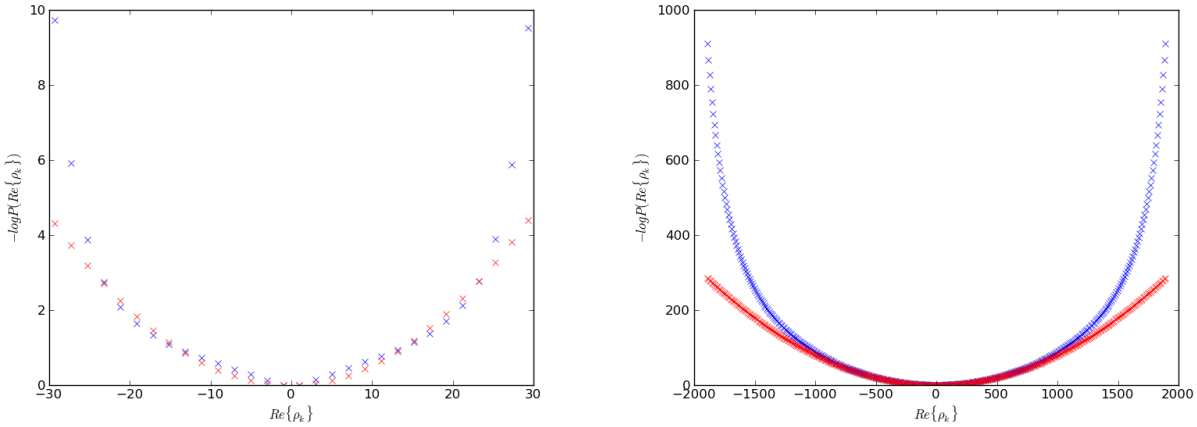
In addition to the lattice based simulations of 5.28, we also conducted Monte Carlo simulations with an explicit bead-spring representation of the polymers as a check on the validity of the field theory results. Although this approach is also not exactly analogous to the field theoretic model, since each polymer is explicitly coarse-grained into a bead-spring chain, it has been shown recently that the approach can still be exceedingly accurate [92]. We first simulated a system of independent, non-interacting Gaussian chains, the system used to calculate $S_{RPA}(k)$, and computed $p(\tilde{x}^R_{k^*})$ for this system. This is an interesting system, because it allows us to qualitatively assess the contribution of bead-spring chain structure to any potential non-Gaussian behavior of $p(\tilde{x}^R_{k^*})$ separately from excluded volume considerations. After this, it is possible to turn on volume excluding interactions in this model and compute $p(\tilde{x}^R_{k^*})$ again. We won't go into great detail regarding the simulation technique in order to maintain focus on the field-theoretic simulations. More extensive details on the model and methods for explicit chain calculations may be found elsewhere [29].

Results

As we show in Figure 5.4, for a single chain, $p(\tilde{x}^R_{k^*})$ is highly non-Gaussian. However, since $\tilde{x}^R_{k^*}$ is a Fourier transformed compositional variable, it is the sum of single particle density operators, and a central limit theorem of sorts ensures that this quantity quickly assumes Gaussian behavior as more chains are added to the system. This is demonstrated in Figure 5.4b, by increasing the number of chains to 64, and shown in further detail in 5.5. As a result, although a single Gaussian chain (confusingly) results in non-Gaussian compositional density fluctuations, for a sufficiently large number of chains, compositional density fluctuations are essentially Gaussian.

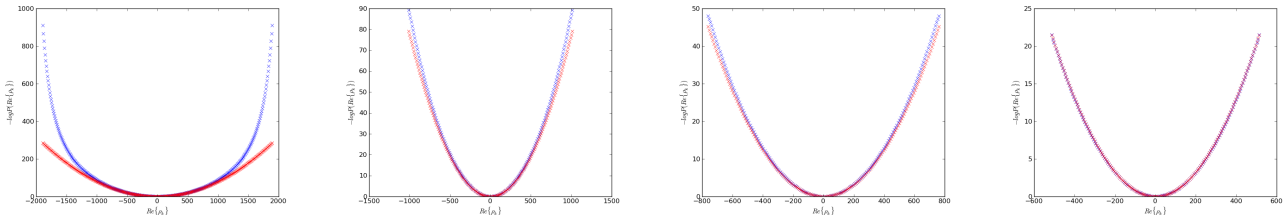
Next, the results for the volume excluding explicit chain calculations are given in Figure 5.6, which clearly demonstrates that the explicit chain model demonstrates behavior that is nearly Gaussian, even once volume excluding interactions are turned on. Given that intra-chain behavior eventually becomes Gaussian, it is not surprising that volume excluding interactions fail to disrupt this Gaussian behavior, since the density fluctuations of volume excluding simple liquids are known to be highly Gaussian, even down to very small length scales [28]. While these results are encouraging, the downside of the explicit chain technique is the heroic computational effort required to simulate and collect the data at realistic chain density parameters. For the data in Figure 5.6, a 640,000 particle system was needed, and many simulations with separate biasing potentials were run in parallel for over 20 days each. As a result, we were unable to push this system towards a regime where it strongly deviates from Gaussian behavior.

