Title

Multi-scale models of amphiphilic assemblies

Author

Liu, Xiaoxuan

Publication Date

2021

Multi-scale models of amphiphilic assemblies

by

Xiaoxuan Liu

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

in

Chemistry

in the

Graduate Division
of the
University of California, Berkeley

Committee in charge:

Professor Phillip Geissler, Chair
Professor Eran Rabani
Professor Daniel Fletcher

Fall 2021

Multi-scale models of amphiphilic assemblies

Abstract

Multi-scale models of amphiphilic assemblies

by

Xiaoxuan Liu
Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor Phillip Geissler, Chair

Nature utilizes amphiphilic assemblies to compartmentalize different chemical environments, modulate interfacial properties and reaction rates, and create complex materials with mesoscopic order. Many models explain the phase behavior of amphiphiles, but our understanding of the non-equilibrium behaviors of amphiphilic assemblies is limited. There exists a vast and rich space of structural and functional variations unexplored by equilibrium methods and models. A recent set of amphiphilic self-assembly experiments have suggested that non-equilibrium pathways of amphiphilic assemblies are complex and poorly understood[1]. In the experiments, assembly of 2-oxooctanoic acid was photoinitiated and eventually yielded monodisperse aggregates of long-term stability. In this dissertation, I explore various states of amphiphilic assemblies and the non-equilibrium processes that interconnect these states in the context of the photoinitiated assembly model system. In Chapter 2, I used molecular dynamics simulations to study structural variations and mechanisms of formation of the early nuclei from aqueous solution. Importance sampling and classical nucleation theory were employed to estimate the rate of nucleation. I found that the kinetics of coaggregation, where a mixture of more than one amphiphilic species, could differ significantly from that of single-species aggregation. Specifically, the participation of a second species could open up new pathways of growth and introduce microscopic phase separation of the two species within the aggregates and as a result modulate the growth rates. In Chapter 3, I computed the solutions to master-equation chemical kinetic models of amphiphilic aggregation pathways. I found that a large critical nucleus size is an important factor that contributes to the production of a narrow aggregate size distribution, a highly desirable characteristic in the preparation of self-assembled nanoparticles. I incorporate the elementary steps of nucleation, growth, and a source of precursor molecules to observe the effects of competing elementary rates on the aggregate size distribution. In Chapter 4, I furthered the development of the charge-frustrated Ising model to represent amphiphilic species with non-zero spontaneous curvatures. I also built in the correct interfacial roughness by using two lattice spacings in the same lattice model.

For my mother.

Contents

List of Figures

List of Tables

Acknowledgments

Many people helped me towards the completion of this dissertation. My advisor Phill Geissler was there for me whenever I needed guidance. He had the most elegant perspective of any complex concept, and he kept my thoughts rigorous when I so often get carried away in my excitement about something new. I also thank Julia Rogers, Layne Frechette, Andreana Rosnik, Rian Kormos, Pratima Satish, Georg Menzl, Jaffar Hasnain, and other Geissler Group members who crossed paths with me over the years. They provided me with the much-needed companionship on my journey into a scientific field that can appear all too overwhelming in the beginning. Towards the end of my time at Berkeley, I also enjoyed many enlightening conversations with Katie Klymko and Ahmad Omar, who generously agreed to rescue my sanity over Zoom during the pandemic.

In middle school, my Physics teacher told me that I should not take the extra-curricular physics lessons because "girls don't need to worry about it." But with extraordinary luck, I had Andrew McKenzie as my Physics teacher. Without the help from help and many others at Chigwell, including Tracy Martin, Peter Clayton, Shakil Chaudhary, and Andy Long, I would never have been able to even dream of attending Cambridge. During college, I studied a wide range of subjects under patient guidance from Paul Wood, Tuomas Knowles, Thomas Michaels, Andela Saric, and Daan Frenkel. It was their example and encouragement that led me to scientific research.

Outside work, my long-time friends Helen Ailun Huang, Sheng Gao, and Yihan Xu listened to my troubles and helped me despite being scattered across the country and around the world. My husband Nathan Odendahl tolerated my temper throughout the four years and fed me delicious meals during the writing of this dissertation. As the first person in my family to go to college I thank my grandparents and great grandparents for cheering for me along this path. I thank my aunt and uncle's family, especially my cousin Yi Nuo for admiring me as I strive to become the big sister that she looks up to. Lastly, I thank my mother for her unwavering support. The challenges during graduate school cannot compare in the slightest to those dark times that we endured together. I'm excited to show you the bright future, Mom.

Chapter 1

Introduction to amphiphilic assemblies

1.1 Amphiphiles and their phase behaviors

Liquids are often classified as hydrophilic (water-loving) and hydrophobic (water-hating). On their own, hydrophobic and hydrophilic molecules have mismatched interactions that make their mixing unfavorable; mixtures of hydrophilic and hydrophobic molecules are therefore in general capable of segregating into distinct phases.

Amphiphilic molecules possess one or more hydrophobic (groups of) atoms and one or more hydrophilic (groups of) atoms. In amphiphilic molecules, the two types of groups are connected by a covalent bond that prevents the two groups from phase separation. Consequently, amphiphilic molecules tend to sit at the interface between macroscopic hydrophobic and hydrophilic phases, such as the air-water or the oil-water interfaces. In addition, supramolecular assemblies in the liquid phase often involve the creation of such interfaces from solution. Due to their ability to modulate liquid interfacial properties and liquid-phase assemblies, amphiphiles are widely seen in nature.

A common household example is milk, which contains casein micelles and milk fat globules as the major insoluble components[2]. Casein micelles are the source of calcium in milk and are not micelles in the conventional sense but large, 50 to 250 nm aggregates coated with peptide chains that prevent coagulation[3]. In the process of making cheese and yogurt, casein is denatured to expose the hydrophobic regions of the polypeptide, yielding a cross-linked gel. Despite their technological relevance, the interior structure of casein micelles, or even their existence, remains a topic of debate.

The milk fat globules on the other hand exemplify a method of fat transport commonly seen in biological systems[4]. Fat molecules such as triacylglycerides are important for energy storage, but their hydrophobic nature would force them to macroscopically phase separate at high concentrations from the aqueous cellular environment. Such separation does not allow the organism to control the equilibrium between increasing and depleting energy storage. Evolution's solution to this problem is to coat the hydrophobic molecules with amphiphilic molecules to form a membrane[5, 6], which stabilizes the fat-water interfaces and prevents macroscale separation[7]. The membrane can be further tagged by amphiphilic proteins for signaling and transporting to specific locations in the cellular environment[8]. In milk, the fat globules secreted into the suspension are coated with three monolayers of phospholipids, whereas in adipose tissues intracellular energy storage utilizes lipid droplets that have only a monolayer of phospholipids.

Amphiphiles can form more than just monolayers at the interfaces of hydrophilic and hydrophobic environments. At high concentrations, the interfacial surface areas are no longer large enough to accommodate all the amphiphilic molecules, and in response to that amphiphilic molecules create increasingly more interfacial area by assembling into clusters and networks with microscopic phase separation[9]. For instance, in water, a highly concentrated solution of amphiphiles will form microscopic micellar aggregates, which can further assemble hierarchically into a lyotropic liquid crystal of micelles, and further fuse into two interconnected networks of water and hydrophobic groups.

The lyotropic liquid crystalline phase has received attention for its potential as a vehicle for drug delivery[10]. Hydrophobic drug molecules can be incorporated into the hydrophobic regions of the amphiphilic assemblies[11]. The existence of a layer of amphiphiles can also protect the carried drug from degradation and aid the entrance of the drug molecules through the endocytosis pathways[12]. A prominent recent example is the mRNA-lipid nanoparticle complex, which has negatively charged nucleic acid conjugated with positively charged lipids[13, 14, 15, 16].

1.2 Models of amphiphiles

Despite their common appearance, our understanding of amphiphilic assemblies is limited. Most current models are only applicable for rationalizing an observed assembly. The ultimate goal of studying self-assembly is to be able to predict and thus design assembled states[17], as well as predict the dynamics with which the system evolves towards the assembled states, but it is often the case that we have neither the resolution nor the predictive capability for the microscopic structure of these assemblies.

1.2.1 Membrane as elastic sheets

In 1925, Gorter and Grendel deposited the constituent lipids of the red blood cell plasma membranes in a monolayer and measured its surface area to prove that the membranes are bilayers[18]. It wasn't until half a century later that Helfrich proposed a continuum theory of membrane elasticity[19] to argue that curvature elasticity controls the shape of assembled lipid bilayers. Seifert and Lipowski then used the Helfrich model to draw phase diagrams for shape transformations of vesicles[20] and explain the mechanism of vesicle adhesion[21]. Safran then resolved the paradox of vesicle thermodynamic stability by examining the Helfrich model of a two-component lipid mixture with opposite curvatures[22]. More recently, Levine et al. showed that the CHARMM36 molecular dynamics force field can reproduce experimentally measured mechanical properties of phospholipid bilayers[23].

What these models all share in common is the treatment of the plasma membrane as elastic sheets with mechanical properties associated with the material. They have been successful in modeling micron-scale behavior such as the shape deformations of cells. They cannot however be used to model transformations that involve the disassembly or dissolution of constituent lipids.

1.2.2 Molecular geometry

Due to their broad definition, amphiphiles are a class of compounds that exhibit a large range of chemical diversity due to the possible hydrophobic and hydrophilic groups that can be combined. In principle, it should be possible to connect the continuum models that treat membranes as elastic sheets to a molecular perspective and predict the membrane material properties from molecular structure.

To do this, Israelachvili proposed a packing parameter that summarizes the molecular structure of an amphiphilic molecule by its shape[24]. The hydrophilic head group can be assigned a size based on the functional groups it contains, its charge, and coordination number with solvents. Similarly, the hydrophobic tail group can be assigned a size primarily based on the number of tails, the number of carbons in the chain, and any unsaturated carbon-carbon bonds that cause the tails to bend. Overall, the amphiphile is therefore considered either a cone (small head, large tail), a cylinder (comparably sized head and tail), or an inverted cone (large head, small tail).

This geometric view is a qualitative proxy to the factors that affect non-translational chemical potential of the monomer. It is then possible to model the chemical equilibrium of various states of assemblies at different concentrations. The conical amphiphiles will preferentially assemble into micelles (also known as oil-in-water micelles) or other structures with the same concave curvature. The inverted conical amphiphiles will preferentially assemble into inverted micelles (also known as water-in-oil micelles). Finally, the cylindrical micelles will preferentially assemble into bicelles (also known as bilayer discs).

1.2.3 Field-theoretic models

The Ising model is a minimal model that encompasses the liquid phase behaviors from which the hydrophobic effect[25] arises. Initially intended for modeling the phase transitions of ferromagnets, the $+1$ and -1 magnetic spins can be instead seen as the occupation or densities of a hydrophobic substance and a hydrophilic substance. Since the assembly of amphiphilic species at its core is also due to the hydrophobic effect, a modified version of the Ising model should apply to the study of amphiphilic assemblies.

Indeed, Wu et al. showed that the density correlation functions of a solution of amphiphiles can be expanded in Fourier space and then truncated to second-order of wavevector k to preserve only the small-k (or large scale in real space) behaviors of assembly[26]. It turned out that the truncated density correlation, and the corresponding effective Hamiltonian found through Gaussian field theory, are identical to that of a solution of charges. More specifically, the density of amphiphilic molecules can be represented in the Ising model by placing opposite effective charges on the up and down spins. That is, the up spins will gain a positive charge, while the down spins will gain a negative charge. As a result, the opposite spins have an unfavorable interaction but the electrostatic interaction that arises between the opposing spins is attractive and favorable.

Under equilibrium conditions, the electrostatic analogy and the charge-frustrated Ising model can ultimately bridge microscopic molecular properties (in the form of intramolecular density correlation functions) with macroscopic treatments of membranes as elastic sheets[27]. The electrostatic analogy is also capable of reproducing other dense lyotropic liquid crystal phases at equilibrium[28].

1.3 Photoinitiated assembly

Due to their significance in biological systems, theories of the conformations of bilayer membranes from molecular to continuum models at or near equilibrium have long attracted the interest of biophysicists. But the conformational space of amphiphilic assemblies is vast, and the out-of-equilibrium pathways that a system may traverse are poorly understood.

Various elements of amphiphile assembly can be studied using a model system of 2-oxooctanoic acid (OOA)[1, 29]. Initially, OOA is dissolved in an aqueous solution at low pH to ensure that the carboxylate group is protonated. UV radiation is introduced to catalyze the formation of a free radical at the second, ketone carbon (C2), which can dimerize with another free radial in its vicinity to form dihexyltartaric acid (DHTA). The structural formulae of the aforementioned molecules are shown in Fig.1.1 and Fig.1.2.

The chemical transition then triggers a series of events that lead to the formation of aggregates that range from 100 to 1000 nm in diameter. The progress of aggregation can be monitored by dynamic light scattering (DLS) because aggregates of different sizes scatter a laser beam differently. Two notable features of the aggregate size time evolution are an initial hour-long induction time and a final aggregate population with a narrow size distribution.

This thesis details an exploration of amphiphilic assemblies with tools of theory and simulation, using the photoinitiated assembly of OOA and DHTA as a model system, to paint a fuller picture of the dynamics and structures of amphiphilic assemblies via a bottom-up understanding of the underlying molecular properties.

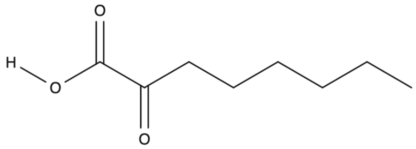

Figure 1.1: Structural formula of 2-oxooctanoic acid (OOA).

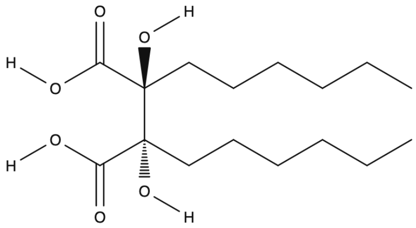

Figure 1.2: Structural formula of one of the stereoisomers of dihexyltartaric acid (DHTA).

Chapter 2

Molecular Dynamics Simulations of OOA and DHTA

2.1 Motivation

The photoinitiated assembly of OOA and DHTA can be studied using molecular dynamics (MD) simulations of fully atomistic detail. The reason for doing so is not to reproduce the events observed over experimental timescales, which would be too long (hours) and the system too large (centimeters) for even the most high-performing computer simulations today; but instead to investigate the building blocks and elementary mechanisms of assembly without making assumptions a priori. Contrary to the intractably long simulation time that a full simulation of the experimental system poses, the elementary steps' timescale (nanoseconds) and molecular lengthscale (nanometers) are well within reach of computer simulations. Once these physical characteristics of the molecular mechanism are determined, they form the foundation for other techniques such as master equations and field theory, which extend the modeling to experimental scales. In this chapter, I present the results of MD simulations using the OPLS-AA force field [30] and discuss their implications for building larger-scale models detailed in subsequent chapters. All simulations were performed using the GROMACS simulation engine[31]. Analysis and biased sampling were carried out using the open-source, community-developed PLUMED library[32], version 2.6[33].

2.2 Molecular configurations in bulk environments

OOA is a linear, 8-carbon alpha-keto carboxylic acid that is fully protonated under experimental conditions (pH 2). Photo-catalyzed dimerization at the keto-carbon produces the double-tailed DHTA with 6 carbons in each tail. The sites of dimerization become chiral, but in the absence of any evidence for enantio-selectivity the mixture is likely racemic with 25% (R,R)-DHTA and 25% (S,S)-DHTA, in addition to the diastereoisomer

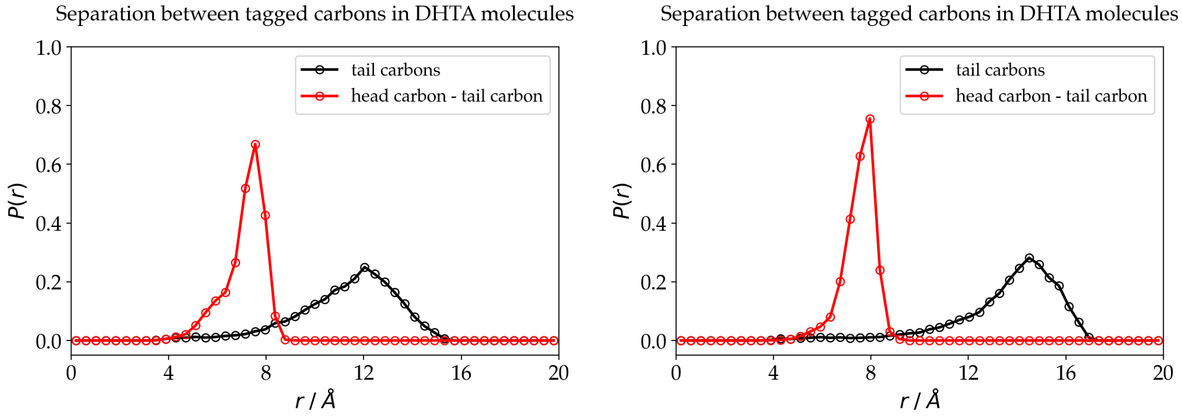

Figure 2.1: Probability density of splay angle of DHTA in the aqueous phase (left) and the organic phase (right).

(R,S)-DHTA that accounts for another 50%.

For the mixture of OOA and DHTA at a fixed total volume fraction in an aqueous solution, dimerization is expected to reduce the solubility of the mixture because the ideal (entropic) part of the solution's chemical potential is reduced by $k_B T \ln 2$.

Both OOA and DHTA are 1 nm long from the carboxylic acid (head) carbon to last carbon of the alkyl tail. This implied the molecular size is borderline for the hydrophobic effect to cause the tails to collapse against each other because water might still be able to accommodate the presence of tails without breaking its hydrogen bonding network. Statistics for tail alignment are shown in Fig.2.1.

In these measurements, a DHTA is solvated in TIP3P water. Its alcohol carbons (C2,C2') are tagged as the head carbon, and the last carbons (C8,C8') on each tail are tagged as the tail carbon. The C2-C8, C2'-C8' and C8-C8' distances are measured. The probability densities show that when a single DHTA molecule is solvated in water, the tail carbons are at about twice the distance between a head carbon and a tail carbon. This can be understood by inspecting the two possible molecular configurations of DHTA, one in parallel alignment and the other one splayed, which are illustrated in Fig.2.2. The C2 and C8 carbons are highlighted in green. The molecular dynamics data point to a preference for a splayed arrangement. The same splay is observed in simulations of the organic bulk liquid that consists purely of DHTA molecules.

When the tails on a DHTA molecule are splayed to more than 90 degrees and sometimes to 180 degrees, the molecule takes up a tail-head-tail, or hydrophobic-hydrophilic-hydrophobic configuration instead of the conventional head-tail, or hydrophilic-hydrophobic configuration. Such a molecular configuration implies that the single-tail OOA and double-tail DHTA are now distinct building blocks. While OOA is a more typical surfactant that has the head-tail configuration, DHTA is not.

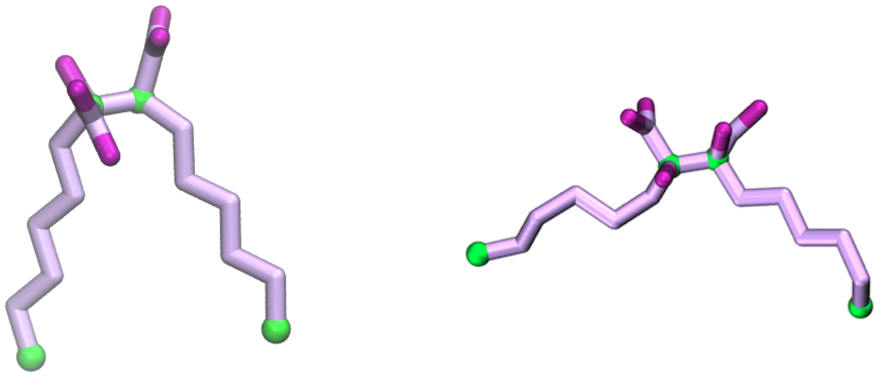

Figure 2.2: Molecular configuration of DHTA in the conventional head-tail arrangement (left) and in the tail-head-tail arrangement (right).

These observations contrast with those of most long-chain biological lipids that contain 12 to 18 carbons. Due to their greater lengths, the tails prefer to align in parallel with minimal splay. These lipids are the main structural components of biological membranes, so parallel alignment is essential for maintaining a stable lamellar structure. It has been hypothesized that tail splay mediates the interactions between two membranes and eventually causes fusion.

2.3 Dynamics in the organic bulk

During the simulations of bulk organic phases at fixed pressure, the density of bulk DHTA was observed to relax more slowly (> 10 ns) than that of bulk OOA (< 0.1 ns). Even after density appears to have reached its steady-state value, the density fluctuations were greater for DHTA than for OOA at the same volume fraction, pressure, and temperature. Therefore, one or more slow molecular timescales must be present in DHTA but not in OOA to form the bottleneck for the bulk organic phase of DHTA to relax at least two orders of magnitude more slowly than OOA does.

By inspection of the simulation trajectories integrated at room temperature (298 K), relaxation of the angle θ between the C2-C8 and C2'-C8' tail vectors appears to be a slow molecular process. As computation becomes increasingly costly for longer simulation time, it was unfeasible to let the simulation run for much longer at room temperature to determine the true relaxation time of order parameter θ.