Finally, in Figure 5.7 we show the results of lattice simulations for the constrained Gaussian field theory, 5.28. In Figure 5.7, we compare the reweighted data (blue) with the log-probability for a Gaussian of the same mean and variance (green). The data are nearly identical until the constrained system has a nearly instantaneous divergence towards in-



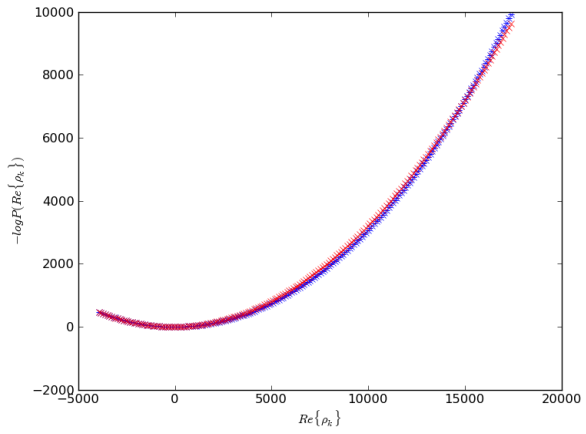
(a) (b)

Figure 5.4: (blue) Negative log-probability histogram obtained from WHAM for a single explicit 32 monomer bead-sping system. (red) The negative log-probability for a Gaussian distribution of the same mean and variance as computed from simulation data via WHAM. For a system of 1 chain (left) deviations from harmonic behavior may already be seen at relatively small kT values. When the system size is increased to 64 chains, $p(\tilde{x}_{k^*})$ remains Gaussian to much higher values of kT.



(a) (b) (c) (d)

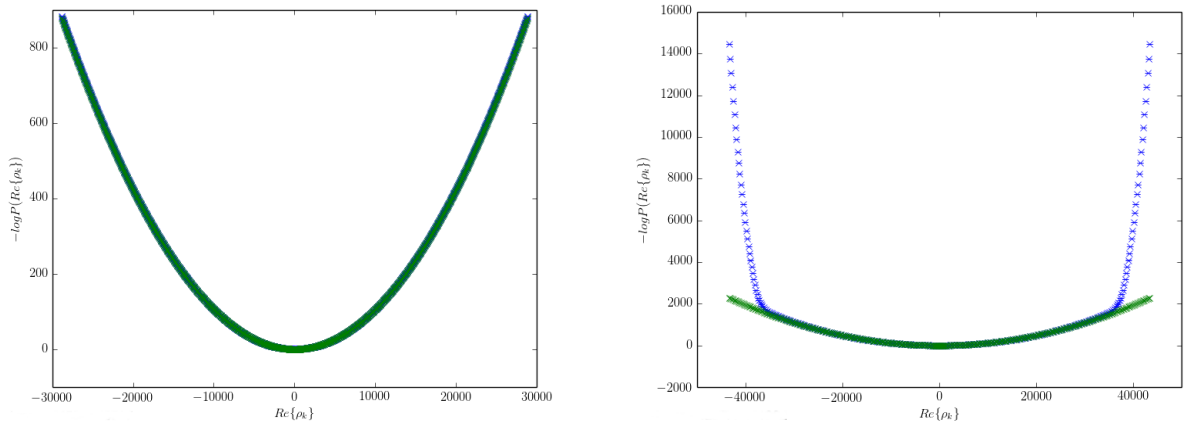
Figure 5.5: A more detailed view of the 64 chain explicit Monte Carlo simulation from 5.4b. Full data set is presented in (a). Panels (b), (c), and (d) present the same calculation applied to increasingly restricted data sets, showing that the data is well fit by a to a harmonic model (blue) that implies a Gaussian distribution.



(a)

Figure 5.6: (blue) Negative log-probability histogram obtained from WHAM for a 640,000 system of 32-bead polymers with volume excluding interactions turned on. Parameters are chosen so that the system would form 4 lamellar layers, were the AB interactions turned on. (red) The negative log-probability for a Gaussian distribution of the same mean and variance as computed from simulation data via WHAM.

finitely large free energy (which corresponds to zero probability). This indicates that the effect of imposing hard constraints on real-space density fluctuations is, to a first approximation, to impose a hard constraint on normal mode fluctuations. This is significant, because it indicates that the normal modes of the field theory retain their original structure, but are truncated, in the constrained model. As a result, the original $\phi^4 - \phi^2$ model can be replaced by a greatly simplified constrained Gaussian field theory which has the character and advantages of simpler Gaussian theories while retaining the stabilizing features of imparted by the ϕ^4 term. Since much is known about how to properly exclude volume and how to develop efficient numerical techniques for Gaussian theories, a promising direction for further extending these results will be to fully develop and apply what is known from liquid Gaussian theories to this system.



(a) (b)

Figure 5.7: (blue) Negative log-probability histogram obtained from WHAM for a 128x128 lattice with Hamiltonian 3.10. (green) The log-probability for a Gaussian distribution of the same mean and variance as computed from simulation data via WHAM. When no constraints are imposed (left) the distribution is Gaussian as expected. When constraints are imposed (right) that no lattice site can exceed ± 10, the data for wavevector k^* remains Gaussian far out into the wings.

Bibliography

[1] Michael P Allen and Dominic J Tildesley. *Computer simulation of liquids*. Oxford University Press, 1989.

[2] J.G. Amar, F.E. Sullivan, and R.D. Mountain. Monte Carlo study of growth in the two-dimensional spin-exchange kinetic Ising model. *Physical Review B*, 37(1):196–208, 1988.

[3] Joshua A. Anderson, Chris D. Lorenz, and A. Travesset. General purpose molecular dynamics simulations fully implemented on graphics processing units. *Journal of Computational Physics*, 227(10):5342–5359, May 2008.

[4] S. Babu, J.C. Gimel, and T. Nicolai. Diffusion limited cluster aggregation with irreversible flexible bonds. *Arxiv preprint arXiv:0801.4447*, 2008.

[5] GT Barkema and T MacFarland. Parallel simulation of the ising model. *Physical Review E*, 50(2):1623, 1994.

[6] Chris Bencher, Jeffrey Smith, Liyan Miao, Cathy Cai, Yongmei Chen, Joy Y. Cheng, Daniel P. Sanders, Melia Tjio, Hoa D. Truong, Steven Holmes, and William D. Hinsberg. Self-assembly patterning for sub-15nm half-pitch: a transition from lab to fab. In Daniel J. C. Herr, editor, *SPIE Advanced Lithography*, pages 79700F–79700F–9. International Society for Optics and Photonics, March 2011.

[7] L. Berthier. Revisiting the slow dynamics of a silica melt using Monte Carlo simulations. *Physical Review E*, 76(1):11507, 2007.

[8] L. Berthier and W. Kob. The Monte Carlo dynamics of a binary Lennard-Jones glass-forming mixture. *Journal of Physics, Condensed Matter*, 19(20):205130, 2007.

[9] E. Bianchi, P. Tartaglia, E. Zaccarelli, and F. Sciortino. Theoretical and numerical study of the phase diagram of patchy colloids: ordered and disordered patch arrangements. *Arxiv preprint arXiv:0802.2466*, 2008.

[10] K. Binder. Applications of Monte Carlo methods to statistical physics. *Reports on Progress in Physics*, 60(5):487–559, 1997.

[11] James J Binney, NJ Dowrick, AJ Fisher, and M Newman. *The theory of critical phenomena: an introduction to the renormalization group*. Oxford University Press, Inc., 1992.

[12] Ion Bita, Joel K W Yang, Yeon Sik Jung, Caroline A Ross, Edwin L Thomas, and Karl K Berggren. Graphoepitaxy of self-assembled block copolymers on two-dimensional periodic patterned templates. *Science (New York, N.Y.)*, 321(5891):939–43, August 2008.

[13] Henk W. J. Blote, Lev. N. Shchur, and Andrei L. Talapov. The Cluster Processor: New Results. *International Journal of Modern Physics C*, 10(06):1137–1148, September 1999.

[14] 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*, 404(6779):746–748, 2000.

[15] August W Bosse. Phase-field simulation of long-wavelength line edge roughness in diblock copolymer resists. *Macromolecular Theory and Simulations*, 19(7):399–406, 2010.

[16] John Philip Boyd. *Chebyshev and Fourier spectral methods*. Courier Dover Publications, 2001.

[17] AJ Bray. Theory of phase-ordering kinetics. *Advances in Physics*, 51(2):481–587, 2002.

[18] SA Brazovskii. Phase transition of an isotropic system to a nonuniform state. *Soviet Journal of Experimental and Theoretical Physics*, 41:85, 1975.

[19] M. G. Buonomenna, W. Yave, and G. Golemme. Some approaches for high performance polymer based membranes for gas separation: block copolymers, carbon molecular sieves and mixed matrix membranes. *RSC Advances*, 2(29):10745–10773, 2012.

[20] M. J. Cawkwell, E. J. Sanville, S. M. Mniszewski, and Anders M. N. Niklasson. Computing the Density Matrix in Electronic Structure Theory on Graphics Processing Units. *Journal of Chemical Theory and Computation*, 8(11):4094–4101, November 2012.

[21] Hector D Ceniceros and George O Mohler. A practical splitting method for stiff sdes with applications to problems with small noise. *Multiscale Modeling & Simulation*, 6(1):212–227, 2007.

[22] D. Chandler. *Introduction to modern statistical mechanics*. Oxford University Press New York, 1987.

[23] D. Chandler. Interfaces and the driving force of hydrophobic assembly. *Nature*, 437(7059):640–647, 2005.

[24] David Chandler. Gaussian field model of fluids with an application to polymeric fluids. *Physical Review E*, 48(4):2898, 1993.

[25] Weng C Chang and Peter White. *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*, volume 222. Oxford University Press on Demand, 2000.

[26] P. Charbonneau and DR Reichman. Systematic characterization of thermodynamic and dynamical phase behavior in systems with short-ranged attraction. *Physical Review E*, 75(1):11507, 2007.

[27] B. Chen and J.I. Siepmann. Improving the efficiency of the aggregation-volume-bias Monte Carlo algorithm. *J. Phys. Chem. B*, 105(45):11275–11282, 2001.