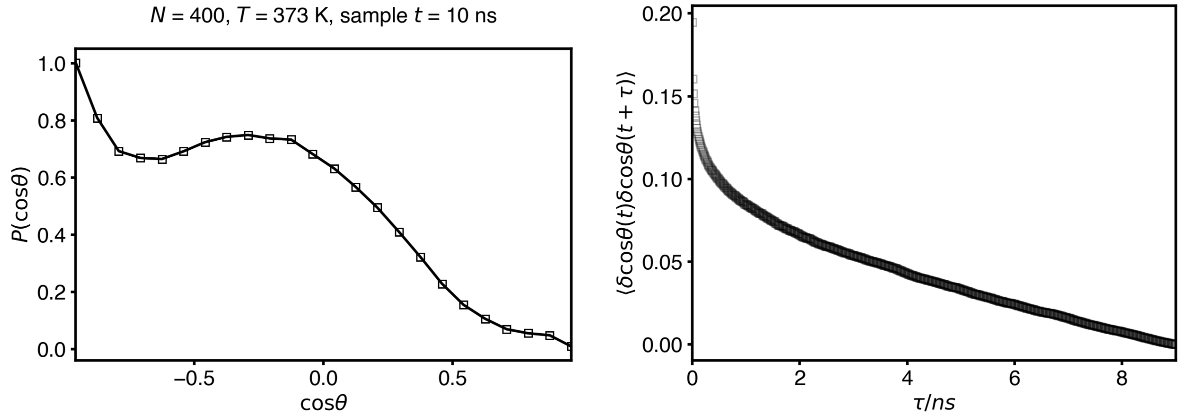

Figure 2.3: Probability distribution of the angle θ between two tails on the same DHTA molecule, averaged over 400 molecules in isotropic bulk environment at 373 K.

Therefore, the simulation box was heated to 373 K to measure a shortened correlation time, which was found to be around 10 ns as shown in Fig.2.3. Over a sampling time of 10 ns, the probability distribution of $\cos\theta$ for the box of 400 DHTA molecules contains up to 800 independent samples. The distribution shows a strong preference for antiparallel alignment of the tails ($\cos\theta = -1$) and a minimal preference for parallel alignment ($\cos\theta = +1$). A second peak is also observed when the two tails are at a slightly greater than right angle to each other.

The data indicate that at room temperature the configurations of DHTA molecules in its bulk organic phase relax over timescales much longer than 10 ns, possibly impeding other processes involved in the nucleation and growth, such as the adsorption-desorption of monomers and surface relaxation pathways necessary for coalescence and fragmentation during aggregation. The precursor OOA, with only one tail, would not be expected to exhibit such behavior.

Another dynamical quantity of interest for studying surface mechanisms and the aggregation pathways, in general, is the diffusion coefficient. Fig.2.4 shows the root mean squared displacement (MSD) and probability distribution of displacements of DHTA molecules in the bulk organic phase at room temperature. DHTA ceased to move after the initial relaxation period of 10 ns. The diffusion coefficient is effectively zero.

Heating the system to higher temperatures (348 K, 398 K, and 500 K) reveals that the linear regime of MSD is gradually regained, as demonstrated in Fig.2.5. At 398 K, a regime of conventional diffusion where MSD is proportional to time appears with a diffusion coefficient of $0.08 \times 10^{-10}\mathrm{m}^2/\mathrm{s}$. When further heated to 500 K, diffusion accelerated to $10.70 \times 10^{-10}\mathrm{m}^2/\mathrm{s}$. For comparison, a typical diffusion coefficient for globular proteins in solution is on the order of $10^{-10}\mathrm{m}^2/\mathrm{s}$, whereas water's self-diffusion coefficient is $23.0 \times 10^{-10}\mathrm{m}^2/\mathrm{s}$ at room temperature. The dramatic increase in the diffusion coefficient upon

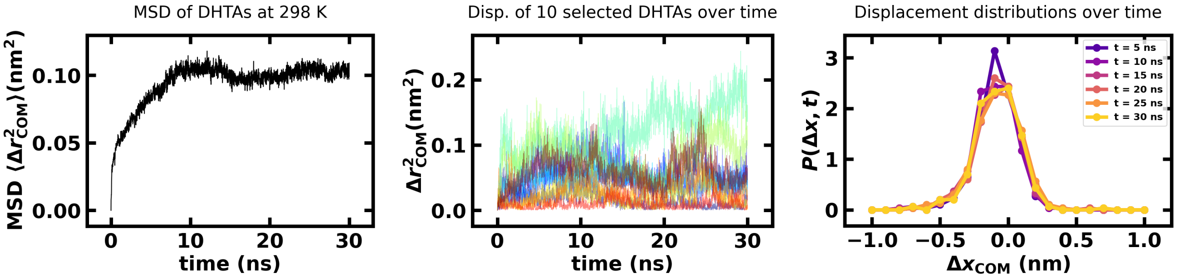

Figure 2.4: Displacement of 10 DHTA molecules, MSD averaged over the whole system, and probability distribution of displacements along the x-axis for bulk DHTA at 298 K over 30 ns.

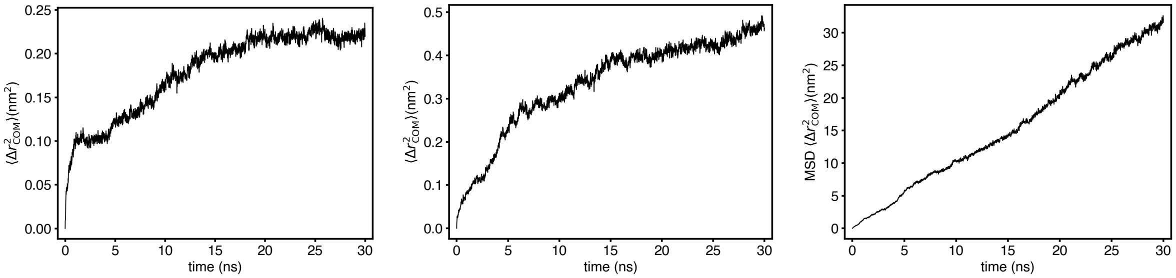

Figure 2.5: MSD of bulk DHTA at increasing temperatures (348 K, 398 K, and 500 K).

increases in temperature implies a strongly T-dependent mobility μ the Einstein relation,

$$D = \mu k_B T$$

where μ is the mobility. So we must conclude that the DHTA is either a solid or a highly viscous fluid at room temperature, but undergoes a transition to a much less viscous liquid state upon heating to higher temperatures (over 398 K).

2.4 Molecular configurations near interfaces with water

To further understand the structural consequences of different head group and tail group arrangements between OOA and DHTA, structural gradations between the aqueous bulk environment and the organic bulk environment can be investigated by molecular dynamics simulations. In these simulations, the systems contain a slab of organic material (either OOA or DHTA) at the center of the box. A slab of water of 2 nm thick is added along the x-axis, padding either side of the slab to create two interfaces with the organic

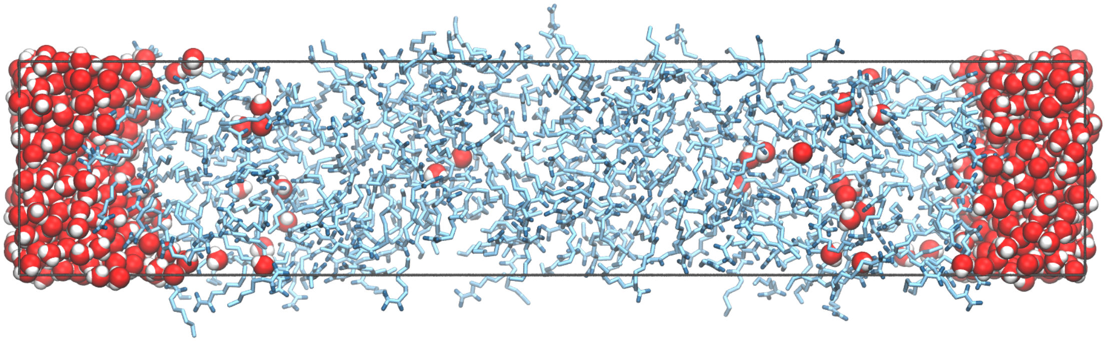

Figure 2.6: Snapshot of OOA at water-organic interface

phase. Snapshots of the simulation boxes are shown in Fig.2.6 and Fig.2.8.

Deep inside the bulk phase, molecules have no orientational preferences in the isotropic environment. Closer to the interface with water, the broken symmetry is expected to induce orientational ordering. This is observed in the case of OOA's interface with water in Fig. 2.7. Densities of C2 head carbons and C8 tail carbons and the angle α between the molecular axis (defined as the vector between C2 and C8) and x-axis of the box are plotted as a function of their x-coordinates. The strongest layering is observed at both interfaces ($x = -6$ nm and $x = +6$ nm). The layer is attenuated deeper into the interface but extends at least 6 nm into the interface.

The same is not observed in the interface simulation data of DHTA (illustrated in Fig.2.8), where the organic phase no longer has a clear preference for the formation of a micellar or lamellar phase despite the broken symmetry at the interface. The density profiles of head carbons (C2, C2') and tail carbons (C8, C8') are not showing any clear layering. Angle α shows essentially no preference for any direction.

Similarly, Fig.2.10 shows the angle θ between two tails inside a DHTA molecule at different depths from the interface. No significant change in angular preference is observed for different depths. The reason behind the absence of interface-induced ordering is likely the slow dynamics, preventing equilibration in our simulations as detailed in the previous section.

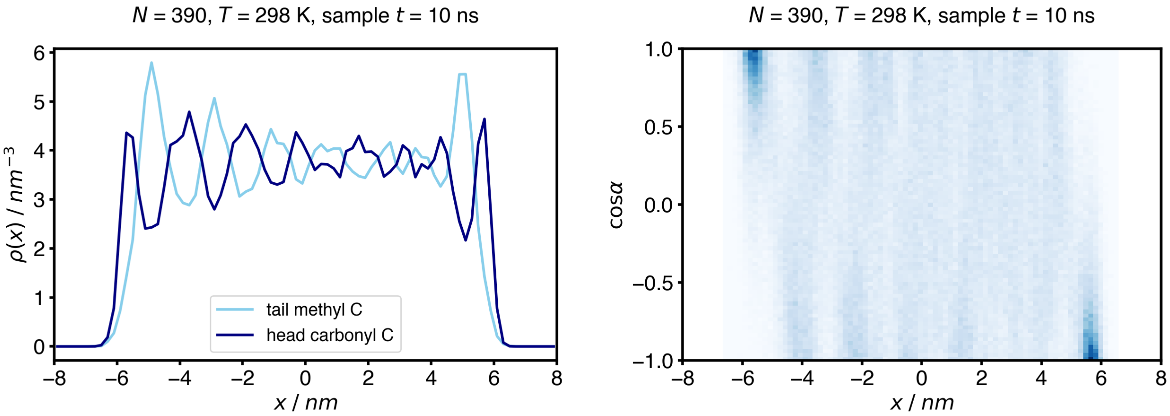

Figure 2.7: Orientation of OOA at water-organic interface

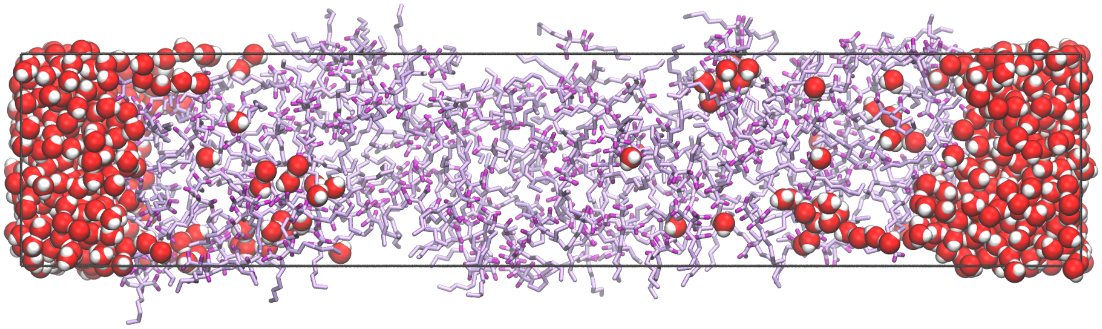

Figure 2.8: Snapshot of DHTA at water-organic interface

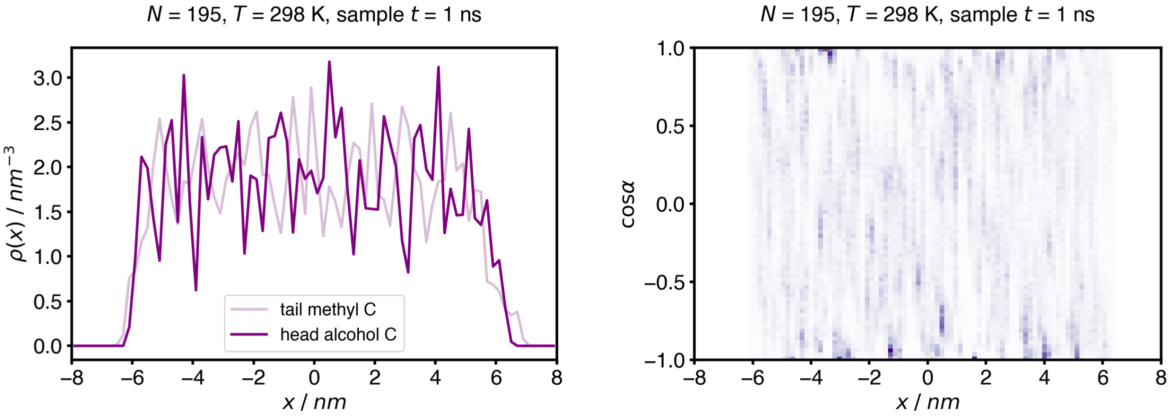

Figure 2.9: Orientation of DHTA at water-organic interface

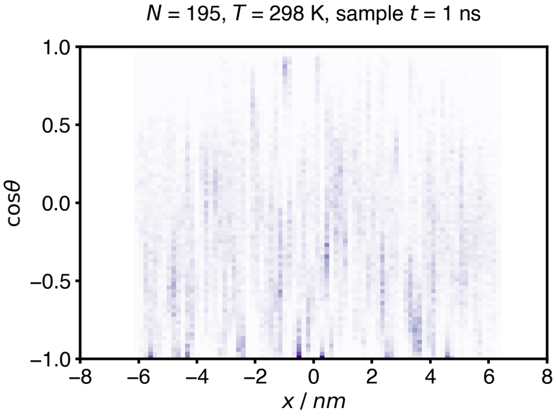

Figure 2.10: Snapshot of DHTA inter-tail angle

2.5 Molecular configurations in smaller clusters

The assembly of OOA and DHTA is a process that starts from the bulk aqueous environment and ends with large ($>$ 100nm) aggregates inside of which a bulk organic environment emerges. Regardless of the specific pathways that could lead from the aqueous to organic phases, the system is likely to go through a series of smaller clusters. Simulations can assess the structure and thermodynamics of such clusters that contain different numbers of OOA (N_1) and DHTA (N_2).

Fig.2.11 shows the radial density profile of carbons in a small cluster that comprises purely DHTA ($N_1 = 0$, $N_2 = 30$). Following the same numbering scheme as in previous sections, the carboxylic acid carbons are numbered C1 (red) and the terminal tail carbons C8 (blue). Farthest away from the center of mass of the cluster, there is no density of carbon but a constant, bulk density of water. Moving towards the center of mass, we first encounter a broad peak of C1, followed by consecutive peaks from C2, C3, ..., and C8, each with successively broader distributions. While the first carbons, C1 to C4, are confined to the outer layer of the cluster and have a negligible density within 0.3 nm of the center, the last carbons, C5 to C8, show significant density throughout the cluster even though the peaks appear near the center. In other words, there is significant tail group density near the cluster surface and thus significant exposure of the tail carbons to water.

Our radial density profiles for DHTA carbons compare well with the radial density profiles of DPPC micelles in literature[34]. This contrasts with the conventional mental image of a surfactant micelle, which has head groups facing strictly outwards and tail groups strictly inwards. Having significant tail group density overlapping with the density of water means that the mental image is only an approximate description of the radial density peaks rather than a complete description of the micellar surface structure.

Similar tail exposure to solvent is seen in mixed micelles of OOA and DHTA. Ten of the DHTA molecules in the previous system are replaced with OOA ($N_1 = 10$, $N_2 = 20$) and equilibrated to produce the radial density profiles shown in Fig.2.12. The OOA carbons and DHTA carbons of the same number generally have the same peak location, which means both species have head groups at the interface with water, and tail groups mostly near the center of the cluster. While the radial densities of this mixed micelle appear similar to those of the pure micelle, the snapshot of the mixed micelle in Fig.2.13 shows that angular distribution of OOA and DHTA might not be homogeneous. Throughout the 30-ns trajectory from which this snapshot was taken, OOA segregates from DHTA into one or two caps on the surface.

Both radial and angular segregation of the two species can be seen in larger clusters such as the $N_1 = 320$, $N_2 = 60$ cluster shown in Fig.2.14. The cluster's composition is chosen to mimic the early molecular composition of the experimental system. The angular segregation is visible in the snapshot, where the cluster surface is studded with around a dozen small groups of DHTA.

The radial density profiles in Fig.2.15 and Fig.2.16 suggest that the large cluster is

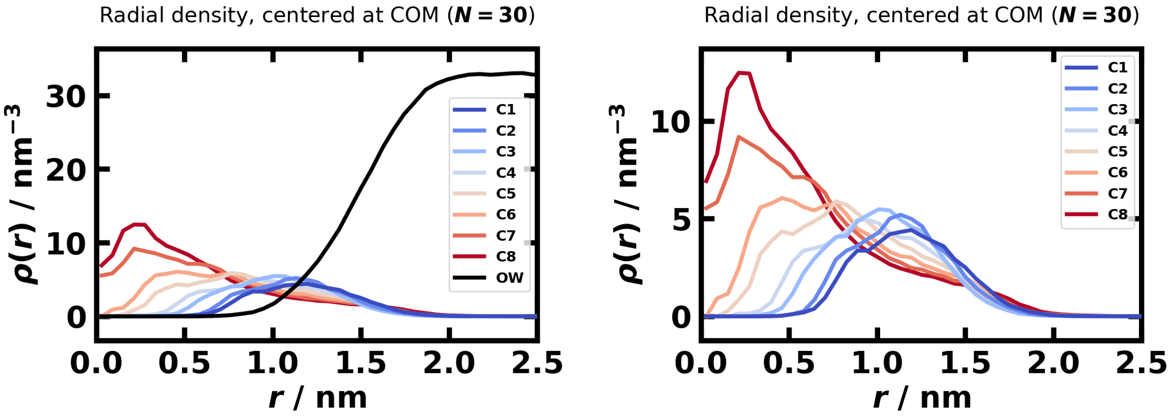

Figure 2.11: Radial density profiles of a $N_1 = 0$, $N_2 = 30$ cluster. Water included in left plot for comparison.

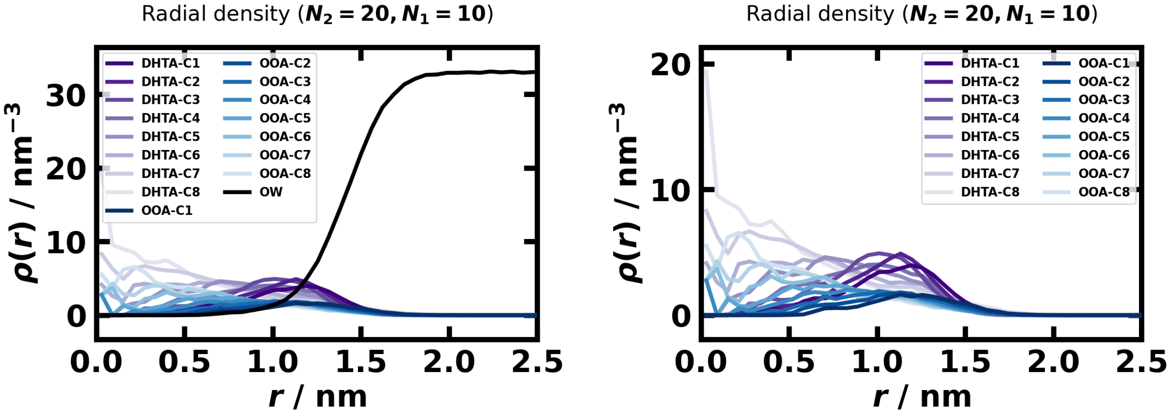

Figure 2.12: Radial density profiles of a $N_1 = 10$, $N_2 = 20$ cluster. Water included in left plot for comparison.