[28] Gavin E Crooks and David Chandler. Gaussian statistics of the hard-sphere fluid. *Physical Review E*, 56(4):4217, 1997.

[29] François A. Detcheverry, Darin Q. Pike, Paul F. Nealey, Marcus Müller, and Juan J. de Pablo. Simulations of theoretically informed coarse grain models of polymeric systems. *Faraday Discussions*, 144:111, October 2010.

[30] C.M. Dobson. Protein folding and misfolding. *Nature*, 426(6968):884–890, 2003.

[31] Oren M. Elrad and Michael F. Hagan. Mechanisms of size control and polymorphism in viral capsid assembly. *arXiv:0807.4344*, 2008.

[32] W A Fenton, Y Kashi, K Furtak, and A L Horwich. Residues in chaperonin GroEL required for polypeptide binding and release. *Nature*, 371(6498):614–9, October 1994.

[33] Y Fink, AM Urbas, MG Bawendi, JD Joannopoulos, and EL Thomas. Block copolymers as photonic bandgap materials. *Journal Of Lightwave Technology*, 17(11):1963–1969, NOV 1999. Workshop on Electromagnetic Crystal Structures, Laguna Beach, California, Jan 04-06, 1999.

[34] H. Fraenkel-Conrat and R.C. Williams. Reconstitution of Active Tobacco Mosaic Virus from Its Inactive Protein and Nucleic Acid Components. *Proceedings of the National Academy of Sciences of the United States of America*, 41(10):690–698, 1955.

[35] Glenn Fredrickson. *The Equilibrium Theory of Inhomogeneous Polymers (International Series of Monographs on Physics)*. Oxford University Press, USA, 2006.

[36] Glenn H Fredrickson and Eugene Helfand. Fluctuation effects in the theory of microphase separation in block copolymers. *The Journal of chemical physics*, 87:697, 1987.

[37] Glenn Harold Fredrickson. *The equilibrium theory of inhomogeneous polymers*. Clarendon Press Oxford, 2006.

[38] D. Frenkel. Speed-up of Monte Carlo simulations by sampling of rejected states. *Proceedings of the National Academy of Sciences*, 101(51):17571, 2004.

[39] D. Frenkel and B. Smit. *Understanding Molecular Simulation: From Algorithms to Applications*. Academic Press, 2002.

[40] Yijie Gao, John O. Thomas, Robert L. Chow, Gwo-Hwa Lee, and Nicholas J. Cowan. A cytoplasmic chaperonin that catalyzes β-actin folding. *Cell*, 69(6):1043–1050, June 1992.

[41] G. Ge and L.E. Brus. Fast Surface Diffusion of Large Disk-Shaped Nanocrystal Aggregates. *Nano Lett*, 1(4), 2001.

[42] Phillip L. Geissler. Dynamics in liquids, 2000. Copyright - Copyright UMI - Dissertations Publishing 2000; Last updated - 2010-08-07; First page - n/a; M3: Ph.D.

[43] MR Ghadiri, JR Granja, RA Milligan, DE Mcree, and N Khazanovich. Self-assembling organic nanotubes based on a cyclic peptide architecture . *Nature*, 372(6507):709, Dec 15 1994.

[44] Valeriy V Ginzburg, Feng Qiu, Marco Paniconi, Gongwen Peng, David Jasnow, and Anna C Balazs. Simulation of hard particles in a phase-separating binary mixture. *arXiv preprint cond-mat/9905284*, 1999.

[45] AW Götz, T Wölfle, and RC Walker. Quantum chemistry on graphics processing units. *Reports in Computational Chemistry*, 6(10):21–35, 2010.

[46] YH Ha and EL Thomas. Deformation behavior of a roll-cast layered-silicate/lamellar triblock copolymer nanocomposite. *Macromolecules*, 35(11):4419–4428, May 21 2002.

[47] M.F. Hagan and D. Chandler. Dynamic Pathways for Viral Capsid Assembly. *Biophysical Journal*, 91(1):42, 2006.

[48] Samuli Hakala, Ville Havu, Jussi Enkovaara, and Risto Nieminen. Multiple Graphics Processing Units (GPUs). pages 63–76, 2013.

[49] W G Hixon and D G Searcy. Cytoskeleton in the archaebacterium Thermoplasma acidophilum? Viscosity increase in soluble extracts. *Bio Systems*, 29(2-3):151–60, January 1993.

[50] Pierre C Hohenberg and Bertrand I Halperin. Theory of dynamic critical phenomena. *Reviews of Modern Physics*, 49(3):435, 1977.

[51] M.A. Horsch, Z. Zhang, and S.C. Glotzer. Self-Assembly of Polymer-Tethered Nanorods. *Physical Review Letters*, 95(5):56105, 2005.

[52] Rami Hourani, Chen Zhang, Rob van der Weegen, Luis Ruiz, Changyi Li, Sinan Keten, Brett A. Helms, and Ting Xu. Processable Cyclic Peptide Nanotubes with Tunable Interiors. *Journal of the American Chemical Society*, 133(39):15296–15299, Oct 5 2011.