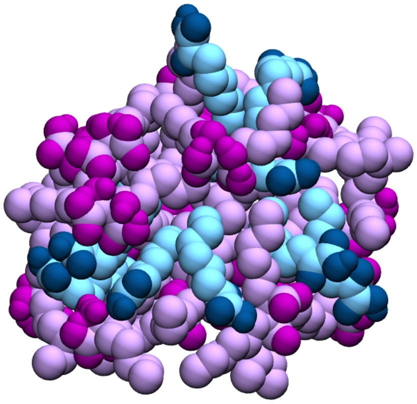

Figure 2.13: Snapshot of a $N_1 = 10$, $N_2 = 20$ cluster showing OOA head carbons (dark blue) and OOA tail carbons (light blue) segregating from DHTA head carbons (dark purple) and DHTA tail carbons (light purple). Solvents omitted for clarity.

made of three OOA layers and two DHTA layers. The DHTA head carbons have peaks at $r = 2.7$ nm and $r = 1.4$ nm, which means they are concentrated at the interface with water and approximately one monolayer's distance from the interface.

The OOA head carbon peaks are broader and hence more evenly spread throughout the cluster, with its outermost peak sitting further into the solvent than the DHTA head carbons do at $r = 2.9$ nm. A second broad peak appears at $r = 1.0$ nm. Surprisingly, the third head carbon density peak appears sharp at the center of mass of the micelle, where both OOA and DHTA tail carbon densities are low. The center of the cluster is occupied by OOA head groups. The radial segregation could contribute to the core-shell organization of large aggregates experimentally observed under the electron microscope.

2.6 Free energy of DHTA association

Homogeneous nucleation is an important molecular mechanism because it necessarily precedes all aggregation pathways. However, simulating nucleation events using all-atom molecular dynamics is difficult because the experimental concentration of OOA and DHTA in the cuvette is no more than 5 mM, which is equivalent to a molar ratio of the order 1 solute to 10000 waters. On top of that, the rare nature of nucleation events further reduces the likelihood of ever observing one event in any computationally feasible trajectory length.

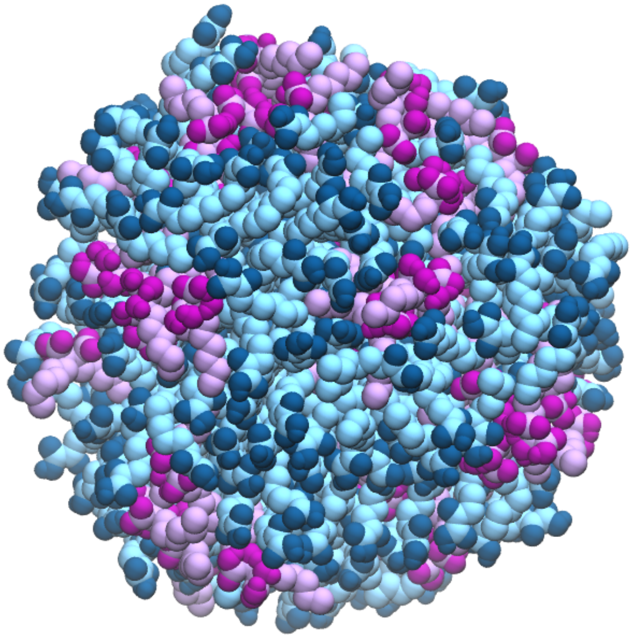

Figure 2.14: Snapshot of a $N_1 = 320$, $N_2 = 60$ cluster. Solvents molecules are omitted for clarity.

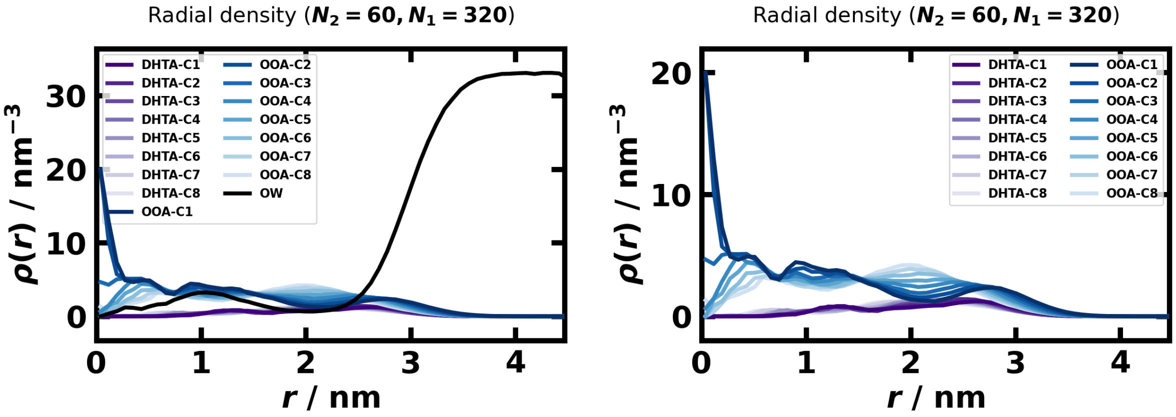

Figure 2.15: Radial density profiles of a $N_1 = 320$, $N_2 = 60$ cluster. Water included in left plot for comparison.

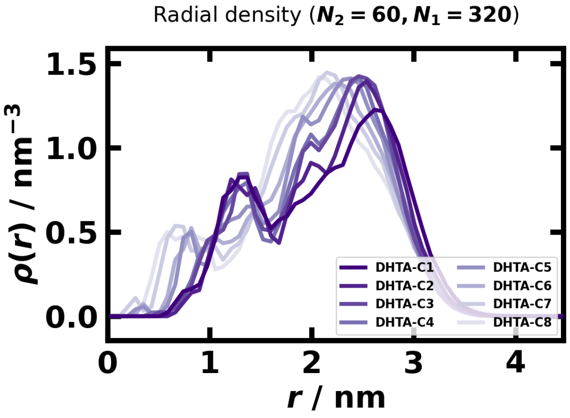

Figure 2.16: Radial density profiles of a $N_1 = 320$, $N_2 = 60$ cluster. Only DHTA carbons are shown for clarity.

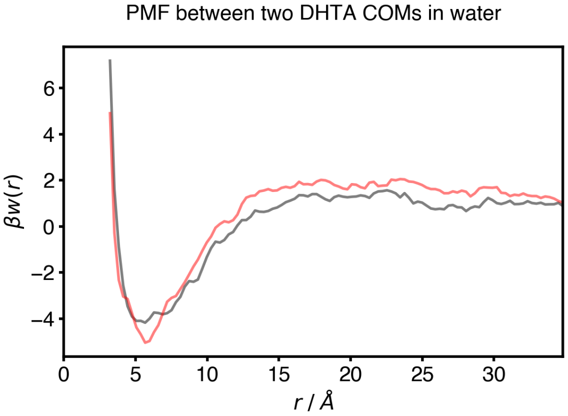

Figure 2.17: Potential of mean force between two DHTA molecules in aqueous solution. Grey and red are two segments of the trajectory for error estimates.

Instead of using brute-force simulations to observe large systems over a long time, umbrella sampling[35] can be used to measure the reversible work, or potential of mean force (PMF), to move the system along a reaction coordinate. For aggregation, the natural coordinate of choice is the distance between the center of mass of the cluster and the center of the mass of the molecule to be added to the cluster, r. Without prior knowledge of the free energy along the coordinate, it is difficult to guess a bias potential $U_{\text{bias}}(r)$ that counteracts the PMF $w(r)$. Therefore, parabolic bias potentials with spring constant $k = 300\text{kJ/mol} \cdot \text{nm}^2$ were placed to be at a series of 12 points to span the whole range of r from 0 to 2.4 nm.

Fig.2.17 shows the PMF between two DHTA molecules in water. There is no barrier to the aggregation of the two molecules. The effective interaction between the centers of mass of two DHTA molecules is attractive and short-ranged, extending through no more than 1.2 nm, which is approximately the length of a tail. The most favorable distance of contact is 0.6 nm with a well depth of $-4k_BT$, but pushing the two molecules any further closer quickly becomes highly unfavorable due to steric clashing.

Adding DHTA monomers to clusters of larger sizes becomes more thermodyanmically favorable. Fig.2.18, Fig. 2.19, Fig.2.20, and Fig.2.21 shows that the well-depth grows from $-6k_BT$ to $-10k_BT$, and the basic shape of the PMF is the same: there are no barriers to monomer addition, and pushing a monomer deep into the core of the cluster is unfavorable. The latter observation is in agreement with the radial density profile of pure DHTA micelles with $N_1 = 0$ and $N_2 = 30$ because for the head-out, tail-in arrangement to be satisfied, a DHTA molecule should not preferentially sit at the core of the cluster.

The PMFs can be integrated to yield equilibrium constants between an $(N-1)$-cluster, a monomer, and an N-cluster:

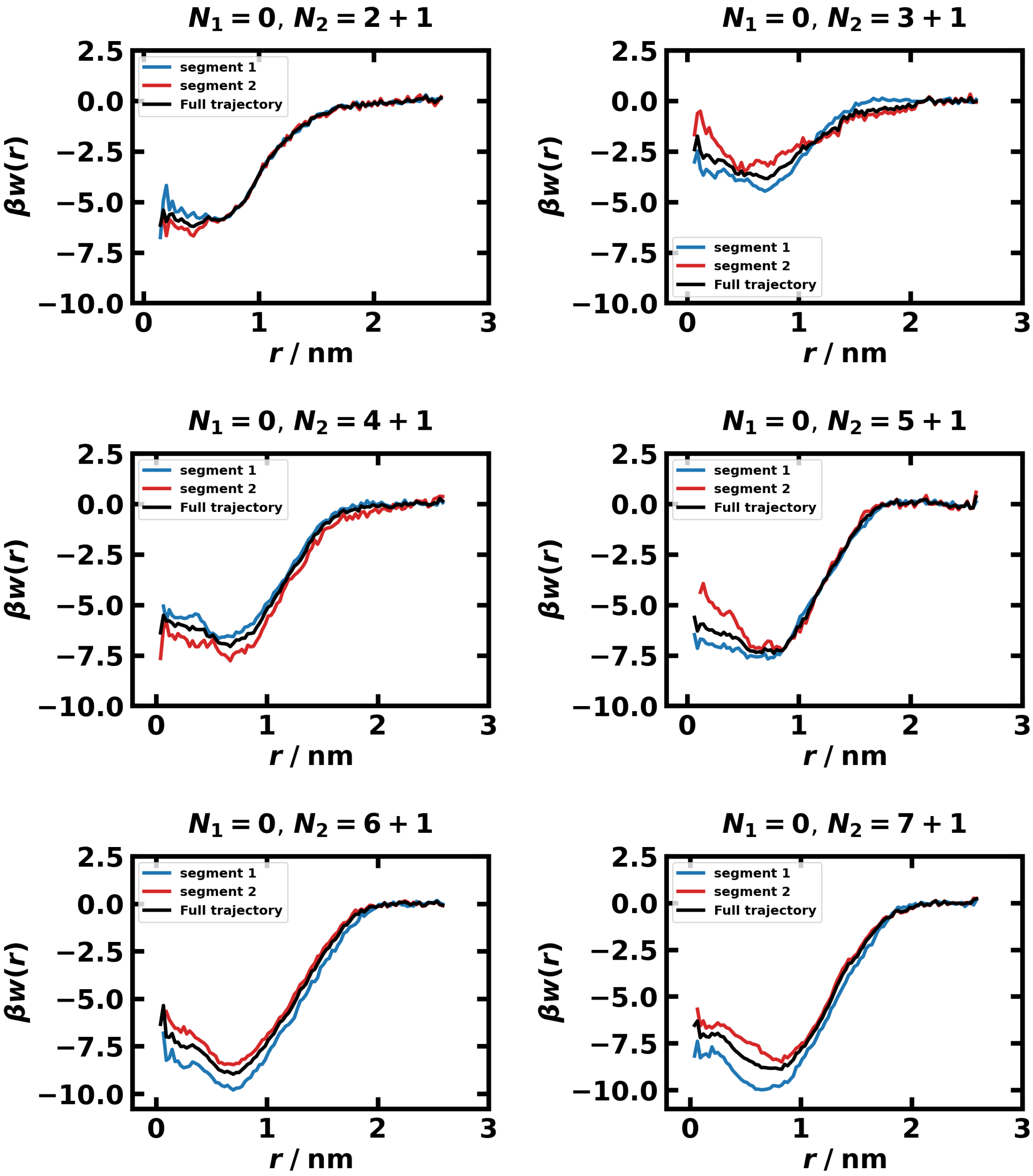

Figure 2.18: Potential of mean force for adding one DHTA to a cluster of 2, 3, 4, 5, 6, and 7 DHTA molecules in aqueous solution. Blue and red are two segments of the trajectory for error estimates.

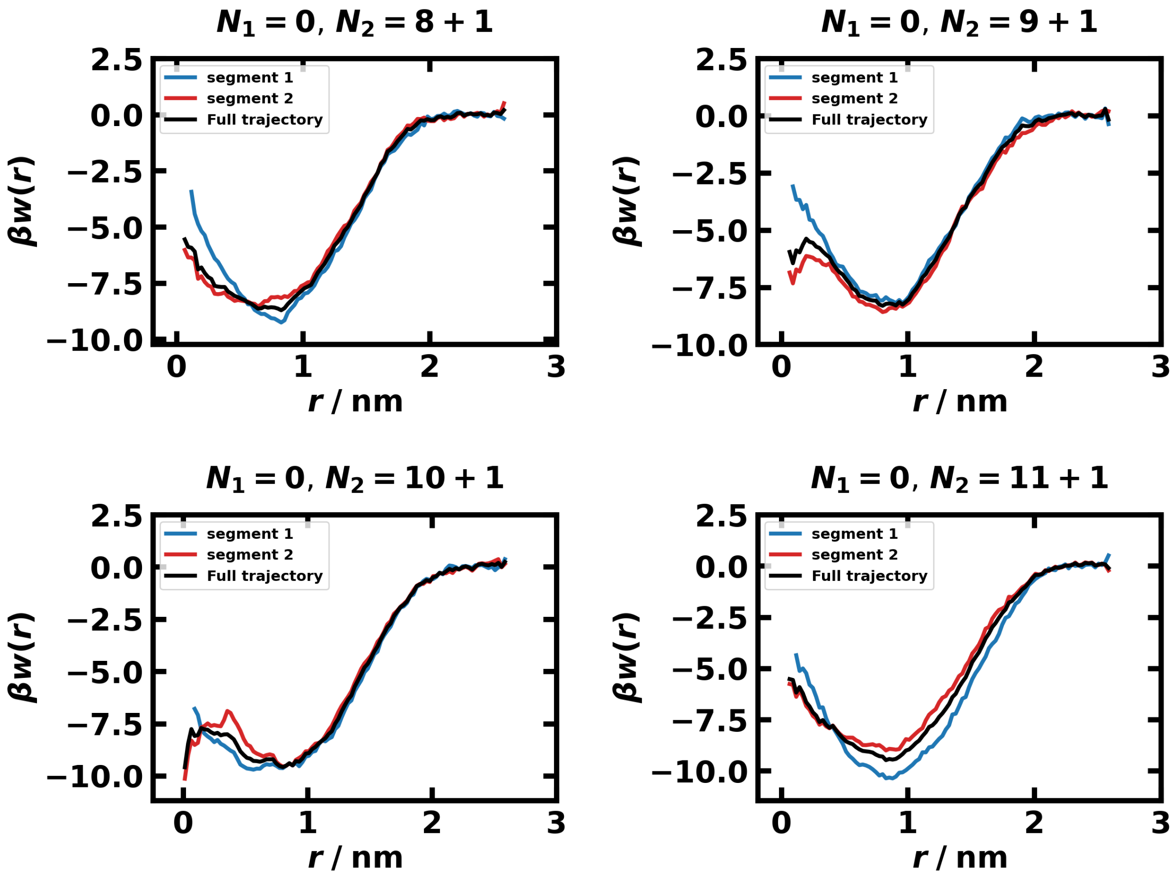

Figure 2.19: Potential of mean force for adding one DHTA to a cluster of 8, 9, 10, and 11 DHTA molecules in aqueous solution. Blue and red are two segments of the trajectory for error estimates.

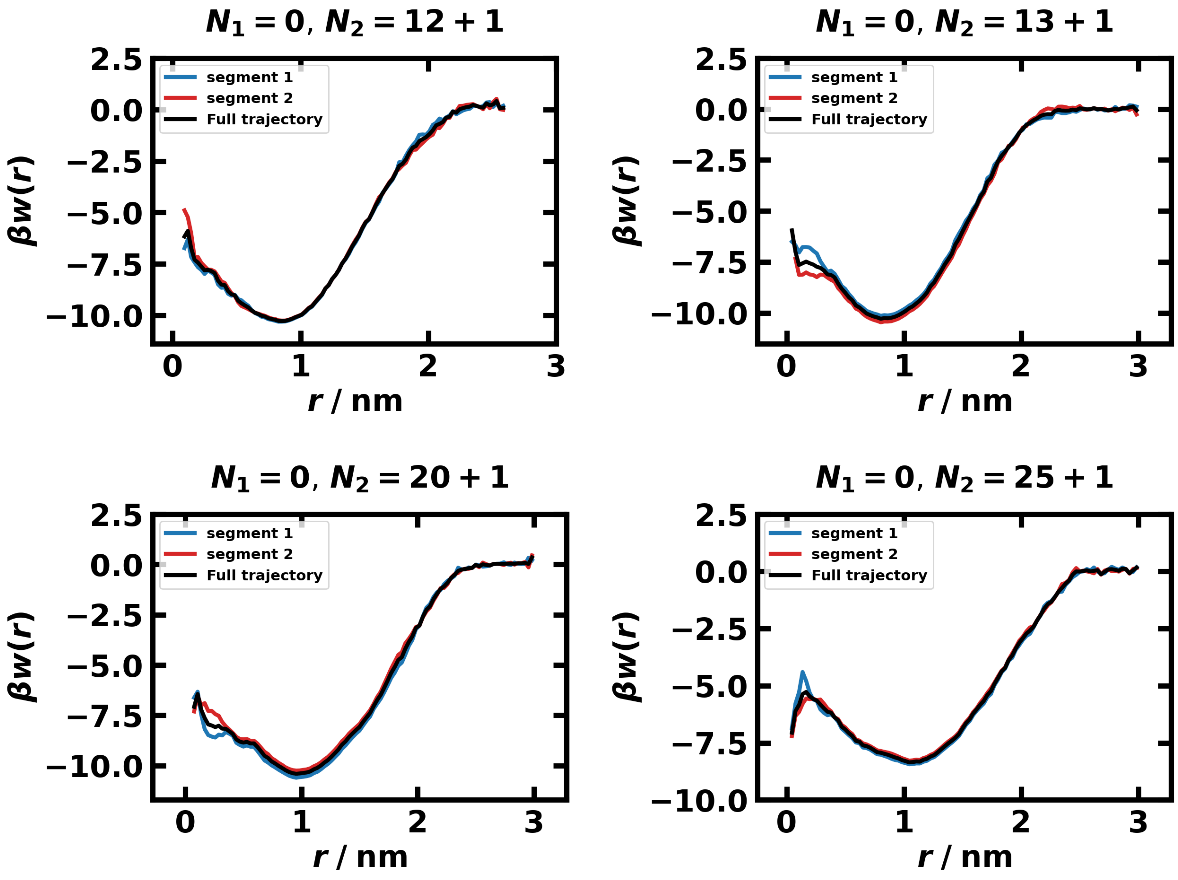

Figure 2.20: Potential of mean force for adding one DHTA to a cluster of 12, 13, 20, and 25 DHTA molecules in aqueous solution. Blue and red are two segments of the trajectory for error estimates.

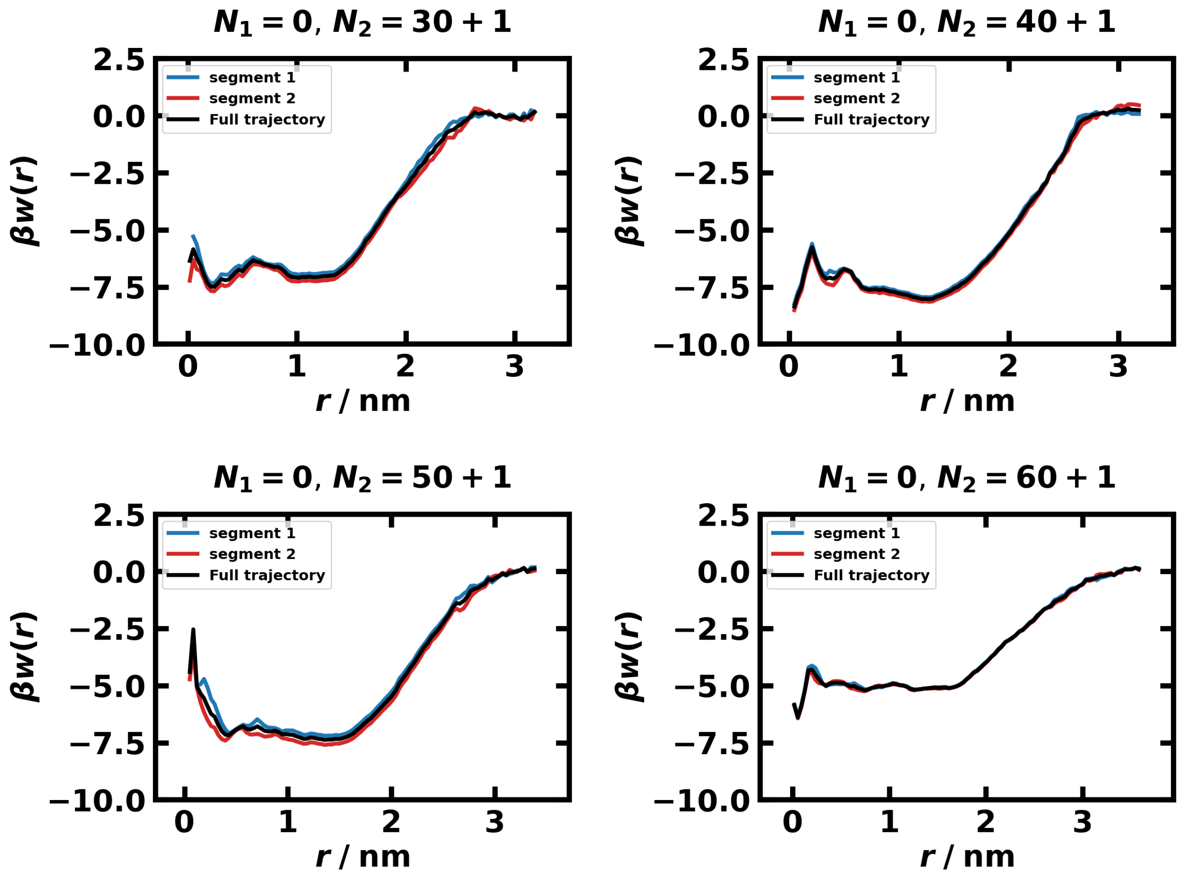

Figure 2.21: Potential of mean force for adding one DHTA to a cluster of 30, 40, 50, and 60 DHTA molecules in aqueous solution. Blue and red are two segments of the trajectory for error estimates.

$$K_N = \frac{1}{N}\int d\mathbf{r}\; e^{-\beta[w(r)-w(\infty)]}$$

and for a cluster of size N, the equilibrium constants $\{K_i\}_{i=1}^{N}$ determine the free energy as a function of N, at any solution monomer concentration ρ_1:

$$\begin{aligned} F(N) &= -k_B T \sum_{i=1}^{N} \ln\left(\rho_1 K_i\right) \\ &= -k_B T \sum_{i=1}^{N} \ln K_i - N k_B T \ln \rho_1 \end{aligned}$$

where we see the solution DHTA monomer concentration adds a linear shift to the function $F(N)$. Fig.2.22 shows the free energy as a function of cluster size at different solution monomer concentrations up to 1.7 mM. For the purpose of calculating the nucleation rate, the relevant features of $F(N)$ are the critical nucleus size at which the barrier appears, $N^\ddagger$, and the height of the barrier $\Delta F^\ddagger$:

$$\Delta F^\ddagger = F(N^\ddagger) - F(1)$$

Due to the linear shift from the $-Nk_BT\ln\rho_1$ term, lower solution monomer concentration causes a higher barrier to appear at larger N. Therefore to observe the barriers for nucleation for ever lower concentrations, ever larger N values will need to be simulated.

The experimental concentration at which homogeneous nucleation of DHTA occurs was not directly monitored throughout the reaction time, but the concentrations of DHTA in the post-photolysis solution as well as the aggregates were measured relative to that the concentrations of OOA in these environments at the end of the assembly reaction. The total mass of inside the reaction vessel was conserved to be equal to the single-tail precursor OOA's initial concentration of 3 mM. At the end of the reaction, two thirds of the organic material was incorporated into aggregates and one third remained in the aqueous solution. The aggregates contained 65% DHTA and 18% OOA in addition to 17% other photochemical products, whereas the aqueous solution contained 38% DHTA and 32% OOA in additions to 30% other photochemical products. Overall, the total (solution and aggregate) molar concentration of DHTA does not exceed 0.84 mM at the end of the reaction, of which 0.65 mM is in the aggregates and 0.19 mM is in solution.

At lower concentrations, such as the 0.19 mM curve shown on in the figure, the free energy of the cluster has an initial barrier around $N = 10$ DHTA molecules, which is expected for a hydrophobic phase nucleating out of solution. However, after the barrier, instead of rapidly decreasing, the the curve reaches a metastable trough and turns back up again. This is the result of the DHTA molecule being amphiphilic rather than fully hydrophobic, and is consistent with our understanding that micelles have a preferred size. The preferred size depends on the concentration of DHTA in the aqueous phase. At the

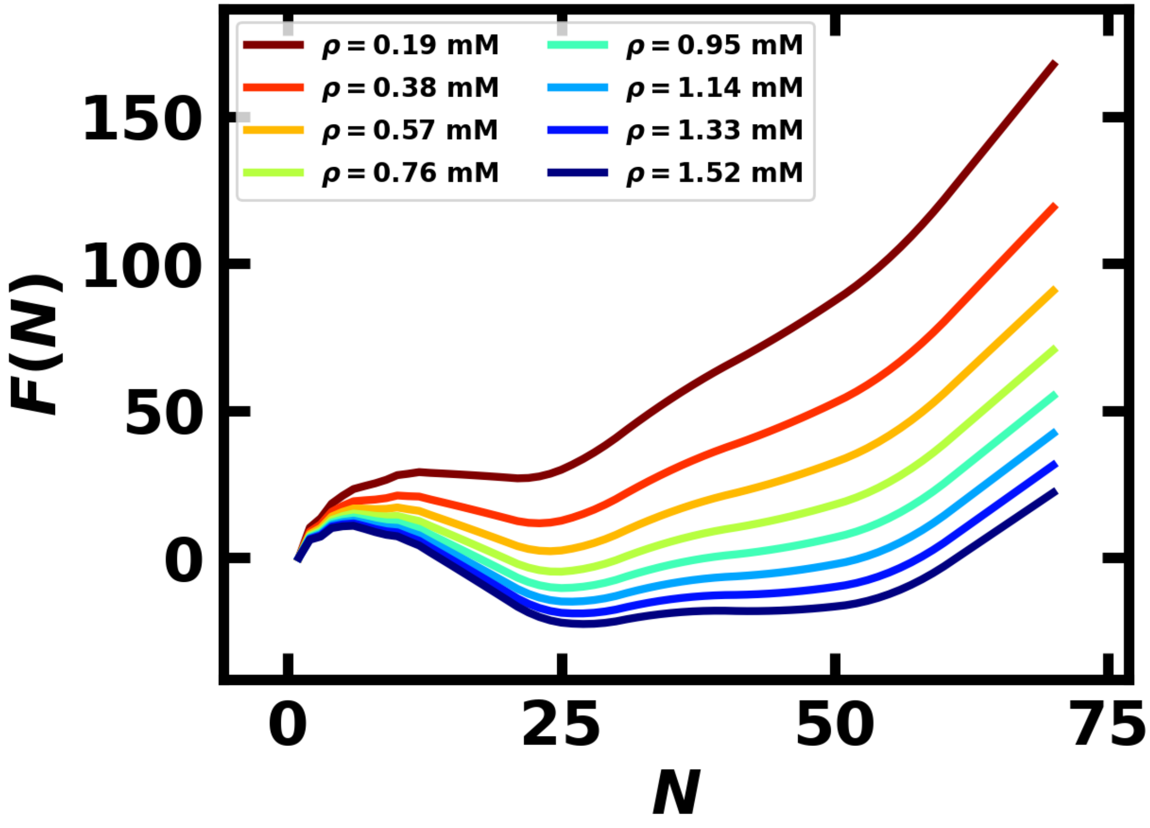

Figure 2.22: Free energy (in units of k_BT) as a function of number of molecules in the cluster. Colors are for different soluble DHTA monomer concentration in equilibrium with the cluster as shown in the legend. The dark red curve (0.19 mM) is the closest to experimental conditions.

experimental concentration (0.19 mM) the metastable size is 24 DHTA molecules, whereas
.

Within the cluster size range of 2 to 61 DHTA molecules, the radii of gyration observed in simulations follow the compact scaling law

$$R_g(N) = R_g(1)N^{1/3}$$

where $R_g(1)$ is 0.402 nm. A cluster of 24 DHTA molecules exhibits a radius of gyration of 1.16 nm. The two-orders-of-magnitude discrepancy between the predicted pure-DHTA cluster radius (1.16 nm) and and experimentally observed mixed OOA and DHTA cluster radii (over 100 nm) suggests homogeneous nucleation of DHTA molecules alone is unlikely to account for all the aggregation in the system.

2.7 Nucleation kinetics of DHTA

In the previous section, we established that the small equilibrium cluster radius predicted from free energy measurements of a pure DHTA system is in disagreement with the large cluster radii observed in experiments of mixtures of OOA and DHTA. It is clear that for the large 100 nm aggregates to form, both OOA and DHTA need to participate in the aggregation pathways.

But the role of homogeneous nucleation of DHTA in relation to cannot be neglected. The first step to any aggregation pathway in solution is homogeneous nucleation. The pure OOA at the experimental concentration is fully soluble, which means the aqueous OOA will not spontaneously nucleate. In this case the nuclei must be composed of either pure DHTA or a mixture of OOA and DHTA. Therefore, regardless of the cluster sizes at the late stages of reaction, the nucleation kinetics of DHTA is relevant to the early stages of the reaction, especially the long induction time.

The free energy profile as a function of cluster size provides the necessary information for estimating the rate of nucleation at various aqueous concentrations of DHTA. One approach is to apply Transition State Theory (TST), which says that once a barrier height has been observed on the free energy $F(N)$, a nucleation rate can then be computed using classical nucleation theory[36, 37, 38]:

$$k_n = Zv\rho_1 e^{-\beta\Delta F^\ddagger}$$

where v is the speed along the reaction coordinate N at the critical nucleus size $N^\ddagger$, and Z is the Zeldovich factor which accounts for the diffusive recrossing. .

An alternative approach to the TST is to use a master equation to model the stepwise association and dissociation that lead up to the critical nucleus size. Given the wealth of information available on the full free energy profile $F(N)$, I will adopt this approach instead of TST.

A master equation describes the time evolution of a system through n possible states. It assigns a probability density to each state that the system may occupy at any time, $\{p_i(t)\}_{i=1}^{n}$, and a transition rate constant to every pair of states, $\{k(i \to j)\}_{i,j=1}^{n}$.

This systematic description applies well to the aggregation pathway of DHTA, where every cluster size N at time t is a state to which we assign a probability density $p_N(t)$. The collection of all such states from cluster size $N = 2$ to cluster size $N = N_{\max}$ written as a vector:

$$\mathbf{p}(t) = \begin{pmatrix} p_2(t) \\ p_3(t) \\ \vdots \\ p_{N_{\max}}(t) \end{pmatrix}$$

The transition rate constants can be written as a matrix:

$$\mathbf{W} = \begin{pmatrix} -w_1^+ & w_2^- & & & \\ w_1^+ & -(w_2^+ + w_2^-) & & & \\ & w_2^+ & & & \\ & & \ddots & & \\ & & & w_{N_{\max}-1}^- & \\ & & & -(w_{N_{\max}-1}^+ + w_{N_{\max}-1}^-) & w_{N_{\max}}^- \\ & & & w_{N_{\max}-1}^+ & -w_{N_{\max}}^- \end{pmatrix}$$

where w_i^+ denotes the rate constant of one DHTA monomer associating to a cluster of size i, and w_i^- denotes the rate constant of dissociating from a cluster of size i. The transitions rate constants with subscripts $5, \ldots, N_{\max} - 2$ are omitted for clarity. Only the diagonal and immediately off-diagonal elements are non-zero because the clusters are assumed to grow one DHTA at a time.

The master equation combines the rate constants and the occupational probabilities:

$$\frac{\partial \mathbf{p}}{\partial t} = \mathbf{W}\mathbf{p}(t)$$

which can be solved to yield the cluster size distribution over time.

2.7.1 Association rates

The rate constants can be found from the free energy profile $F(N)$ presented in the previous section. For the association *rate constants* (with superscript $+$), we are interested in the rate of transition from a state of cluster size N to $N + 1$ by adding a monomer. The *rate* for this process is

$$w_N^+ p_N(t) = k_d(N)\rho_1 p_N(t)$$

where on the right-hand side the transition rate constant w_N^+ has been decomposed into a monomer-independent part $k_d(N)$ and the monomer concentration itself ρ_1. Since $F(N)$ shows no barrier to association, the monomer-independent part $k_d(N)$ can be assumed diffusion-limited,

$$k_d(N) = 4\pi D(N)R(N)$$

where $D(N)$ is the diffusion coefficient of the N-cluster, $R(N)$ is the radial distance from the center of the cluster within which a monomer will experience a net-inward pull.

To estimate the diffusion coefficients $D(N)$, we consider that water has a molecule size of 0.270 nm, which means its radius is about 0.135 nm. Its self-diffusion coefficient is $2.29 \times 10^{-9}\mathrm{m}^2/\mathrm{s}$ at room temperature. The DHTA monomer's radius of gyration is 0.402 nm. The Einstein-Stokes relation says that the diffusion coefficient is inversely proportional to the hydrodynamic radii, which we approximate to be 0.135 nm and 0.402 nm for water and DHTA respectively. So the diffusion coefficient of a DHTA monomer is approximately

$$D(1) = \frac{D_{\text{water}} R_{\text{water}}}{R_{\text{DHTA}}} = \frac{2.29 \times 10^{-9}\mathrm{m}^2/\mathrm{s} \times 0.135\,\mathrm{nm}}{0.402\,\mathrm{nm}} = 7.7 \times 10^{-10}\mathrm{m}^2/\mathrm{s}$$

Given that the DHTA clusters follow the compact scaling law $R \sim N^{1/3}$, we can further write

$$D(N) = \frac{D(1)}{N^{1/3}}\,.$$

From the free energy profiles, $R(N)$ is observed to scale as

$$R(N) = R(1)N^{0.21}$$

where $R(1) = 1.4\,\mathrm{nm}$.

Combining all elements, the association rate constant is

$$w_N^+ = \rho_1 k_d(N) = \rho_1 N^{-0.123} 1.35 \times 10^{-17}\mathrm{m}^3/\mathrm{s}$$

which has the dimension of inverse time.

2.7.2 Dissociation rates

The free energy profile $F(N)$ contains information about the equilibrium among all cluster sizes $2, ..., N_{\max}$. Detailed balance must be obeyed at equilibrium, so the dissociation *rate constants* (with superscript $-$) are related to the association rate constants by

$$\bar{\rho}_N w_N^+ = \bar{\rho}_{N+1} w_{N+1}^-$$

where $\bar{\rho}_N$ is the density of N-clusters at equilibrium. At equilibrium, the left-hand side of this equality can be expanded to

$$\bar{\rho}_N k_d(N)\bar{\rho}_1 = \bar{\rho}_{N+1} w^-_{N+1}$$

and rearranged to

$$w^-_{N+1} = \frac{\bar{\rho}_N \bar{\rho}_1 k_d(N)}{\bar{\rho}_{N+1}}$$

from which we identify the equilibrium constant

$$K_{N+1} = \frac{\bar{\rho}_{N+1}}{\bar{\rho}_N \bar{\rho}_1}$$

and it is related to the free energy profile by

$$K_{N+1} = \frac{\bar{\rho}_{N+1}}{\bar{\rho}_N} \times \frac{1}{\bar{\rho}_1} = e^{-\beta[F(N+1)-F(N)]} \times \frac{1}{\bar{\rho}_1}\,.$$

So the dissociation rate constant of the $(N+1)$-cluster is

$$w^-_{N+1} = \frac{\bar{\rho}_1 k_d(N)}{e^{-\beta[F(N+1)-F(N)]}}$$

for which we also previously found

$$k_d(N) = N^{-0.123} 1.35 \times 10^{-17} \mathrm{m}^3/\mathrm{s}\,.$$

and equilibrium DHTA monomer concentration is

$$\bar{\rho}_1 = 0.19\,\mathrm{mM} = 1.14 \times 10^{23}\,\mathrm{m}^{-3}$$

We can now use the dissociation and association rate constants to infer the non-equilibrium nucleation flux.

2.7.3 Numerical propagation

Numerically propagating the master equation

$$\frac{\partial \mathbf{p}}{\partial t} = \mathbf{W}\mathbf{p}(t)$$

gives the cluster size distribution $\mathbf{p}(t)$ shown in Fig.2.23 and Fig.2.24 for two different equilibrium aqueous concentrations, 1.5 mM and 0.76 mM respectively. The numerical integrations are carried out at a timestep size of 1 fs and the total time propagated is 1 ms. The pre-critical cluster sizes under $N = 10$ are omitted for clarity. The figures also show a linear increase to the total nuclei probability on the right-hand side, from which a nucleation rate can be calculated.

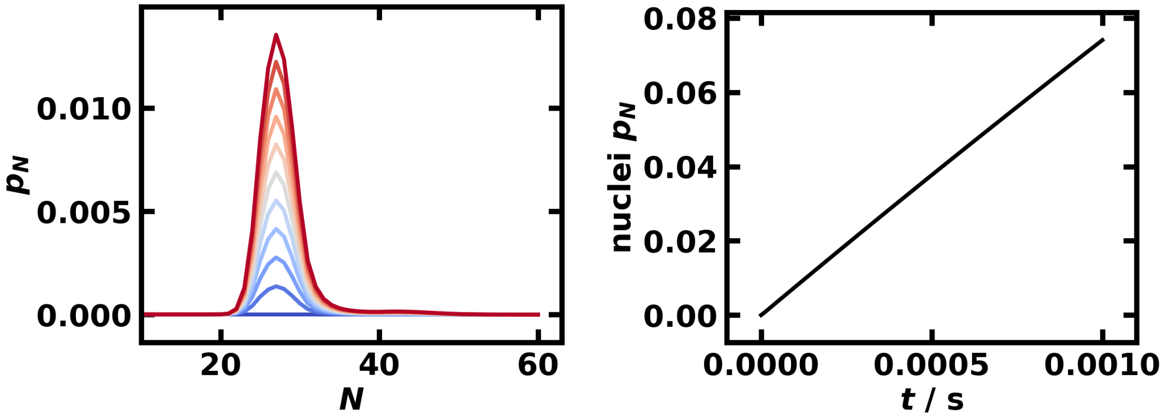

Figure 2.23: Numerical solution to the master equation evaluated at an equilibrium aqueous concentration of 1.5 mM.

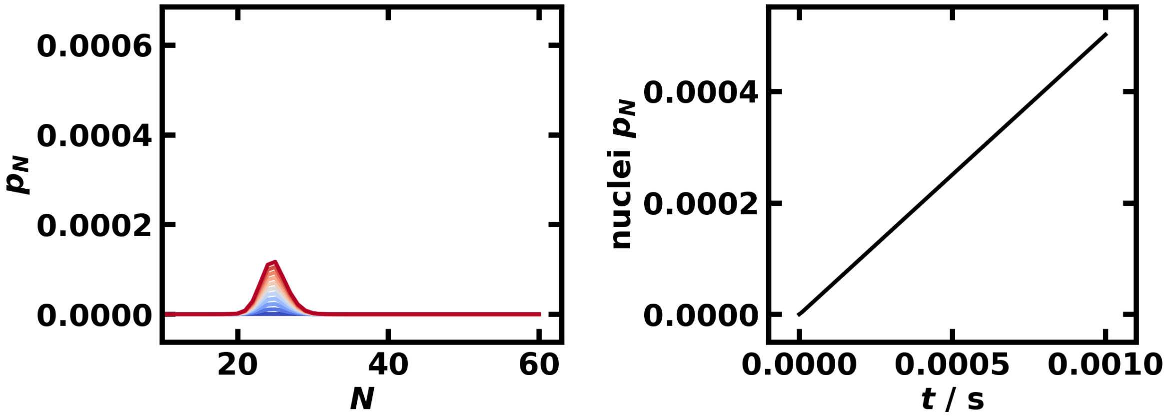

Figure 2.24: Numerical solution to the master equation evaluated at an equilibrium aqueous concentration of 0.76 mM.

While not shown in the figures, the numerical solutions at the terminal experimental aqueous concentration of 0.19 mM did not yield any nuclei during the 1 ms. This is a consequence of $F(N)$'s global minimum sitting in the monomer state $N = 1$ at 0.19 mM (Fig.2.22). This means the aqueous solution will not spontaneously nucleate new DHTA clusters and is consistent with the final stability of the experimental system.

The numerical solutions also inform us about the early stages of the reaction. If all 3 mM of OOA in the starting aqueous solution was converted to DHTA before nucleation occurs, the solution would contain 1.5 mM of DHTA. In that case, there is a 8% probability that the solution has nucleated after 1 ms, which means on average 0.12 mM of DHTA is incorporated into DHTA clusters of average size 24 molecules, or equivalently a cluster density of 5 μM. This implies a nucleation rate of 3×10^{18} events per second per liter.

If only half of the OOA is converted into DHTA before nucleation occurs, the solution would contain around 0.76 mM of DHTA. In that case, there is a 0.05% probability that the solution has nucleated after 1 ms, which means on average 0.38 μM of DHTA is incorporated into DHTA clusters of average size 22 molecules, or equivalently a cluster density of 17 nM. This implies a nucleation rate of 1.0×10^{16} events per second per liter, 300 times slower than that at 1.5 mM.