[53] K. Huang. *Statistical Mechanics, J.* Wiley, New York, 1987.

[54] J.C. Huie. Guided molecular self-assembly: a review of recent efforts. *Smart Materials and Structures*, 12(2):264–271, 2003.

[55] BAH Huisman, PG Bolhuis, and A. Fasolino. Phase transition to bundles of flexible supramolecular polymers. *Arxiv preprint arXiv:0711.4704*, 2007.

[56] Farzana Hussain, Mehdi Hojjati, Masami Okamoto, and Russell E. Gorga. Review article: Polymer-matrix nanocomposites, processing, manufacturing, and application: An overview. *Journal Of Composite Materials*, 40(17):1511–1575, SEP 2006.

[57] C. Hyeon, G.H. Lorimer, and D. Thirumalai. Dynamics of allosteric transitions in GroEL. *Proceedings of the National Academy of Sciences*, 103(50):18939, 2006.

[58] R.L. Jack, M.F. Hagan, and D. Chandler. Fluctuation-dissipation ratios in the dynamics of self-assembly. *Physical Review E*, 76(2):21119, 2007.

[59] H.K. Kagawa, T. Yaoi, L. Brocchieri, R.A. McMillan, T. Alton, and J.D. Trent. The composition, structure and stability of a group II chaperonin are temperature regulated in a hyperthermophilic archaeon. *Molecular Microbiology*, 48(1):143–156, 2003.

[60] Joseph Kao, Peter Bai, J. Matthew Lucas, A. Paul Alivisatos, and Ting Xu. Size-Dependent Assemblies of Nanoparticle Mixtures in Thin Films. *Journal of the American Chemical Society*, 135(5):1680–1683, Feb 6 2013.

[61] Joseph Kao, Kari Thorkelsson, Peter Bai, Benjamin J. Rancatore, and Ting Xu. Toward functional nanocomposites: taking the best of nanoparticles, polymers, and small molecules. *Chemical Society Reviews*, 42(7):2654–2678, 2013.

[62] N Khazanovich, JR Granja, DE Mcree, RA Milligan, and MR Ghadiri. Nanoscale tubular ensembles with specified internal diameters - Design of a self-assembled nanotube with a 13-Angstrom Pore. *Journal of the American Chemical Society*, 116(13):6011–6012, Jun 29 1994.

[63] K. Kikuchi, M. Yoshida, T. Maekawa, and H. Watanabe. Metropolis Monte Carlo method as a numerical technique to solve the FokkerPlanck equation. *Chemical Physics Letters*, 185(3-4):335–338, 1991.

[64] DB Kirk and WH Wen-mei. *Programming Massively Parallel Processors A Hands-on Approach*. 2010.

[65] John L Klepeis, Kresten Lindorff-Larsen, Ron O Dror, and David E Shaw. Long-timescale molecular dynamics simulations of protein structure and function. *Current opinion in structural biology*, 19(2):120–7, April 2009.

[66] W. Krauth. *Statistical mechanics: algorithms and computations*. Oxford University Press, Oxford, 2006.

[67] Shankar Kumar, John M Rosenberg, Djamal Bouzida, Robert H Swendsen, and Peter A Kollman. The weighted histogram analysis method for free-energy calculations on biomolecules. i. the method. *Journal of Computational Chemistry*, 13(8):1011–1021, 1992.

[68] David P. Landau and Kurt Binder. *A Guide to Monte Carlo Simulations in Statistical Physics*. Cambridge University Press, 2009.

[69] Anthony Lee, Christopher Yau, Michael B. Giles, Arnaud Doucet, and Christopher C. Holmes. On the Utility of Graphics Cards to Perform Massively Parallel Simulation of Advanced Monte Carlo Methods. *Journal of Computational and Graphical Statistics*, 19(4):769–789, January 2010.

[70] Ludwik Leibler. Theory of microphase separation in block copolymers. *Macromolecules*, 13(6):1602–1617, 1980.

[71] Erin M Lennon, George O Mohler, Hector D Ceniceros, Carlos J García-Cervera, and Glenn H Fredrickson. Numerical solutions of the complex langevin equations in polymer field theory. *Multiscale Modeling & Simulation*, 6(4):1347–1370, 2008.

[72] Y. Li, CD Paavola, H. Kagawa, SL Chan, and JD Trent. Mutant chaperonin proteins: new tools for nanotechnology. *Nanotechnology*, 18(455101):455101, 2007.

[73] N.A. Licata and A.V. Tkachenko. Colloids with key-lock interactions: Nonexponential relaxation, aging, and anomalous diffusion. *Physical Review E*, 76(4):41405, 2007.

[74] J. Liu and E. Luijten. Rejection-Free Geometric Cluster Algorithm for Complex Fluids. *Physical Review Letters*, 92(3):35504, 2004.

[75] Weiguo Liu, Bertil Schmidt, Gerrit Voss, and Wolfgang Müller-Wittig. Accelerating molecular dynamics simulations using Graphics Processing Units with CUDA. *Computer Physics Communications*, 179(9):634–641, November 2008.