The experiments yield aggregates of radius 100 nm from a solution of 3 mM of OOA, 2 mM of which is incorporated into aggregates either in the form of OOA or DHTA. Given that the organic material has a density of 900 g/L and the molar weight of OOA is 158 g/mol, the volume of OOA or DHTA in aggregates occupies

$$\frac{2 \times 10^{-3} \frac{\text{mol}}{\text{L}} \times 158 \frac{\text{g}}{\text{mol}}}{900 \frac{\text{g}}{\text{L}}} = 3.5 \times 10^{-4} \text{ of aqueous solution volume.}$$

Each liter of aqueous solution therefore contains

$$\frac{3.5 \times 10^{-4} \text{L} \times 10^{-3} \frac{\text{m}^3}{\text{L}}}{\frac{4\pi}{3} \times \left(100\text{nm} \times 10^{-9} \frac{\text{m}}{\text{nm}}\right)^3} = 8 \times 10^{13} \text{ aggregates / L.}$$

This aggregation concentration takes only 30 μs at 1.5 mM and 8 ms at 0.76 mM. Nucleation of DHTA must have taken place at even lower concentrations than 0.76 mM.

Chapter 3

Aggregation kinetics

3.1 Motivation

The experiments observed the light-activated production of surface-active molecules coupled to subsequent aggregation of these product molecules in solution. At the highest level, the steps leading from solvated 2-OOA molecules to aggregates around 100nm in size can be divided into three categories: photochemistry, interfacial reactions, and aggregation.

In the following, I will list all the possible elementary steps of each category of reactions, and postulate the relative importance of each step for limiting the overall rate of aggregate formation based on what we know about intermolecular interactions of these molecules and what experiments with additional perturbations might have implied. I will then collect this rate-determining subset of elementary mechanisms, combine them into systems of differential rate equations, and solve them numerically to demonstrate simple scaling laws whenever possible.

Given the values of rate constants, these theoretical calculations can inform us how the system's chemical composition, interfacial partitioning, aggregate mass, average aggregate size, and aggregate size distributions evolve. However, it is the reverse procedure that elucidates the underlying physics: given the experimentally measured time evolutions of chemical composition, interfacial partitioning, aggregate mass, average aggregate size, and aggregate size distributions, it should be possible to fit experimental kinetic curves to theoretical calculations and ultimately extract rate constants of the elementary mechanisms.

3.1.1 Photochemistry

The post-photolysis solution appears homogeneous, so we assume that the system is well-mixed. Mean-field kinetics are assumed instead of diffusion-controlled reaction kinetics. Single-tailed molecules S are converted into activated molecules S^* by UV light, which then react with un-activated monomers to yield radicals S°. We assume only single-

reaction	equation	rate
Photoactivation	$\mathrm{S} \to \mathrm{S}^*$	$k_*[\mathrm{S}]$
Radical formation	$\mathrm{S}^* + \mathrm{S} \to \mathrm{S}^\circ + \mathrm{S}$	$k_\circ[\mathrm{S}][\mathrm{S}^*]$
Radical combination	$2\mathrm{S}^\circ \to \mathrm{D}$	$k_D[\mathrm{S}^\circ]^2$

Table 3.1: Elementary photochemical mechanisms that convert OOA to DHTA.

tailed molecules are photoactivated and turned into radicals because there are very few double-tails available as reactants for further radical polymerization. This means 2-OOA polymerization does not go further than double tails to any significant extent within the experimental timescale, so I will neglect the triple-tail formation steps. The single-tails are then combined into double-tails in the termination step.

These steps, summarized in Table 3.1, can be written as the following rate equations:

$$\begin{cases} \frac{d}{dt}[\mathrm{S}] & = -k_*[\mathrm{S}] \\ \frac{d}{dt}[\mathrm{S}^*] & = k_*[\mathrm{S}] - k_\circ[\mathrm{S}][\mathrm{S}^*] \\ \frac{d}{dt}[\mathrm{S}^\circ] & = k_\circ[\mathrm{S}][\mathrm{S}^*] - 2k_D[\mathrm{S}^\circ]^2 \\ \frac{d}{dt}[\mathrm{D}] & = k_D[\mathrm{S}^\circ]^2 \end{cases}$$

With the mass conservation constraint $[\mathrm{S}]+[\mathrm{S}^*]+[\mathrm{S}^\circ]+2[\mathrm{D}] = \mathrm{S}_{\mathrm{tot}}$ and initial conditions $[\mathrm{S}^*](0) = [\mathrm{S}^\circ](0) = [\mathrm{D}](0) = 0$ and $[\mathrm{S}^*](0) = \mathrm{S}_{\mathrm{tot}}$.

When both the photoactivation and radical formation steps are extremely fast, this process may be reduced to simple second-order mean-field kinetics $2\mathrm{S} \to \mathrm{D}$. The rate equations may be easily integrated to give:

$$[\mathrm{S}](t) = \left(\frac{1}{[\mathrm{S}](0)} + 2k_D t\right)^{-1}$$

$$[\mathrm{D}](t) = [\mathrm{D}](0) - \frac{1}{2}\left[\left(\frac{1}{[\mathrm{S}](0)} + 2k_D t\right)^{-1} - [\mathrm{S}](0)\right]$$

In this case, the appearance $[\mathrm{D}](t)$ is immediate without any incubation period. However, if the first two steps are not extremely fast, it takes the system some time to go through the intermediate states S^* and $\mathrm{S}^{\cdot}$ before the single-tailed radicals have accumulated to significant concentration for combination into double tails to start. This means there is an incubation period as a result of slow photoactivation and radical formation.

reaction	equation	rate
Adsorption	$\mathrm{X} \to \mathrm{X}_{(\mathrm{ads})}$	$k^{X}_{\mathrm{ads}}[\mathrm{X}]$
Desorption	$\mathrm{X}_{(\mathrm{ads})} \to \mathrm{X}$	$k^{X}_{\mathrm{des}}[\mathrm{X}_{(\mathrm{ads})}]$

Table 3.2: Elementary physical mechanisms at interfaces.

3.1.2 Interfacial chemistry

All molecules undergoing photoreaction and aggregation in this system are surfactants, which preferentially partition to the interfaces (in this experiment, solution/glass, and air/solution). For any species X, the flux between bulk solution and the interface is the sum of adsorption and desorption rates, as listed in Table 3.2.

This has two implications for the aggregation rates: firstly, according to Langmuir trough experiments for the air/solution interface, the photoreaction's product (double tails) is more surface active than its reactant (single tails). So the interfaces compete with the aggregates to serve as an additional sink for the double tails.

Secondly, the interfaces might be catalytic for the chemical conversion steps and/or the aggregate formation steps. The loss in translational entropy due to the combination of radicals or nucleation is much reduced at an interface (2D translation) compared to in bulk solution (3D translation).

3.1.3 Aggregation

On their own, the first two categories of reactions have relatively well-understood elementary steps and reaction intermediates. The kinetics of aggregation is more opaque because the aggregates' structures are varied and the associated energy landscape much more complicated than those of small molecule reaction dynamics[39]. If non-trivial structural conversions beyond the addition of a monomer to the aggregate occur during nucleation and growth of aggregates, we will need to have some knowledge of the aggregates' structure and ask about how to incorporate the rates of these structural conversion steps before writing down the elementary aggregation mechanisms for kinetic modeling. The structure of the aggregates is of interest, but in the current calculations, I assume no kinetically important structural conversion is occurring beyond simple nucleation and growth so that we can understand the implications of photochemistry and interfacial adsorption before examining the possibility of more complicated pathways involving structural conversion. The elementary steps are listed in Table 3.3.

Here $f(t, 1) = [\mathrm{D}](t)$ is the number of solvated double tails in the system (monomer in the aggregation reaction), and $f(t, j) = [\mathrm{D}_j](t)$ is the size (volume) distribution of aggregates at any instant t. If the size distribution is too noisy to be informative, the time evolution of average aggregate size is related to the solution to the master equation by the following equations:

reaction	equation	rate
Nucleation	$n\mathrm{D} \rightarrow \mathrm{D}_n$	$k_n[\mathrm{D}]^n$
Monomer addition	$\mathrm{D}_i + \mathrm{D} \rightarrow \mathrm{D}_{i+1}$	$k_p(1,i)[\mathrm{D}_i][\mathrm{D}]$
Monomer dissociation	$\mathrm{D}_i \rightarrow \mathrm{D}_{i-1} + \mathrm{D}$	$k_f(i-1,1)[\mathrm{D}_i]$
Aggregate coalescence	$\mathrm{D}_i + \mathrm{D}_j \rightarrow \mathrm{D}_{i+j}$	$k_p(i,j)[\mathrm{D}_i][\mathrm{D}_j]$
Aggregate fragmentation	$\mathrm{D}_{i+j} \rightarrow \mathrm{D}_i + \mathrm{D}_j$	$k_f(i,j)[\mathrm{D}_i][\mathrm{D}_j]$

Table 3.3: Elementary mechanisms of aggregation.

$$P(t) = \sum_{j=n}^{\infty} f(t,j) \qquad \text{Number of aggregates}$$

$$M(t) = \sum_{j=n}^{\infty} j f(t,j) \qquad \text{Total aggregate mass}$$

$$\langle j \rangle (t) = M(t)/P(t) \qquad \text{Average aggregate size}$$

$$\langle R \rangle (t) \propto (\langle j \rangle (t))^{1/3} \qquad \text{Average aggregate radius}$$

3.1.4 Possible rate-determining steps

The kinetic curves have an initial incubation period, which is a particularly interesting characteristic suggestive of a slow step before runaway growth happens. They also have a final asymptotic growth period, which could be attributed to the depletion of material with aggregation propensity.

There are multiple steps involved in both the chemical conversions and the aggregation. When analyzing the kinetic curves that show how the average radius of clusters changes over time, one should keep in mind that it is not solely aggregation and its numerous possible nucleation and addition pathways that produce the interesting features. The chemical conversion alone or the balance between the input from chemical conversion and output to aggregation could also explain the appearance of these features.

In the next section, I discuss how photochemistry, interfacial reactions, and aggregation pathways may interact with one another to produce the two essential features: induction time and monodispersity. I will start with a system that grows from a fixed total concentration of monomers, calculate its aggregate populations and induction time. I will then add in the elements of nucleation and source input to observe the effects of these steps on the aggregate population.

3.2 Aggregate population from stepwise growth

The aggregation pathways outlined above form a vast reaction network. To understand its essential consequences of reaction rate constants on the population of different aggregate sizes, I will begin by considering stepwise monomer addition as the only mechanism of growth in the absence of source input, nucleation, or fragmentation. The absence of fragmentation steps implies that aggregation is fast enough that fragmentation does not take place to any significant extent. The absence of nucleation, which is a nonlinear term, allows us to solve the master equation easily. Lastly, ignoring the effect of the source is equivalent to postulating that photochemical reactions take place on a much faster timescale than that of aggregation. In the later sections, I will add in nucleation and the source input to isolate the effects of these potentially slow steps.

For a system that grows with simple stepwise addition from a fixed total concentration of aggregates, the total amount of aggregation-prone material is initially $f(t=0, j=1)$ and the aggregate concentration $f(t=0, j) = 0$ for all $j > 1$. The rate constant for combining aggregates of size i and size j is $k_p(i, j)$. The size distribution evolves according to

$$\frac{\partial f(t,j)}{\partial t} = \sum_{i=1}^{j-1} k_p(i, j-i) f(t,i) f(t, j-i) - \sum_{i=1}^{\infty} (1+\delta_{ij}) k_p(i,j) f(t,i) f(t,j)$$

This set of master equations assumes spatial homogeneity in a dilute system (so that three- or higher-body interactions can be ignored), and that the aggregate mass is the only determinant for the addition rate. We can further simplify the system by assuming that coalescence is negligible, in which case the only non-zero elements in the growth rate matrix $k_p(i, j-i)$ are those with $i = 1$:

$$\frac{\partial f(t,j)}{\partial t} = \begin{cases} -2k_p(1,1) f(t,1)^2 - \sum_{i=2}^{\infty} k_p(1,i) f(t,1) f(t,i) & j = 1 \\ k_p(1, j-1) f(t,1) f(t,j-1) - k_p(1,j) f(t,j) f(t,1) & j > 1 \end{cases}$$

From the microscopic concentrations $f(i,j)$, I define number density $N(t) = \sum_{j\geq 1} f(t,j)$:

$$\begin{aligned}
\frac{dN}{dt} &= \sum_{j=1}^{\infty} \frac{\partial f(t,j)}{\partial t} \\
&= -2k_p(1,1)f(t,1)^2 - \sum_{i=2}^{\infty} k_p(1,i)f(t,1)f(t,i) \\
&\quad + \sum_{j=2}^{\infty} \left[k_p(1,j-1)f(t,1)f(t,j-1) - k_p(1,j)f(t,j)f(t,1)\right] \\
&= -2k_p(1,1)f(t,1)^2 - f(t,1)\sum_{i=2}^{\infty} k_p(1,i)f(t,i) \\
&\quad + \sum_{j=2}^{\infty} k_p(1,j-1)f(t,1)f(t,j-1) - \sum_{j=2}^{\infty} k_p(1,j)f(t,j)f(t,1) \\
&= -k_p(1,1)f(t,1)^2 - f(t,1)\sum_{i=2}^{\infty} k_p(1,i)f(t,i)
\end{aligned}$$

where the cancellation of two terms within the j-sum implies that the chain of stepwise aggregation does not change the total number of entities in the system beyond the loss of a monomer. When aggregates are large and their diffusion slows, we may further assume that the monomer addition rate constants $k_p(1,i)$ are independent of i, such that $k_p(1,i) = k_p$:

$$\begin{aligned}
\frac{dN}{dt} &= -k_p f(t,1)^2 - k_p f(t,1)\sum_{i=2}^{\infty} f(t,i) \\
&= -k_p f(t,1)\sum_{i=1}^{\infty} f(t,i) \\
&= -k_p f(t,1)N(t)
\end{aligned}$$

where the monomer concentration obeys

$$\frac{\partial f(t,1)}{\partial t} = -k_p f(t,1)^2 - k_p f(t,1)N(t) = -k_p f(t,1)^2 + \frac{dN}{dt}\,.$$

The time of the reaction network is scaled by $1/mk_p$ and all densities can be scaled by the total density of aggregation-prone material m, which in this case is equal to initial monomer concentration $f(0,1)$. Scaling rate constants and concentrations this way allows for the master equation to be written in its dimensionless form, such that all systems with various m and k_p can be mapped to the same unique solution.

Solving this system of equations numerically yields Fig. 3.1, where we see that the aggregation rapidly comes to a halt as monomers are depleted. Since every growth step requires at least one monomer, the depletion of monomers happens at the same time as

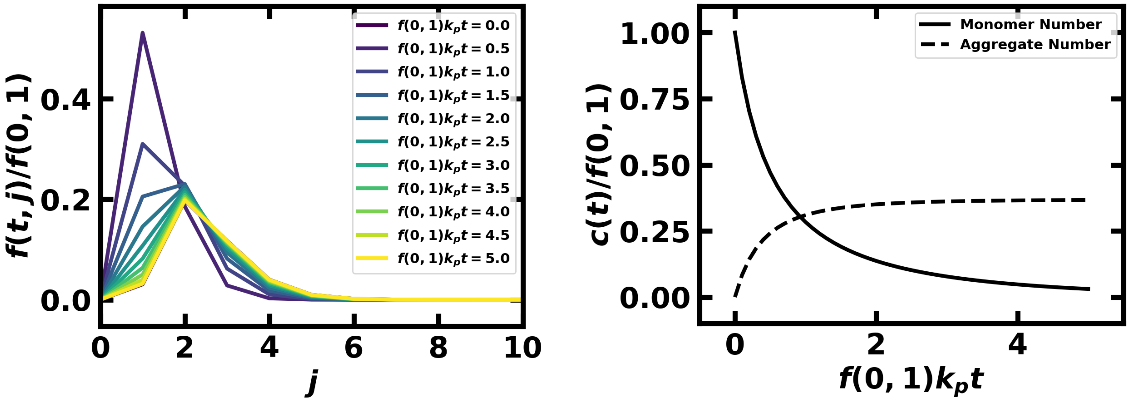

Figure 3.1: (Left) Time evolution of the size distribution from a mechanism that exhibits only stepwise growth through monomer incorporation. (Right) Number density of monomers and non-monomeric ($j > 1$) aggregates.

the growth stops. The largest aggregates in the system do not exceed $j = 10$. To produce larger aggregates, we must either supply the system with a source of monomers or build in a nucleation mechanism that increases the size of the smallest aggregates to be that of the critical nucleus. These two mechanisms will be explored in the following sections.

3.3 Stepwise growth from a step-function source

To model an aggregating solution with a source that continually introduces monomers, the master equations can be written with a source term $s(t)$:

$$\frac{\partial f(t,j)}{\partial t} = \begin{cases} s(t) - 2k_p f(t,1)^2 - \sum_{i=2}^{\infty} k_p f(t,1) f(t,i) & j = 1 \\ k_p f(t,1) f(t,j-1) - k_p f(t,1) f(t,j) & j > 1 \end{cases}$$

We are interested in aggregates made of more than one monomer, with number density $P(t) = N(t) - f(t,1)$:

$$\frac{dP}{dt} = \frac{\partial}{\partial t} \sum_{j>1} f(t,j) = k_p f(t,1) f(t,j-1)$$

and aggregate mass excluding monomer mass $M(t)$:

$$\begin{aligned}
\frac{dM}{dt} &= \frac{\partial}{\partial t}\sum_{j>1} jf(t,j) \\
&= \sum_{j>1} j\frac{\partial}{\partial t}f(t,j) \\
&= k_p\sum_{j>1} j\left[f(t,1)f(t,j-1) - f(t,1)f(t,j)\right] \\
&= k_p\sum_{j>1} jf(t,1)f(t,j-1) - k_p\sum_{j>1} jf(t,1)f(t,j) \\
&= k_p\sum_{j>0}(j+1)f(t,1)f(t,j) - k_p\sum_{j>1} jf(t,1)f(t,j) \\
&= k_pf(t,1)\sum_{j>0} f(t,j) + k_pf(t,1)^2 \\
&= 2k_pf(t,1)^2 + k_p\sum_{j=2}^{\infty} f(t,1)f(t,j)
\end{aligned}$$

which confirms the mass conservation,

$$\frac{dM}{dt} + \frac{\partial f(t,1)}{\partial t} = s(t)$$

In other words, the total amount of material in the system is unchanged by the aggregation pathways and only depends on the source input rate. The simplest source to be considered is the step-function source, from which k_s monomers are injected into the system per unit volume and per unit time, for a finite time m/k_s. The total amount of aggregation-prone material, m, is fixed so that we may compare the aggregate population yielded by the step-function source with the results from the previous section, where all material is already present in the system at the beginning of aggregation. The mechanism is summarized by:

$$s(t) = \begin{cases} k_s & 0 < t < m/k_s \\ 0 & \text{otherwise} \end{cases}$$

By varying k_s, we may then observe the effects of different rates of input on the aggregate population. In dimensionless form, the source input rate is k_s/k_pm^2. The numerical solutions for a series of decreasing input rates ($k_s/k_pm^2 = 1, 0.1, 0.01, 0.001$) are shown in Fig. 3.2 and Fig. 3.3. It is clear that we retrieve the constant initial concentration limit when the source inputs monomers rapidly into the system ($k_s/k_pm^2 = 1$). With reduced input rates, larger aggregates can now be observed because fewer small aggregates form during the early stages of the reaction; then during the later stages of the reaction, the added monomers are incorporated into the existing aggregates instead of forming new aggregates.

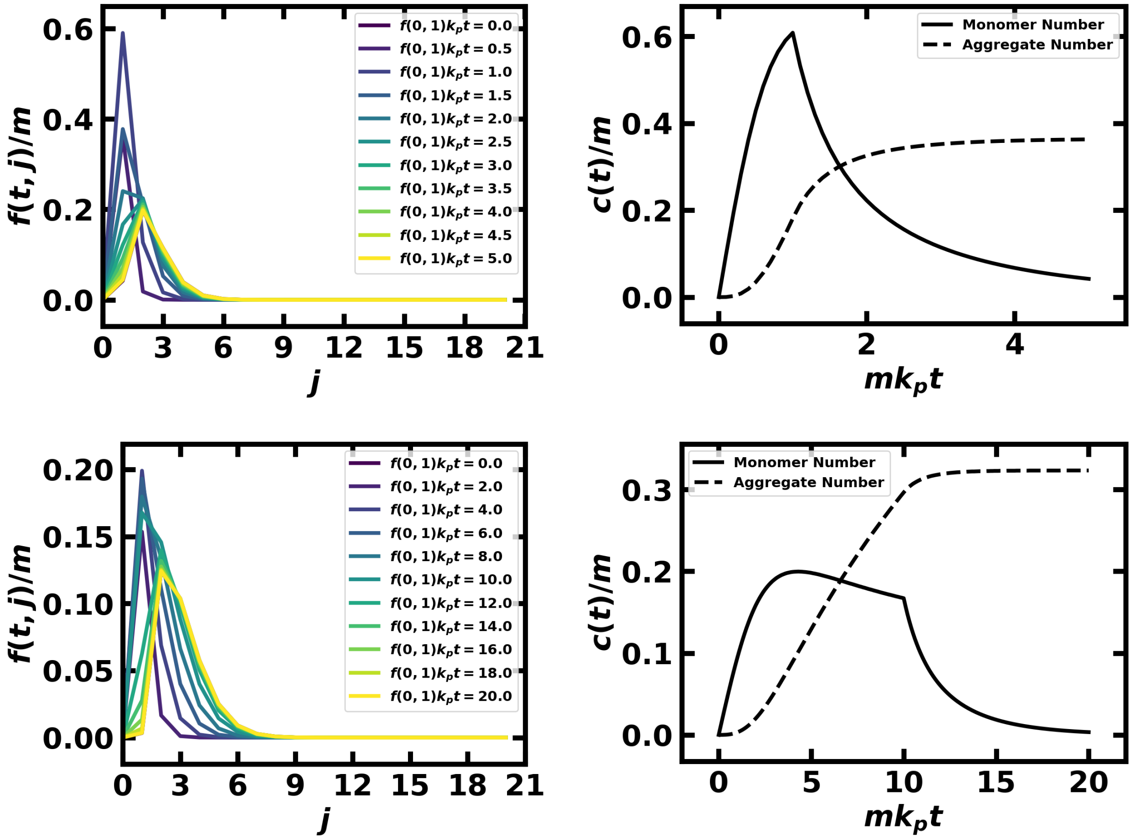

Figure 3.2: (Left) Time evolution of the size distribution from a step-function source. (Right) Number density of monomers and non-monomeric ($j > 1$) aggregates. From top to bottom, the source input rate is $k_s/k_pm^2 = 1, 0.1$ respectively.

The width of the size distribution increases with decreasing k_s/k_pm^2. While the large-j tail of the size distribution has a smooth decay to zero concentration, the small-j tail is abruptly cut off at the moment of monomer depletion, beyond which the system enters a steady state again because all growth steps require monomers. Monomer depletion is therefore the cause of finite aggregate size in this system. As expected, the aggregate number remains low while monomers are being introduced into the system. The induction time scales less than linearly with the monomer input rate.

3.4 Nucleated growth from fixed initial concentration

One of the more widely accepted model for the synthesis of monodisperse solid particles from homogeneous solution was proposed by Victor La Mer[40, 41, 42]. The key for

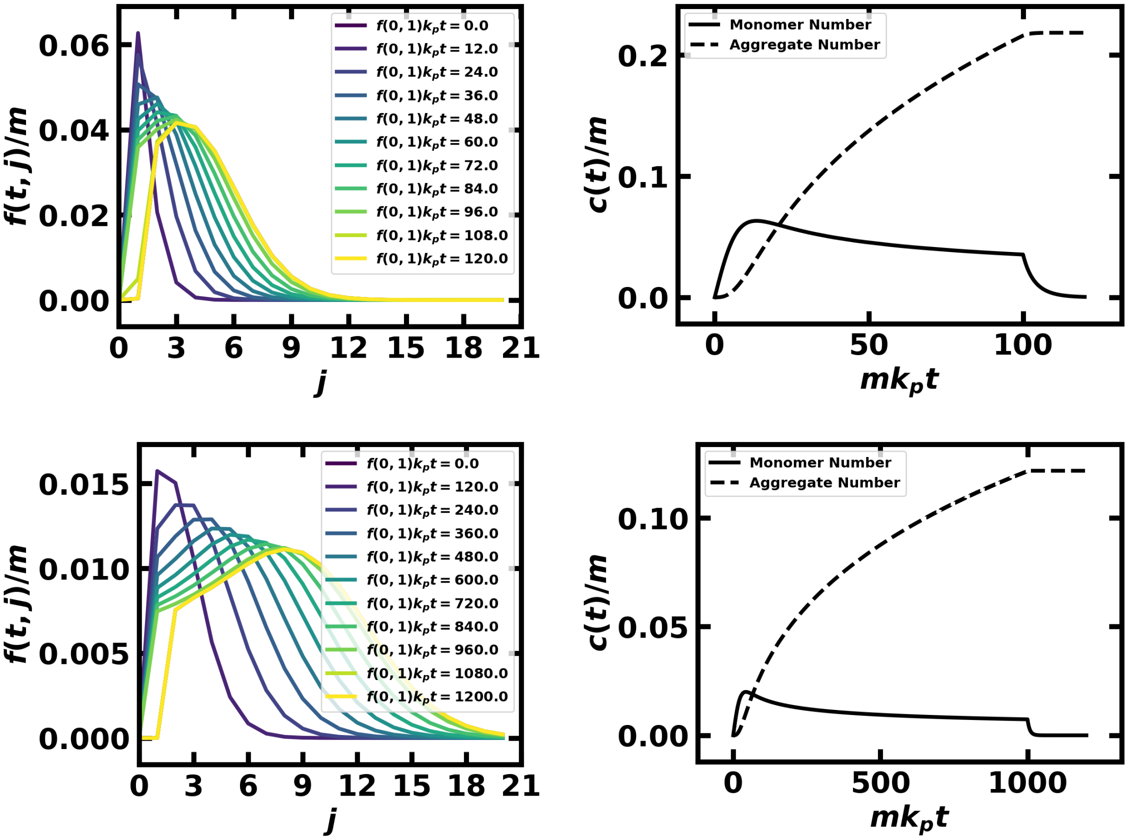

Figure 3.3: (Left) Time evolution of the size distribution from a step-function source. (Right) Number density of monomers and non-monomeric ($j > 1$) aggregates. From top to bottom, the source input rate is $k_s/k_pm^2 = 0.01, 0.001$ respectively.

monodispersity from the La Mer mechanism is that homogeneous nucleation is decoupled from growth by the sharp concentration dependence of the nucleation rate so that nucleation only occurs for a short period at the beginning of the reaction when the monomer concentration is higher. The critical nucleus size should be $j \gg 1$ for a sharp increase in nucleation rate with monomer concentration.

The mechanism in the previous section does not capture this nucleation mechanism as it assumes that an aggregate can grow stepwise from a single monomer without having to form a nucleus initially. In order to isolate the effect of nucleation, the source term is removed and all aggregation-prone material is set to be present as monomers in the system at $t = 0$, such that $f(0,1) = m$. Let n be the critical nucleus size and k_n be the nucleation rate. To add in an irreversible nucleation step, we insert a nonlinear term into the master equation. Each nucleation event consumes n monomers and produces one n-mer. There are no aggregates of size between monomers and critical nuclei. The master equation reads

$$\frac{\partial f(t,j)}{\partial t} = \begin{cases} -nk_n f(t,1)^n - \sum_{i=n}^{\infty} k_p f(t,1) f(t,i) & j = 1 \\ 0 & n > j > 1 \\ k_n f(t,1)^n - k_p f(t,1) f(t,n) & j = n \\ k_p f(t,1) f(t,j-1) - k_p f(t,1) f(t,j) & j > n \end{cases}$$

In the limit of $k_p = k_n$ and $n = 2$, the stepwise growth from fixed initial concentration is retrieved. The dimensionless nucleation rate constant is $k_n m^{n-2}/k_p$. The solutions to this set of master equations when $k_n m^{n-2}/k_p = 1$ and $n = 2, 10, 100$ is shown in Fig. 3.4. All aggregates are formed in the short initial period, as shown by the aggregate number concentration $P(t)$ quickly reaching its asymptotic value $P(\infty)$. Systems with larger critical nuclei exhibit shorter bursts of nucleation relative to the timescale of growth and produce a smaller total number of nuclei; mass conservation, therefore, requires that a large critical nucleus size n produces larger aggregates, and this trend is indeed observed in the numerical solutions. The shorter bursts of nucleation also produce narrower aggregate distributions, which can be quantified by smaller polydispersity indices (PDIs):

$$\mathrm{PDI} = \frac{\langle j^2 \rangle}{\langle j \rangle^2} = \frac{\langle \delta j^2 \rangle}{\langle j \rangle^2} + 1$$

As the initial nuclei form, both nucleation and growth from these nuclei deplete the monomer concentration as evidenced by the almost vertical drop of monomer concentration near $t = 0$. Nucleation is then quenched by this drop due to the high monomer concentration dependency of the homogeneous nucleation rate. Gradual growth then gradually increases the sizes of all nuclei, similar to the propagation of an impulse through j-space. Slight spreading of the aggregate size distribution's left tail is visible in the $n = 10$ and $n = 100$ systems; this can be understood by noting that the sizes of lower concentrations will grow more slowly due to their lower concentration.

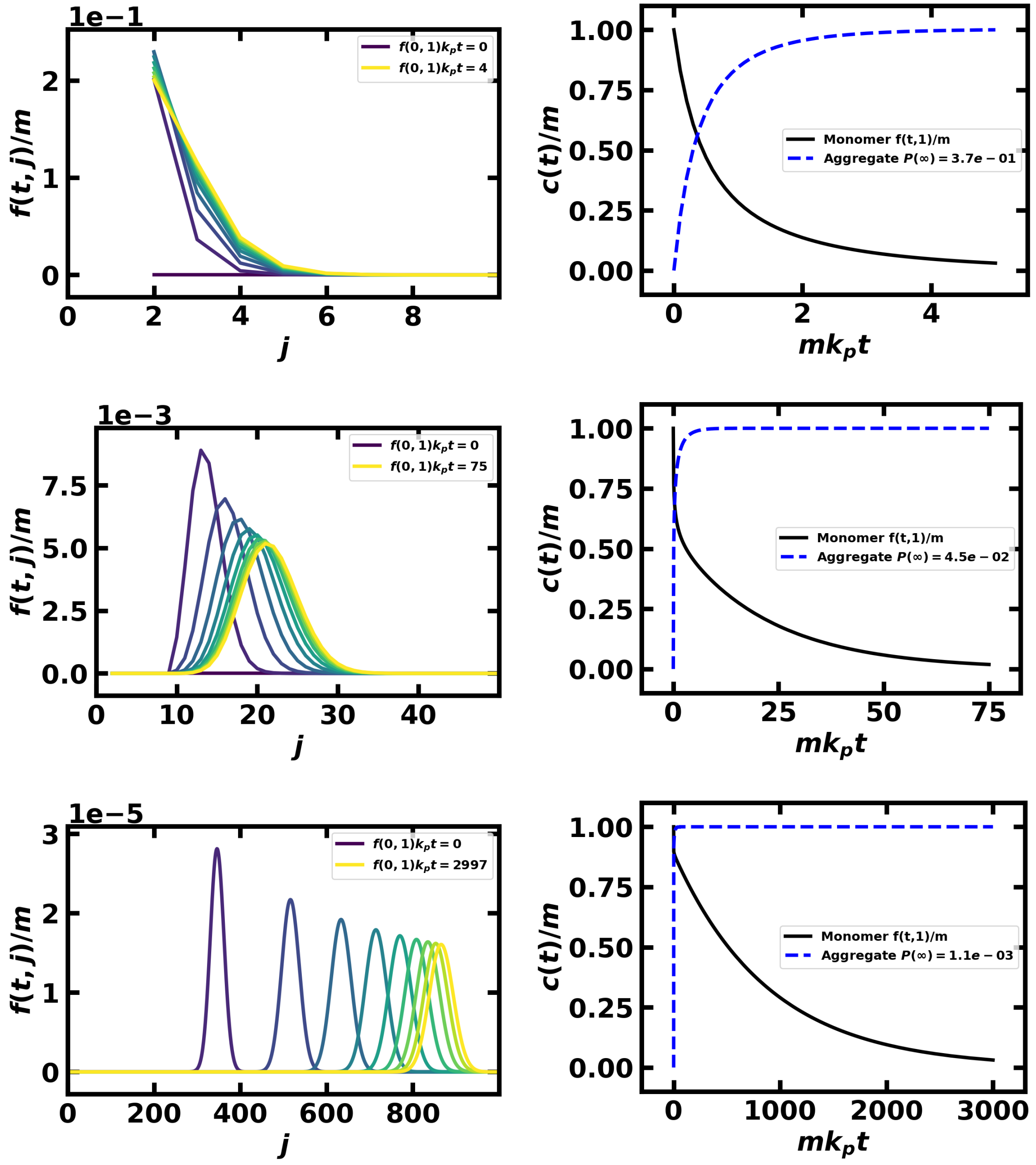

Figure 3.4: (Left) Time evolution of the size distribution of nucleated aggregation from a fixed initial concentration of monomers m. Monomers are not shown for clarity. (Right) Number density of monomers and non-monomeric ($j \geq n$) aggregates. All numerical solutions are generated for $k_n/k_p m^2 = 1$. From top to bottom, the critical nucleus size are $n = 2, 10, 100$ respectively.

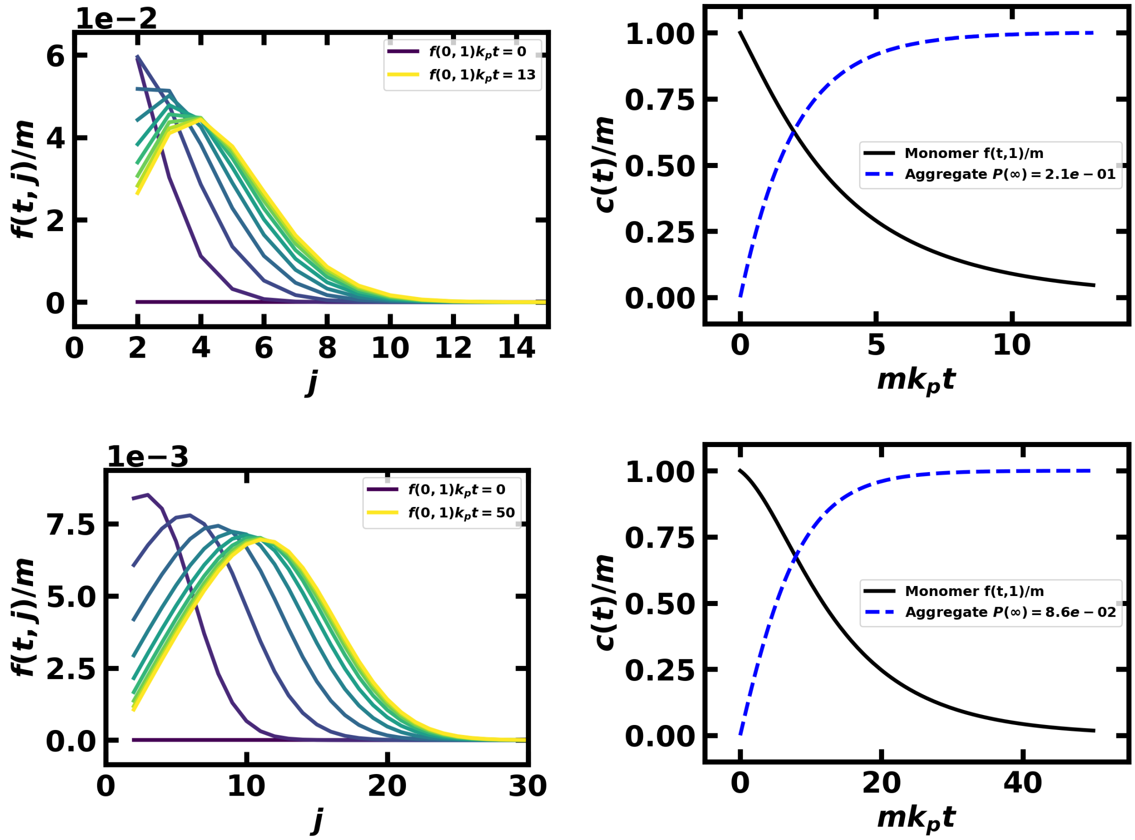

Figure 3.5: (Left) Time evolution of the size distribution of nucleated aggregation from a fixed initial concentration of monomers m. Monomers are not shown for clarity. (Right) Number density of monomers and non-monomeric ($j \geq n$) aggregates. All numerical solutions are generated for critical nucleus size $n = 2$. From top to bottom, the nucleation rate constants are $k_n m^{n-2}/k_p = 10^{-1}, 10^{-2}$ respectively.

As shown in Fig. 3.5 and Fig. 3.6, reducing the nucleation rate constant $k_n m^{n-2}/k_p$ while keeping the order of the nucleation reaction constant also increases aggregate size by reducing the total number of nuclei that form; however, a slower $k_n m^{n-2}/k_p$ does not concomitantly narrow the size distribution but instead widens it because the monomer depletion can only quench nucleation through a high reaction order n. The small number of nuclei are formed gradually and simultaneously with growth, resulting in a polydisperse size distribution. The slowdown of nucleation also introduces a concave region near $t = 0$ for the monomer concentration, though it is not as pronounced as the induction time observed in experiments; no induction time is seen in the aggregate number $P(t)$.

It is therefore expected that slow $k_n m^{n-2}/k_p$ coupled with high n will produce large, monodisperse aggregates without an induction period. As shown in Fig. 3.7, using a small $k_n = 0.0001$ and a large $n = 100$ results in large and highly monodisperse aggregates with

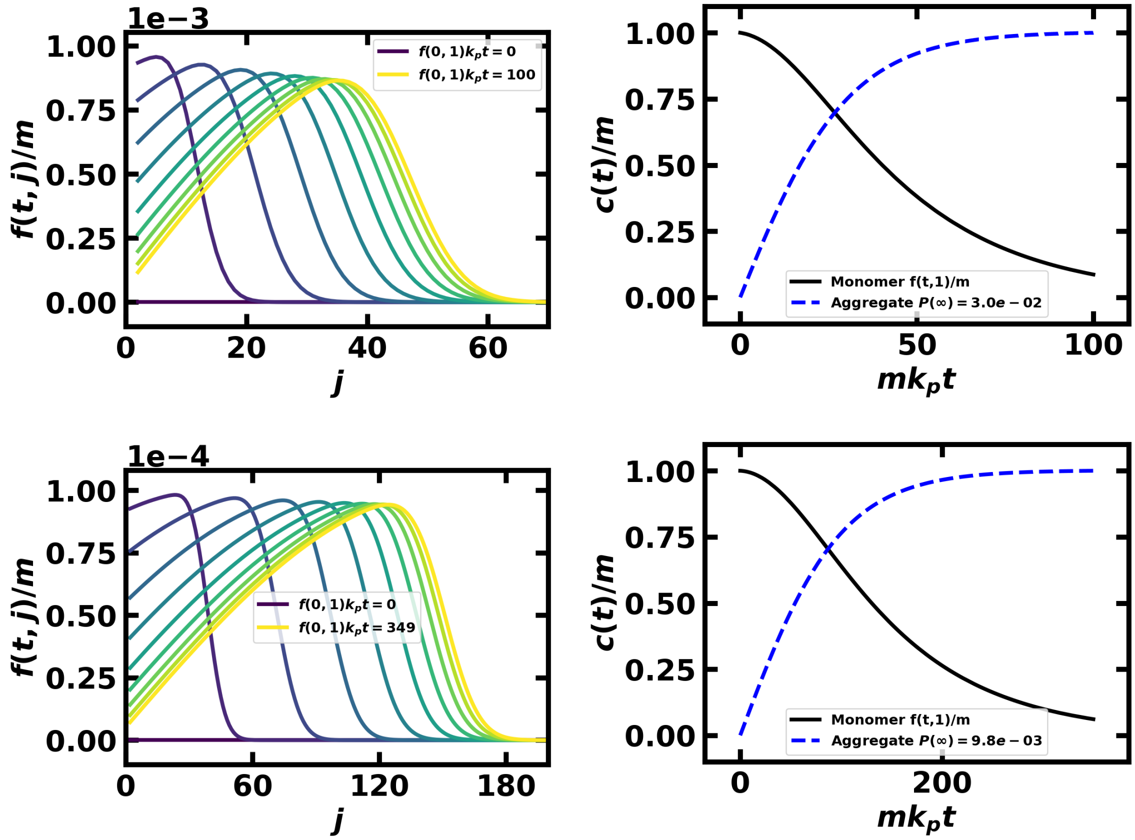

Figure 3.6: (Left) Time evolution of the size distribution of nucleated aggregation from a fixed initial concentration of monomers m. Monomers are not shown for clarity. (Right) Number density of monomers and non-monomeric ($j \geq n$) aggregates. All numerical solutions are generated for critical nucleus size $n = 2$. From top to bottom, the nucleation rate constants are $k_n m^{n-2}/k_p = 10^{-3}, 10^{-4}$ respectively.