[76] Weiguo Liu, Bertil Schmidt, Gerrit Voss, and M Wolfgang. Molecular Dynamics Simulations on Commodity GPUs with CUDA. pages 185–196, 2007.

[77] S. A. Maier, M. L. Brongersma, P. G. Kik, S. Meltzer, A. A. G. Requicha, and H. A. Atwater. Plasmonics-A Route to Nanoscale Optical Devices. *Advanced Materials*, 13(19):1501–1505, September 2001.

[78] Stefan Maier, Mark Brongersma, Pieter Kik, and Harry Atwater. Observation of near-field coupling in metal nanoparticle chains using far-field polarization spectroscopy. *Physical Review B*, 65(19):193408, May 2002.

[79] CH Mak. Stochastic potential switching algorithm for Monte Carlo simulations of complex systems. *The Journal of Chemical Physics*, 122:214110, 2005.

[80] M. W. Matsen and F. S. Bates. Unifying weak- and strong-segregation block copolymer theories. *Macromolecules*, 29(4):1091–1098, 1996.

[81] R.A. McMillan, C.D. Paavola, J. Howard, S.L. Chan, N.J. Zaluzec, and J.D. Trent. Ordered nanoparticle arrays formed on engineered chaperonin protein templates. *Nature Materials*, 1(4):247–252, 2002.

[82] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics*, 21(6):1087, June 1953.

[83] Paul Norman, Paolo Valentini, and Thomas Schwartzentruber. GPU-accelerated Classical Trajectory Calculation Direct Simulation Monte Carlo applied to shock waves. *Journal of Computational Physics*, 247:153–167, August 2013.

[84] Nvidia. Cuda C programming guide. (October), 2012.

[85] D. Nykypanchuk, M.M. Maye, D. van der Lelie, and O. Gang. Main navigation. *Nature*, 451:549–552, 2008.

[86] T. Opplestrup, V.V. Bulatov, G.H. Gilmer, M.H. Kalos, and B. Sadigh. First-Passage Monte Carlo Algorithm: Diffusion without All the Hops. *Physical Review Letters*, 97(23):230602, 2006.

[87] G. Orkoulas and A.Z. Panagiotopoulos. Phase behavior of the restricted primitive model and square-well fluids from Monte Carlo simulations in the grand canonical ensemble. *The Journal of Chemical Physics*, 110:1581, 1999.

[88] Thomas E. Ouldridge, Iain G. Johnston, Ard A. Louis, and Jonathan P. K. Doye. Minimal models of dna and the self-assembly of dna nanostructures. *arXiv:0807.3280*, 2008.

[89] JD Owens and David Luebke. A Survey of GeneralPurpose Computation on Graphics Hardware. *Computer graphics ...*, (August):21–51, 2007.

[90] CD Paavola, SL Chan, Y. Li, KM Mazzarella, RA McMillan, and JD Trent. A versatile platform for nanotechnology based on circular permutation of a chaperonin protein. *Nanotechnology*, 17(5):1171–1176, 2006.

[91] SY Park, AK Lytton-Jean, B. Lee, S. Weigand, GC Schatz, and CA Mirkin. DNA-programmable nanoparticle crystallization. *Nature*, 451(7178):553–6, 2008.

[92] Darin Q Pike, François A Detcheverry, Marcus Müller, and Juan J de Pablo. Theoretically informed coarse grain simulations of polymeric systems. *The Journal of chemical physics*, 131(8):084903, August 2009.

[93] Guillem Pratx and Lei Xing. GPU computing in medical physics: A review. *Medical Physics*, 38(5):2685, 2011.

[94] Tobias Preis, Peter Virnau, Wolfgang Paul, and Johannes J Schneider. Gpu accelerated monte carlo simulation of the 2d and 3d ising model. *Journal of Computational Physics*, 228(12):4468–4477, 2009.

[95] Andrew D. Presley, Joseph J. Chang, and Ting Xu. Directed co-assembly of heme proteins with amphiphilic block copolymers toward functional biomolecular materials. *Soft Matter*, 7(1):172, December 2011.

[96] V.F. Puntes, K.M. Krishnan, and A.P. Alivisatos. Colloidal Nanocrystal Shape and Size Control: The Case of Cobalt. *Science*, 291(5511):2115–2117, 2001.

[97] S Puri and Y Oono. Study of phase-separation dynamics by use of cell dynamical systems. ii. two-dimensional demonstrations. *Physical Review A*, 38(3):1542–1565, 1988.

[98] E Quaite-Randall and A Joachimiak. Purification of archaeal chaperonin from Sulfolobus shibatae. *Methods in molecular biology (Clifton, N.J.)*, 140:1–14, January 2000.

[99] E. Quaite-Randall, J.D. Trent, R. Josephs, and A. Joachimiak. Conformational Cycle of the Archaeosome, a TCP1-like Chaperonin from Sulfolobus shibatae. *Journal of Biological Chemistry*, 270(48):28818, 1995.

[100] E. Rabani, D.R. Reichman, P.L. Geissler, L.E. Brus, et al. Drying-mediated self-assembly of nanoparticles. *Nature*, 426(6964):271–274, 2003.