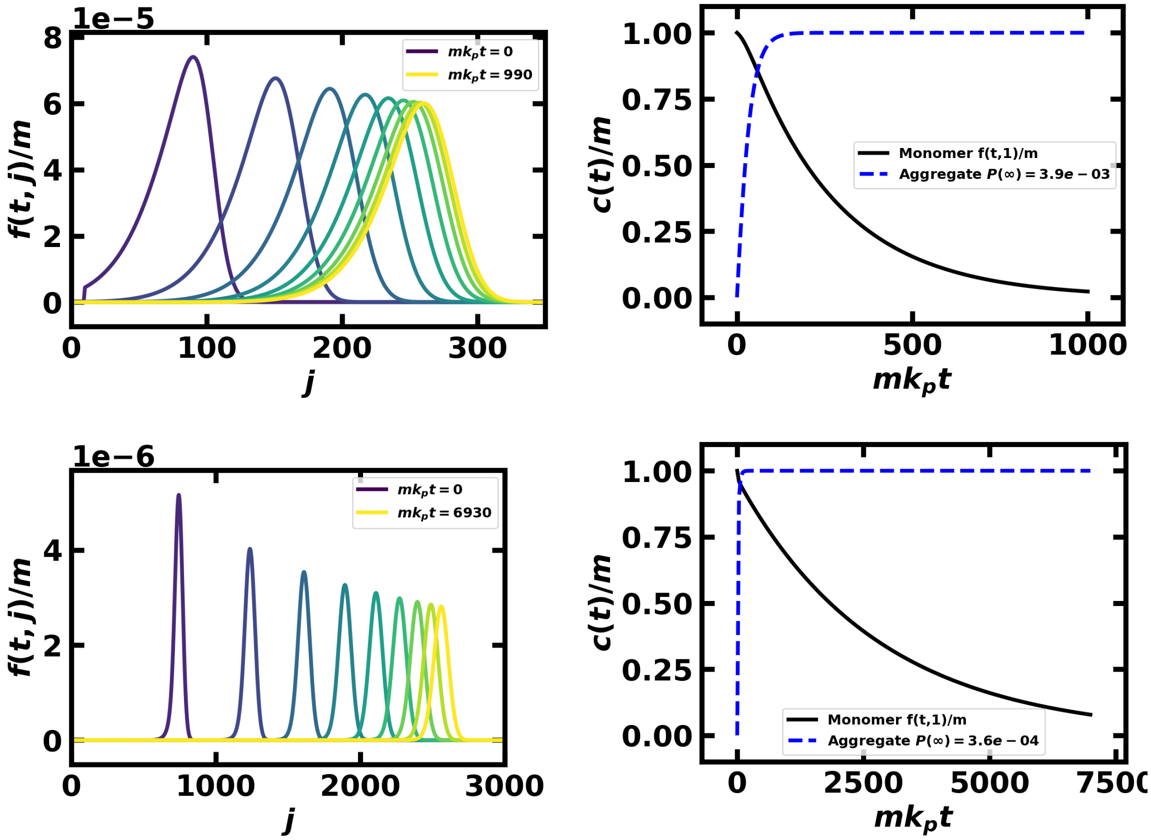

Figure 3.7: (Left) Time evolution of the size distribution of nucleated aggregation from a fixed initial concentration of monomers m. Monomers are not shown for clarity. (Right) Number density of monomers and non-monomeric ($j \geq n$) aggregates. All numerical solutions are generated for $k_n m^{n-2}/k_p = 10^{-4}$. From top to bottom, the critical nucleus sizes are $n = 10$ and $n = 100$ respectively.

an average size of over 2500 monomers and a PDI as low as 1.0004.

3.5 Nucleated growth from a step-function source

On the one hand, homogeneous nucleation with large critical nucleus sizes n can lead to burst nucleation, but it is unlikely to produce an induction time as seen in the case of OOA and DHTA assembly. On the other hand, un-nucleated growth from a step function source showed a limited induction time in the aggregate number density but was unable to yield a monodisperse aggregate population. We now consider nucleated growth from a step-function source to see if they may couple to yield monodisperse aggregates after an induction period:

$$\frac{\partial f(t,j)}{\partial t} = \begin{cases} s(t) - nk_n f(t,1)^n - \sum_{i\geq n} k_p(1,i) f(t,1) f(t,i) & j = 1 \\ 0 & n > j > 1 \\ k_n f(t,1)^n - k_p(1,n) f(t,1) f(t,n) & j = n \\ k_p(1,j-1) f(t,1) f(t,j-1) - k_p(1,j) f(t,1) f(t,j) & j > n \end{cases}$$

This set of master equations can be summarized into an equation for monomers and a set of equations for the aggregate j-mers:

$$\frac{\partial f(t,1)}{\partial t} = s(t) - nk_n f(t,1)^n - f(t,1) \sum_{i\geq n} k_p(1,i) f(t,i)$$

$$\frac{\partial f(t,j)}{\partial t} = k_p(1,j-1) f(t,1) f(t,j-1) - k_p(1,j) f(t,1) f(t,j) + \delta_{j,n} k_n f(t,1)^n$$

where the aggregates of size j larger than 1 and smaller than the critical nucleus size n are no longer part of the master equations because $k_p(1,j) = 0$ for $j \in (1, n-1)$ and $f(t,j) = 0$ for $j \in (1, n-1)$.

Solving this set of master equations yields Fig. 3.8. When $k_s/k_p m^2 = 1$ and $k_n m^{n-2}/k_p = 10^{-4}$ with $n = 100$, the nucleation timescale is much slower than the source input and is, therefore, the rate-limiting step that separates aggregation pathways from the source input mechanism. In this fast input limit, we retrieve the solutions we found for fixed total concentration. Any decrease of $k_s/k_p m^2$ in this limit leads to the appearance of an induction time without interfering with the nucleation pathways, as shown in the second row where $k_s/k_p m^2 = 10^{-3}$ and $k_n m^{n-2}/k_p = 10^{-4}$ with $n = 100$. The aggregate size distribution is unperturbed in either case, with a PDI of 1.0004.

Further lowering the monomer input rate to be comparable or even smaller than the nucleation rate has the effect of shortening the nucleation time window because monomer concentration is kept from reaching m by rapid growth as soon as a small number of nuclei form. This increases the induction time, decreases the total number of nuclei, and increases the size of the aggregates formed, as shown in Fig. 3.9 for $k_s/k_p m^2 = 10^{-4}$ and $k_n m^{n-2}/k_p = 10^{-4}$. Due to the high reaction order of the nucleation step (n), monomer concentration needs to reach almost m for nucleation events to take place; therefore the induction time is $\tau = m/k_s$. The PDI remains to be very monodisperse at 1.0009.

3.6 Nucleated growth from a decaying source

In the photoinitiated assembly reaction, monomers (DHTA) are not introduced at a constant rate but instead converted from a constant supply of precursors (OOA). Therefore we need to consider a second-order decaying source:

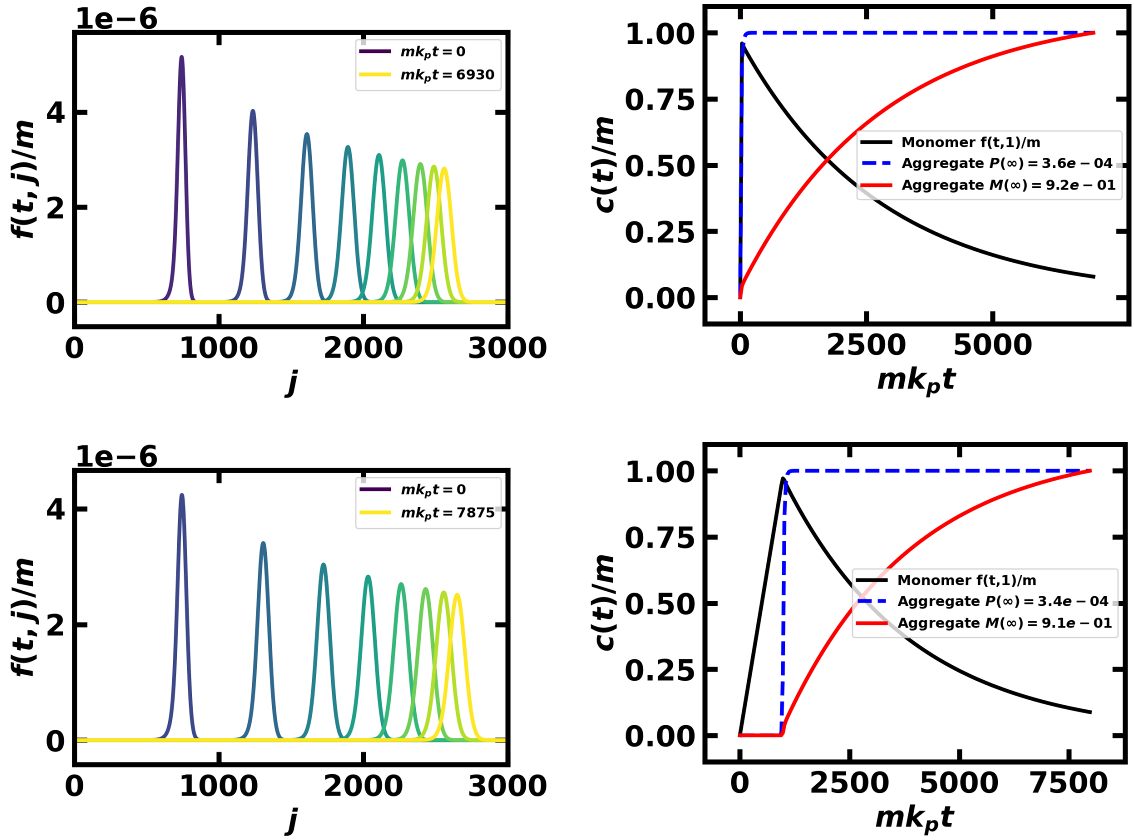

Figure 3.8: (Left) Time evolution of the size distribution of nucleated aggregation from a step-function source that inputs monomer at a constant rate of k_s for time m/k_s. Monomers are not shown for clarity. (Right) Number density of monomers and non-monomeric ($j \geq n$) aggregates. All numerical solutions are generated for $k_n m^{n-2}/k_p = 10^{-4}$. From top to bottom, the input rates are $k_s = 1$ and $k_s = 10^{-3}$ respectively.

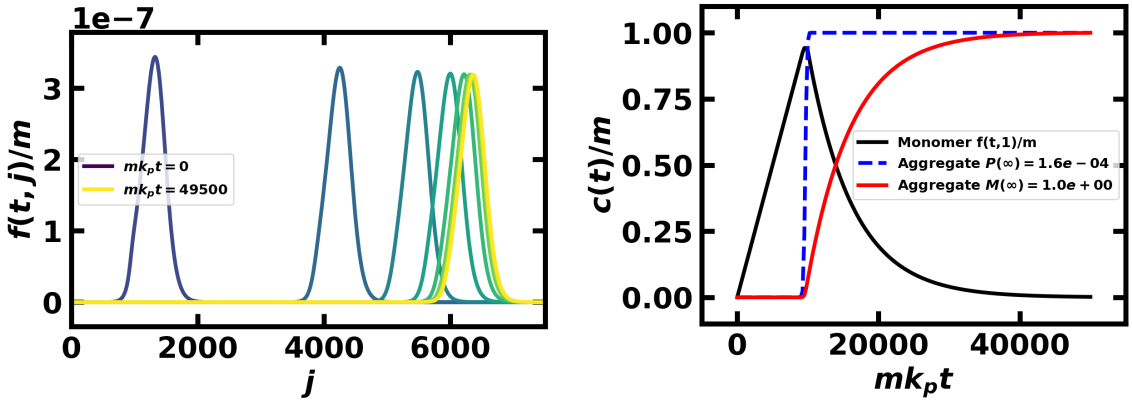

Figure 3.9: (Left) Time evolution of the size distribution of nucleated aggregation from a step-function source that inputs monomer at a constant rate of k_s for time m/k_s. Monomers are not shown for clarity. (Right) Number density of monomers and non-monomeric ($j \geq n$) aggregates. All numerical solutions are generated for $k_n m^{n-2}/k_p = 10^{-4}$ and $k_s/k_p m^2 = 10^{-4}$.

$$s(t) = k_d \left[\frac{1}{2m} + 2k_d t\right]^{-2}$$

where k_d is the rate constant of OOA dimerization into DHTA. This source will again input a total of m concentration units of monomers. The dimensionless version is

$$\frac{s(t)}{k_p m^2} = \frac{k_d}{k_p}\left(\frac{1}{2} + 2mk_d t\right)^{-2} .$$

Integrating gives $\int_0^\infty s(t)\,dt = m$, the same amount of monomers as used in previous mechanisms.

The decaying monomer input rate can compete with either time scales present in the system: the nucleation time scale $1/k_n m^{n-1}$ and the growth time scale $1/k_p m$. When $k_d/k_p \sim \mathcal{O}(1)$ or greater, monomers input into the system completes instantly. As shown in Fig. 3.10, the aggregate population dynamics are identical to the case when monomers were present at a concentration m at the beginning of the reaction.

Also shown in Fig. 3.10 is the system's behavior when k_d is reduced below the growth rate k_p. An induction period appears as a result of the delay before sufficient monomers are present to feed a significant number of nucleation events. Immediately after the induction period, the short burst of nucleation creates a small number of nuclei that rapidly grow with minimal spreading of their size distribution. The decrease in spreading was also observed in previous sections as the constant input rate k_s was lowered. By comparing the step function source and the decaying source, the two sets of solutions appear similar, suggesting that the exact functional form of the source input rate is unimportant for the induction time. To a first approximation, the induction time is proportional to the time scale of the input.

In conclusion, various kinetic mechanisms have been explored to determine the possible elements of the photoinitiated assembly reaction network:

1. A large nucleation rate order n causes burst nucleation that occurs rapidly at high monomer concentration but immediately ceases as monomer concentration decreases due to growth from the nuclei that just formed. A large n also causes the formation of large, monodisperse aggregates because fewer nuclei form during the short burst of nucleation as opposed to a longer nucleation window.

2. A slow nucleation rate constant k_n causes fewer nuclei to form, but a slow k_n on its own without a large nucleation order n cannot support burst nucleation and therefore cannot cause large, monodisperse aggregates to form. This is because the nucleation rate with a small order n is not sensitive enough to monomer concentration to be quenched. Instead, it continuously generates nuclei and only gradually declines as the monomer concentration declines.

3. A slow source input of monomers (k_s or k_d) causes an induction time. It also further shortens the nucleation window because monomer concentration is kept lower for longer than having all monomers present in the solution at the beginning of the reaction.

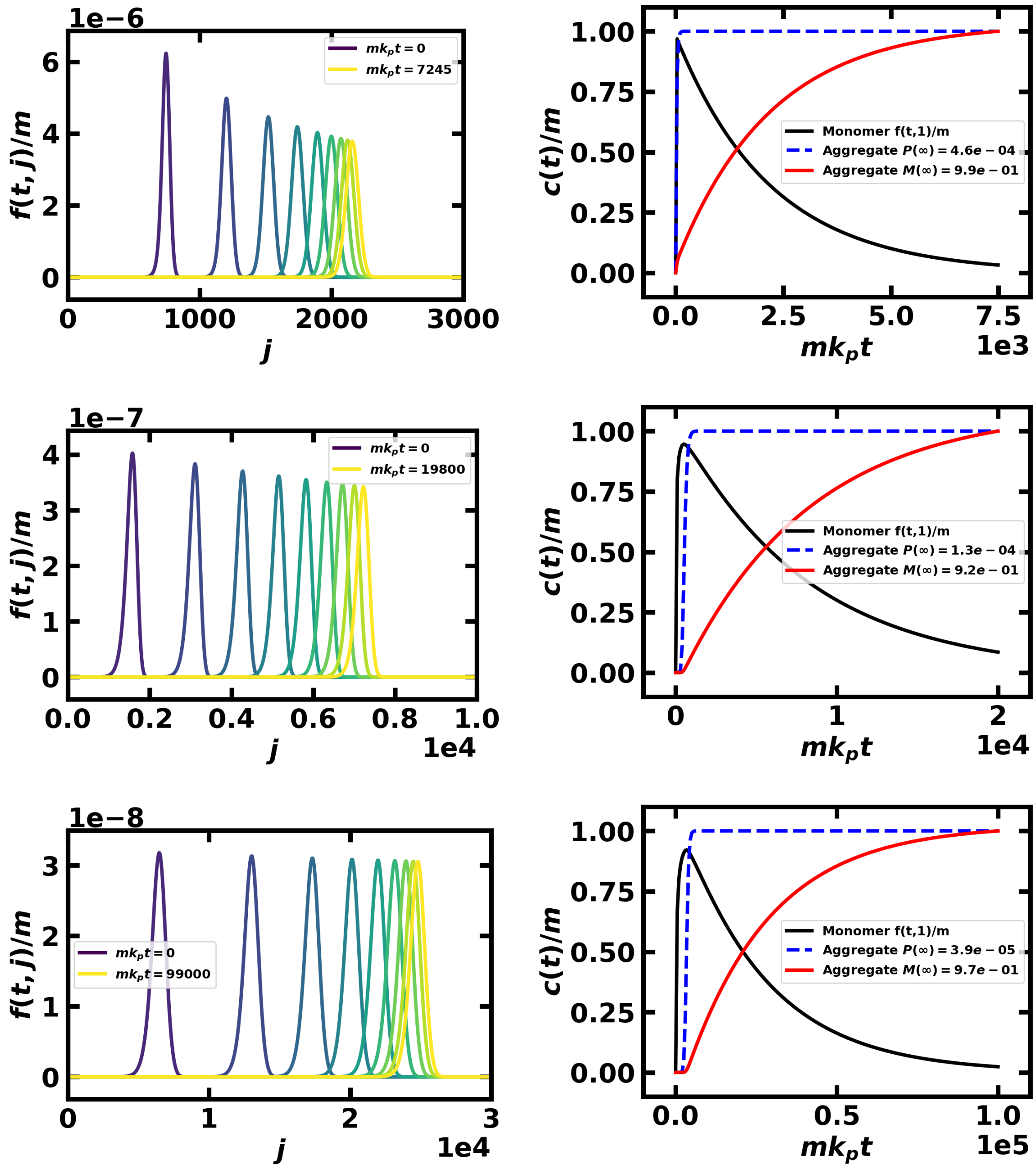

Figure 3.10: (Left) Time evolution of the size distribution of nucleated aggregation from a step-function source that inputs monomer at a constant rate of k_s for time m/k_s. Monomers are not shown for clarity. (Right) Number density of monomers and non-monomeric ($j \geq n$) aggregates. The numerical solutions are generated for $k_n m^{n-2}/k_p = 10^{-4}$, $n = 100$, and $k_d/k_p = 1$ (top) or $k_d/k_p = 10^{-2}$ (middle) or $k_d/k_p = 10^{-3}$ (bottom).

Chapter 4

Charge-frustrated Ising model

4.1 Motivation

The Ising model[43], or equivalently the lattice gas model, captures the universal characteristics of macroscopic phase separation between hydrophobic and hydrophilic liquids. Each lattice cell has a binary occupation number $n_i = 0, 1$ that represents the presence of either a hydrophobic group of atoms (or molecules) or a hydrophilic group of atoms (or molecules). The cells have nearest-neighbor interactions, which are favorable when the two neighboring cells have the same occupation number (both hydrophobic or both hydrophilic), and unfavorable when the two neighboring cells have different occupation numbers (one hydrophobic, and the other hydrophilic). Under the appropriate thermodynamic conditions, these interactions are strong enough to cause the system to phase separate.

Inside amphiphilic molecules, however, hydrophobic and hydrophilic groups are forced to be in proximity by covalent bonds, thus preventing the groups to separate from each other macroscopically. So to represent the amphiphilic connection that *frustrates* the phase separation within the framework of the Ising model, some additional correlations between the hydrophobic and hydrophilic lattice cells need to be built into the model.

There are many ways to add cell-cell interactions that would cause such correlations. For example, a simple construction is to require a hydrophobic cell to always be the immediate neighbor of a hydrophilic cell. Wu et al. built in the necessary correlations by adding a Coulomb-like effective potential term to the existing Ising spin pair interactions.[26] The lattice cells with different occupation numbers were given opposite effective charges, so the favorable opposite-charge interactions counteract the unfavorable different-occupation interactions.

This more sophisticated treatment mapped directly from the molecular structure of the amphiphile to the long-range correlations in the mesoscopic and macroscopic configurations of amphiphilic assemblies. It was then possible to extend the existing methods of studying the Ising model to study the amphiphilic assemblies, such as Monte Carlo simulations and field-theoretic calculations[28, 44, 27]. The resulting model is called the *charge-frustrated Ising model*.

To apply the charge-frustrated Ising model to realistic systems like the photoinitiated 2-OOA assembly, this chapter starts by tuning the charge-free Ising model to represent both the air-liquid interface and the oil-water interface in one system with the correct surface roughness. I then examine the consequences of imposing such realism on the charge-frustrated Ising model, including the representation of mixtures of more than one amphiphilic species as well as fluctuation-free estimates of the relative stabilities of various assembled structures. I also expand the charge-frustrated Ising model to include a treatment of non-zero spontaneous curvature. In the Appendix, I include a description of the simulation methods that speed up computations of the charge-frustrated Ising model.

4.2 Putting oil, water and air on a lattice

4.2.1 Surface tension relationships

To represent water at liquid-vapor coexistence, a one-component lattice gas should have an energy of cohesion $1.25 \leq \epsilon/k_BT \leq 1.4$ and a lattice spacing $1.7\text{Å} \leq \delta \leq 1.9\text{Å}$.[45] The reasoning for restricting the lattice energy and the lattice spacing to such a narrow range is as follows: On one hand, if the cohesion is too strong, then the liquid-vapor interface will not reproduce the correct roughness on the capillary wave spectrum; on the other hand, if it is too weak, then the system becomes supercritical and its coexistence cannot be maintained. Furthermore, the surface tension of water constrains the ratio between ϵ and the lattice spacing δ by

$$\gamma = \frac{\beta\epsilon^2}{\pi\delta^2}\,.$$

Knowing water's surface tension, the range of lattice spacing corresponding to the range of cohesion above is then $1.7\text{Å} \leq \delta \leq 1.9\text{Å}$.