[101] E. Rabani, D.R. Reichman, P.L. Geissler, L.E. Brus, et al. Drying-mediated self-assembly of nanoparticles. *Nature*, 426(6964):271–274, 2003.

[102] Benjamin J. Rancatore, Clayton E. Mauldin, Shih-Huang Tung, Cheng Wang, Alexander Hexemer, Joseph Strzalka, Jean M. J. Frechet, and Ting Xu. Nanostructured Organic Semiconductors via Directed Supramolecular Assembly. *ACS Nano*, 4(5):2721–2729, MAY 2010.

[103] DC Rapaport. *The Art of Molecular Dynamics Simulation*. Cambridge University Press, 2004.

[104] DC Rapaport. The role of reversibility in viral capsid growth: A new paradigm for self-assembly. *Arxiv preprint arXiv:0803.0115*, 2008.

[105] M.C. Rechtsman, F.H. Stillinger, and S. Torquato. Optimized Interactions for Targeted Self-Assembly: Application to a Honeycomb Lattice. *Physical Review Letters*, 95(22):228301, 2005.

[106] JX Ren, AS Silva, and R Krishnamoorti. Linear viscoelasticity of disordered polystyrene-polyisoprene block copolymer based layered-silicate nanocomposites. *Macromolecules*, 33(10):3739–3746, MAY 16 2000.

[107] MBO Riekerink, MB Claase, GHM Engbers, DW Grijpma, and J Feijen. Gas plasma etching of PEO/PBT segmented block copolymer films. *Journal of Biomedical Materials Research Part A*, 65A(4):417–428, JUN 15 2003.

[108] M. Rottereau, J.C. Gimel, T. Nicolai, and D. Durand. Influence of the Brownian step size in off-lattice Monte Carlo simulations of irreversible particle aggregation. *The European Physical Journal E-Soft Matter*, 18(1):15–19, 2005.

[109] Ricardo Ruiz, Huiman Kang, François A Detcheverry, Elizabeth Dobisz, Dan S Kercher, Thomas R Albrecht, Juan J de Pablo, and Paul F Nealey. Density multiplication and improved lithography by directed block copolymer assembly. *Science (New York, N.Y.)*, 321(5891):936–9, August 2008.

[110] Toshihiko Sato, Haroon Ahmed, David Brown, and Brian F. G. Johnson. Single electron transistor using a molecularly linked gold colloidal particle chain. *Journal of Applied Physics*, 82(2):696, July 1997.

[111] G. Schoehn, E. Quaite-Randall, J.L. Jiménez, A. Joachimiak, and H.R. Saibil. Three conformations of an archaeal chaperonin, TF55 from Sulfolobus shibatae. *Journal of Molecular Biology*, 296(3):813–819, 2000.

[112] F. Sciortino, E. Bianchi, J.F. Douglas, and P. Tartaglia. Self-assembly of patchy particles into polymer chains: A parameter-free comparison between Wertheim theory and Monte Carlo simulation. *The Journal of Chemical Physics*, 126:194903, 2007.

[113] D G Searcy and W G Hixon. Cytoskeletal origins in sulfur-metabolizing archaebacteria. *Bio Systems*, 25(1-2):1–11, January 1991.

[114] Jessica Y. Shu, Brian Panganiban, and Ting Xu. Peptide-Polymer Conjugates: From Fundamental Science to Application. In Johnson, MA and Martinez, TJ, editor, *Annual Review of Physical Chemistry, Vol. 64*, volume 64 of *Annual Review of Physical*

Chemistry, pages 631–657. Annual Reviews, 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0897 USA, 2013.

[115] Scott W Sides, Bumjoon J Kim, Edward J Kramer, and Glenn H Fredrickson. Hybrid particle-field simulations of polymer nanocomposites. *Physical Review Letters*, 96(25):250601, 2006.

[116] Mark Somervell, Roel Gronheid, Joshua Hooge, Kathleen Nafus, Paulina Rincon Delgadillo, Chris Thode, Todd Younkin, Koichi Matsunaga, Ben Rathsack, Steven Scheer, and Paul Nealey. Comparison of Directed Self-Assembly Integrations. *Proc. SPIE*, 8325:83250G–83250G–14, March 2012.

[117] PP Soo, BY Huang, YI Jang, YM Chiang, DR Sadoway, and AM Mayes. Rubbery block copolymer electrolytes for solid-state rechargeable lithium batteries. *Journal of the Electrochemical Society*, 146(1):32–37, JAN 1999.

[118] H. Sternlicht. The t-Complex Polypeptide 1 Complex is a Chaperonin for Tubulin and Actin in vivo. *Proceedings of the National Academy of Sciences*, 90(20):9422–9426, October 1993.

[119] TR Strick, J.F. Allemand, D. Bensimon, and V. Croquette. Stress-Induced Structural Transitions In DNA And Proteins. *Annual Reviews in Biophysics and Biomolecular Structure*, 29(1):523–543, 2000.

[120] R.H. Swendsen and J.S. Wang. Nonuniversal critical dynamics in Monte Carlo simulations. *Physical Review Letters*, 58(2):86–88, 1987.

[121] G. Tiana, L. Sutto, and RA Broglia. Use of the Metropolis algorithm to simulate the dynamics of protein chains. *Physica A: Statistical Mechanics and its Applications*, 380:241–249, 2007.

[122] VI Tokar and H. Dreyssé. Accelerated kinetic Monte Carlo algorithm for diffusion-limited kinetics. *Physical Review E*, 77(6):66705, 2008.

[123] Stanimire Tomov, Michael McGuigan, Robert Bennett, Gordon Smith, and John Spiletic. Benchmarking and implementation of probability-based simulations on programmable graphics cards. *Computers & Graphics*, 29(1):71–80, February 2005.

[124] J.D. Trent, H.K. Kagawa, T. Yaoi, E. Olle, and N.J. Zaluzec. Chaperonin filaments: The archaeal cytoskeleton? *Proceedings of the National Academy of Sciences*, 94(10):5383, 1997.

[125] J.D. Trent, E. Nimmesgern, J.S. Wall, F.U. Hartl, and A.L. Horwich. A molecular chaperone from a thermophilic archaebacterium is related to the eukaryotic protein t-complex polypeptide-1. *Nature*, 354:490–493, 1991.

[126] F. Ulrich Hartl and Jörg Martin. Molecular chaperones in cellular protein folding. *Current Opinion in Structural Biology*, 5(1):92–102, February 1995.

[127] M Weigel. GPU accelerated Monte Carlo simulations of lattice spin models. *Physics Procedia*, 15:92–96, 2011.

[128] Martin Weigel. Simulating spin models on GPU. 2011.

[129] MS Wertheim. Fluids with highly directional attractive forces. I. Statistical thermodynamics. *Journal of Statistical Physics*, 35(1):19–34, 1984.

[130] S. Whitelam and P.L. Geissler. Avoiding unphysical kinetic traps in Monte Carlo simulations of strongly attractive particles. *The Journal of Chemical Physics*, 127:154101, 2007.

[131] G.M. Whitesides and M. Boncheva. Beyond molecules: Self-assembly of mesoscopic and macroscopic components. *PNAS*, 99(8):4769–4774, 2002.

[132] G.M. Whitesides and B. Grzybowski. Self-Assembly at All Scales. *Science*, 295(5564):2418–2421, 2002.

[133] A.W. Wilber, J.P.K. Doye, A.A. Louis, E.G. Noya, M.A. Miller, and P. Wong. Reversible self-assembly of patchy particles into monodisperse icosahedral clusters. *The Journal of Chemical Physics*, 127:085106, 2007.

[134] U. Wolff. Collective Monte Carlo Updating for Spin Systems. *Physical Review Letters*, 62(4):361–364, 1989.

[135] D. Wu, D. Chandler, and B. Smit. Electrostatic analogy for surfactant assemblies. *The Journal of Physical Chemistry*, 96(10):4077–4083, 1992.

[136] Y. Xia, B. Gates, Y. Yin, and Y. Lu. Monodispersed Colloidal Spheres: Old Materials with New Applications. *Advanced Materials*, 12(10):693–713, May 2000.

[137] Ting Xu, Nana Zhao, Feng Ren, Rami Hourani, Ming Tsang Lee, Jessica Y. Shu, Samuel Mao, and Brett A. Helms. Subnanometer Porous Thin Films by the Co-assembly of Nanotube Subunits and Block Copolymers. *ACS Nano*, 5(2):1376–1384, FEB 2011.

[138] Taras Yavors and Martin Weigel. Optimized GPU simulation of continuous-spin glass models. pages 1–15.

[139] Haitao Yu, Pengcheng Xu, Xiaoyuan Xia, Dong-Weon Lee, and Xinxin Li. Micro-/Nanocombined Gas Sensors With Functionalized Mesoporous Thin Film Self-Assembled in Batches Onto Resonant Cantilevers. *IEEE Transactions on Industrial Electronics*, 59(12):4881–4887, December 2012.

[140] S. Zhang. Fabrication of novel biomaterials through molecular self-assembly. *Nature Biotechnology*, 21(10):1171–1178, 2003.

[141] Z. Zhang, A.S. Keys, T. Chen, and S.C. Glotzer. Self-assembly of patchy particles into diamond structures through molecular mimicry. *Langmuir*, 21(25):11547–11551, 2005.

[142] ZL 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*, 3(10):1341–1346, 2003.

[143] Yue Zhao, Kari Thorkelsson, Alexander J. Mastroianni, Thomas Schilling, Joseph M. Luther, Benjamin J. Rancatore, Kazuyuki Matsunaga, Hiroshi Jinnai, Yue Wu, Daniel Poulsen, Jean M. J. Frechet, A. Paul Alivisatos, and Ting Xu. Small-molecule-directed nanoparticle assembly towards stimuli-responsive nanocomposites. *Nature Materials*, 8(12):979–985, DEC 2009.

[144] V.V. Zhirnov and D.J.C. Herr. New frontiers: self-assembly and nanoelectronics. *Computer*, 34(5):34–43, May 2001.

[145] M. Zinke-Allmang, LC Feldman, and MH Grabow. Clustering on Surfaces. *Surface Science Reports(The Netherlands)*, 16(8):377–463, 1992.