Similarly, using the charge-frustrated Ising model to study the structure and dynamics of amphiphiles at soft interfaces requires the interface roughness to be represented accurately through a careful choice of energies of cohesion and the lattice spacing. The charge-frustrated Ising model that represents amphiphiles at air-water interfaces, however, is a three-component lattice gas made of water, water-like head groups and oil-like tail groups. The parameters in this model are: a universal lattice spacing δ, and six cohesion energies $\{\epsilon_{\text{tt}}, \epsilon_{\text{ww}}, \epsilon_{\text{hh}}, \epsilon_{\text{ht}}, \epsilon_{\text{wt}}, \epsilon_{\text{hw}}\}$ that can be collapsed down to three, $\{\epsilon_{\text{ow}}, \epsilon_{\text{ww}}, \epsilon_{\text{oo}}\}$, if we assume heads are equivalent to water and tails are equivalent to oil.

To determine these four parameters, the constraints of criticality and interfacial roughness are useful. Additionally, this system should have additional constraints implied by the criticality and roughness of all other interfaces:

1. The air-water or air-head interface have the correct surface tensions.

2. The water-tail or head-tail interfaces have the correct surface tensions.
3. The air-tail interface has the correct surface tension.

So we end up with three equality constraints that determine three energy parameters. The criticality constraints are inequalities and will be discussed later to show how they determine the universal lengthscale δ.

The oil-air and water-air interfaces follow the formulae,

$$\gamma_{\mathrm{aw}} = \frac{\beta\epsilon_{\mathrm{ww}}^2}{\pi\delta^2} \quad \text{and} \quad \gamma_{\mathrm{ao}} = \frac{\beta\epsilon_{\mathrm{oo}}^2}{\pi\delta^2} \, .$$

The oil-water surface tension is more complicated because both oil and water cells have cohesion within themselves, unlike air cells which do not have cohesion. To apply the oil-water surface tension constraint, consider an incompressible, binary mixture of binary oil and water. The lattice Hamiltonian is a function of occupation variables $n_i = 0$ for oil sites and $n_i = 1$ for water sites:

$$\begin{aligned} \mathcal{H}\left[\{n_i\}\right] &= \epsilon_{\mathrm{oo}} \sum_{i,j}{}'(1-n_i)(1-n_j) + \epsilon_{\mathrm{ow}} \sum_{i,j}{}'\left[(1-n_i)n_j + n_i(1-n_j)\right] + \epsilon_{\mathrm{ww}} \sum_{i,j}{}' n_i n_j \\ &= \sum_{i,j}{}'\left(\epsilon_{\mathrm{oo}} + \epsilon_{\mathrm{ww}} - 2\epsilon_{\mathrm{ow}}\right) n_i n_j + \ldots \end{aligned}$$

where the chemical potential terms (functions of number of oils and waters) and constants are omitted. From the last line, it is clear that the relative difference of cohesions $(\epsilon_{\mathrm{oo}} + \epsilon_{\mathrm{ww}} - 2\epsilon_{\mathrm{ow}})$ determines the phase behavior of this "lattice gas" system. It also shows that the oil-water surface tension in the lattice representation is related to these cohesion parameters by

$$\gamma_{\mathrm{ow}} = \frac{\beta}{\pi\delta^2}\left(\epsilon_{\mathrm{oo}} + \epsilon_{\mathrm{ww}} - 2\epsilon_{\mathrm{ow}}\right)^2 \, .$$

4.2.2 Choice of lattice cell size(s)

Up to this point, if we are given the three surface tensions and a universal lengthscale δ, we may calculate the energy parameters $\{\epsilon_{\mathrm{ow}}, \epsilon_{\mathrm{ww}}, \epsilon_{\mathrm{oo}}\}$. But δ remains undetermined because we have not yet applied the roughness and criticality constraints - that is if our choice of δ is too small, one or more of $\{\epsilon_{\mathrm{ow}}, \epsilon_{\mathrm{ww}}, \epsilon_{\mathrm{oo}}\}$ will be too weak to hold the liquids at coexistence with their vapors; and if our choice of δ is too large, the interfaces will no longer be rough, in which case the relationship between surface tension and energy and length scales we used will no longer hold.

Fig.4.1 and Fig.4.2 illustrate how a universal lattice size δ is determined:

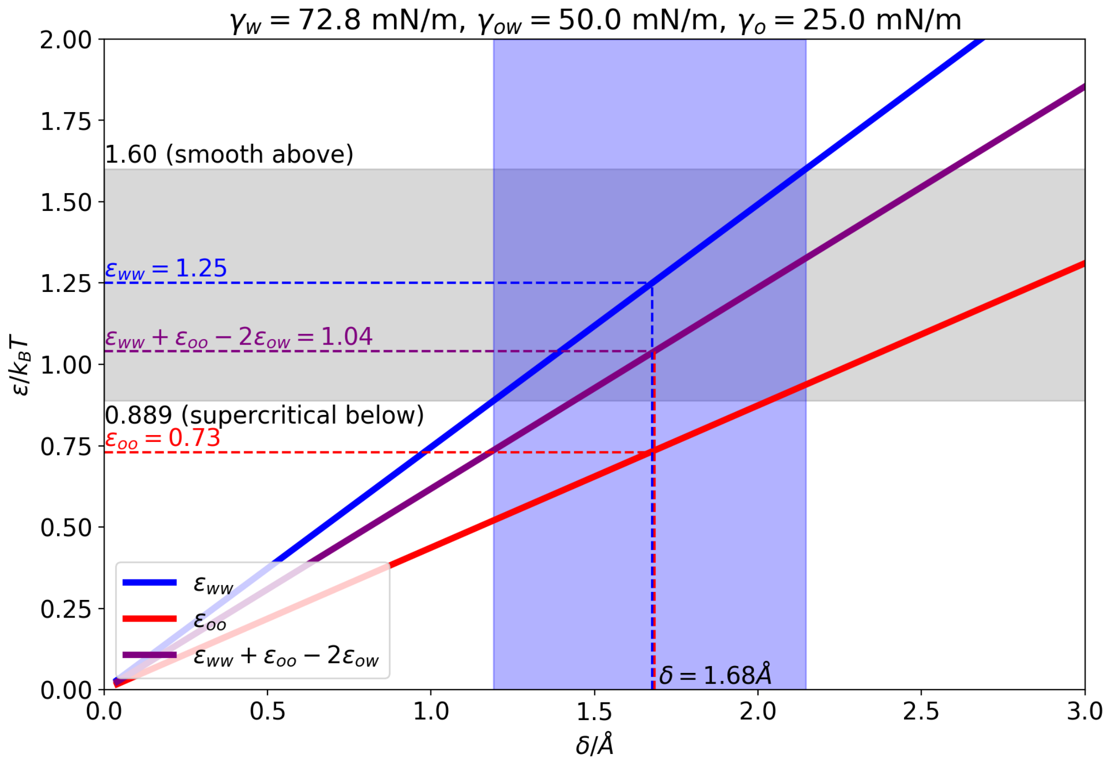

Figure 4.1: Procedure for determining different lattice cell sizes for oil and water with given surface tension $\gamma_{\text{water}} = 72.8$ mN/m and typical values $\gamma_{\text{alkane,water}} = 50$ mN/m and $\gamma_{\text{alkane,air}} = 25$ mN/m at room temperature, using one universal lattice size for all components of the system.

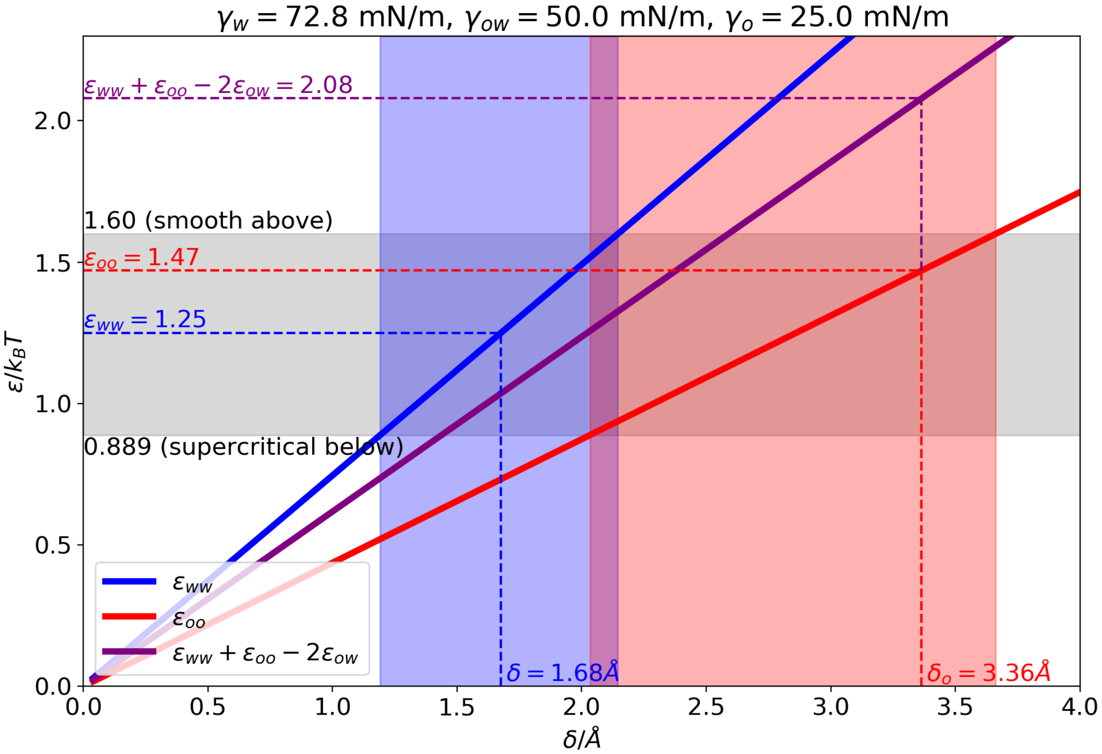

Figure 4.2: Relaxing the constraint of a universal lattice cell size, now choose lattice cell size for oil to be twice that of water.

Interface	Phase coexistence	Interfacial roughness	Surface tension matching
air/water	✓	✓	✓
oil/water	✓	✓	✓
air/oil	× (no interface)	× (no interface)	× (no interface)

Table 4.1: The consequences of using one lattice spacing, 1.68 Å, for all pair interactions.

- The blue, purple, and red solid lines plot the relationship between ϵ and δ for fixed surface tensions, in this case $\gamma_{\text{water}} = 72.8$ mN/m and typical values $\gamma_{\text{alkane,water}} = 50$ mN/m and $\gamma_{\text{alkane,air}} = 25$ mN/m at room temperature. This may seem like a very particular choice of parameters for surface tensions, but the resulting physical phenomena are expected to be generalizable to other surface tension including some that would prohibit a convenient lattice model representation.
- The grey zone spans across the range of cohesion that gives a rough interface without the system being supercritical, $\epsilon/T \in [-0.89, -1.60]$. More strictly, for the interface to be rough at all scales without the formation of large domains of one phase in the other, we prefer the cohesion to be as close to the middle of the interval as possible.
- The blue zone corresponds to the range of lattice sizes δ that will allow the air-water surface tension relationship to be maintained, while at the same time obeying the criticality and roughness constraints.

A single δ value for both oil and water will not let all three ϵ values fall in the range $1.25 \leq \epsilon/k_BT \leq 1.4$, where both the roughness and criticality constraints are satisfied with just one δ without getting too close to the boundaries. Specifically, the shortfalls and advantages of this combination of choices are summarised in table 4.1.

To continue, we must have separate lattice cell sizes, δ_o and δ_w, for oil and water, as illustrated in the bottom graph of Fig. 4.1. The constraints can now be easily satisfied with oils cells being twice the size of water cells:

$$\delta_{\text{o}} = 2\delta_{\text{w}} = 3.36\text{Å}$$

and energies

$$\epsilon_{\text{oo}}^{\text{block}}/T = -1.47 \quad \epsilon_{\text{ww}}^{\text{site}}/T = -1.25\,.$$

In addition to the features previously described for the top graph, we also have:

- The red zone corresponds to the range of lattice sizes δ that will allow the air-water surface tension relationship to be maintained, while at the same time obeying the criticality and roughness constraints.

Table 4.2 puts this into the same tabular form as before for comparison. We can see that in the second graph, we only sacrifice the interfacial roughness of oil and wate, which is preferable to the mixing oil and air in the first graph.

Interface	Phase coexistence	Interfacial roughness	Surface tension matching
air/water	✓	✓	✓
oil/water	✓	×	✓
air/oil	✓	✓	✓

Table 4.2: The consequences of using two different lattice spacings for pair interactions.

4.2.3 Lattice energies

So far we have established the lattice spacing and energy for two systems. The first system is liquid water in coexistence with its own vapor, with a small lattice spacing $\delta_{\mathrm{w}} = 1.68$Å and cohesion $\epsilon_{\mathrm{ww}}/T = -1.25$. The second system is liquid oil in coexistence with oil vapor, with lattice spacing twice as large, $\delta_{\mathrm{o}} = 3.36$Å and a stronger cohesion between the bigger blocks of oil, $\epsilon_{\mathrm{oo}}^{\mathrm{block}}/T = -1.47$. Each oil block effectively occupies eight lattice sites ($2 \times 2 \times 2$) if placed in the water system, and so the interaction between the faces of two oil *sites* is only 1/4 of that between oil *blocks*, $\epsilon_{\mathrm{oo}}^{\mathrm{site}} = \epsilon_{\mathrm{oo}}^{\mathrm{block}}/4 = -0.37T$. In practice this energy need only be calculated between different oil blocks; no accounting for the interactions among oil sites in the same block is necessary because the block is never broken apart.

Next, we move onto the final free parameter: given a fixed surface tension between oil and water (for simplicity we may reasonably assume $\gamma_{\mathrm{ow}} = \gamma_{\mathrm{o}}$) and the previously determined lattice spacings and cohesions, what should the cohesion between oil and water be? For this purpose, it is natural to consider the surface tension of an interface in an incompressible system of oil and water as derived above,

$$\gamma_{\mathrm{ow}} = \frac{\beta}{\pi\delta^2}\left(\epsilon_{\mathrm{oo}} + \epsilon_{\mathrm{ww}} - 2\epsilon_{\mathrm{ow}}\right)^2$$

and the question then becomes which δ this lengthscale corresponds to and which ϵ, the site-site or block-block ones, should be used. In such an incompressible system where oil block movements are restricted to a coarser lattice than the water sites, the water sites are effectively constrained to a coarser lattice too, so the expression above should be interpreted as

$$\gamma_{\mathrm{ow}} = \frac{\beta}{\pi\delta_o^2}\left(\epsilon_{\mathrm{oo}}^{\mathrm{block}} + \epsilon_{\mathrm{ww}}^{\mathrm{block}} - 2\epsilon_{\mathrm{ow}}^{\mathrm{block}}\right)^2$$

where $\epsilon_{\mathrm{oo}}^{\mathrm{block}}/T = -1.47$ and $\epsilon_{\mathrm{ww}}^{\mathrm{block}}/T = -5.00$, and the $\epsilon_{\mathrm{ow}}^{\mathrm{block}}$ is then

$$\begin{aligned}
\epsilon_{\text{ow}}^{\text{block}} &= \frac{1}{2}\left(\epsilon_{\text{oo}}^{\text{block}} + \epsilon_{\text{ww}}^{\text{block}} + \delta_{\text{o}}\sqrt{\frac{\pi\gamma_{\text{ow}}}{\beta}}\right) \\
\epsilon_{\text{ow}}^{\text{site}} &= \frac{1}{4}\epsilon_{\text{ow}}^{\text{block}} \\
&= \frac{1}{2}\left(\epsilon_{\text{oo}}^{\text{site}} + \epsilon_{\text{ww}}^{\text{site}}\right) + \frac{1}{8}\delta_{\text{o}}\sqrt{\frac{\pi\gamma_{\text{ow}}}{\beta}} \\
&= \frac{1}{2}\left(-1.25T - 0.3675T\right) + \frac{1}{8} \times 2.08T \\
&= -0.55T\,.
\end{aligned}$$

4.3 Turning oil and water into surfactants

These parameters in the three-state lattice gas model (water, oil, and vacuum) should now be capable of faithfully representing the interfacial roughness. On top of the pure three-state lattice gas, amphiphilic molecules can then be represented using the electrostatic analogy - by imposing charges onto some of the water and oil molecules to turn them into head and tail groups. However, it is not clear whether we have gained anything on the length- or time-scale reachable by this model compared to other methods, because periodic box simulations of electrolyte solutions often require us to take extra care of calculating the energy of long-ranged electrostatic interactions, and the procedure for book-keeping (e.g. Ewald sum) is demanding on CPU time.

The speed of this simulation is therefore limited by the range of electrostatic interactions. In general, the longer the range of these electrostatic interactions, the longer the book-keeping process will take because we have to sum up energetic contributions from more and more periodic images of the system, which eventually requires a greater number of wavevectors to be included in the Ewald sum implementation.

In this section, I will estimate the range of effective electrostatic interactions in the lattice representations of these surfactant systems (the Debye length) using the common model surfactant molecule sodium dodecyl sulfate (SDS) near its critical micelle concentration (CMC) of 8.2 mM.

4.3.1 Debye length

The Debye length[46] of an electrolyte with concentration c_i and valence z_i for $i = 1, ..., n$ ionic species in a solvent with dielectric constant $\epsilon_0\epsilon_r$ is:

$$\lambda_D = \kappa^{-1} = \sqrt{\frac{\epsilon_r\epsilon_0 k_B T}{e^2 \sum_i \rho_i z_i^2}}$$

where the charges in gaussian units are $z_i e/\sqrt{\epsilon_0\epsilon_r}$. The size of effective charges are established by comparing the $\mathcal{O}(k^2)$ hamiltonian of the surfactant system and an analogous point charge system:

$$\begin{aligned}F_S =&\frac{1}{2V}\sum_{\mathbf{k}}\sum_{i,j=+,-}\frac{\hat{\rho}_i(\mathbf{k})\hat{\rho}_j(-\mathbf{k})}{n_i n_j}\cdot\frac{3k_BT}{\rho k^2\Delta^2}\gamma_{ij}\\ F_C =&\frac{1}{2V}\sum_{\mathbf{k}}\sum_{i,j=+,-}\frac{\hat{\rho}_i(\mathbf{k})\hat{\rho}_j(-\mathbf{k})}{n_i n_j}\cdot\frac{4\pi q_i q_j}{k^2}\end{aligned}$$

where $\gamma_{ij} = 1$ if $i = j$ and $\gamma_{ij} = -1$ if $i \neq j$. The second line is in cgs units. In SI units, the Coulomb energy would be

$$F_C = \frac{1}{2V}\sum_{\mathbf{k}}\sum_{i,j=+,-}\frac{\hat{\rho}_i(\mathbf{k})\hat{\rho}_j(-\mathbf{k})}{n_i n_j}\cdot\frac{q_i q_j}{\epsilon_r\epsilon_0 k^2}\,.$$

By comparing the two free energies, if we constrain that the two sums are equal on a term-by-term basis, we see the effective valence of charges ($z_i = q_i/e$) is related to the molecular density of the surfactant $\rho = \frac{\rho_+}{n_+} = \frac{\rho_-}{n_-}$ by

$$z_+ = -z_- = \sqrt{\frac{3k_BT\epsilon_r\epsilon_0}{\rho\Delta^2 e^2}}$$

where Δ is the average separation of head and tail groups. I estimate the average separation to be $\Delta = n\delta$ where $n \approxeq 3$, and assume for SDS the stoichiometry is $n_+ = n_- = s$. The ionic strength is then

$$\sum_i \rho_i z_i^2 = \left(\rho_+ z_+^2 + \rho_- z_-^2\right) = 2s\rho z^2 = 2s\rho \times \frac{3k_BT\epsilon_r\epsilon_0}{\rho\Delta^2 e^2} = \frac{6sk_BT\epsilon_r\epsilon_0}{n^2\delta^2 e^2}$$

As shown in the equation above, although we intended to carry out this calculation near the critical micelle concentration of SDS, $\rho/N_A = 8.2\,\text{mM}$, this condition has turned out to be irrelavent for the ionic strength because of the existing relationship between effective charges and density cancels with the density dependence of ionic strength. That is, in a more concentrated system, the effective charges will be proportionately weaker, giving no net changes in ionic strength. Finally, we can combine these to find the Debye length

$$\lambda_D = \sqrt{\frac{\epsilon_r\epsilon_0 k_BT}{e^2\sum_i \rho_i z_i^2}} = \frac{n}{\sqrt{6s}}\delta$$

which for most reasonable values of s and n is similar to or smaller than the size of a lattice cell. From this analysis, we see the effective range of electrostatics in these systems with 1:1 stoichiometry is extremely short regardless of the sufactant concentration.

4.3.2 Charge neutrality

Although we have identified the effective valences z_+ and z_- using the the structure of the amphiphile , it is not clear whether these are the total charge on all the head cells or the charge of a single head cell. This section discusses how charges are distributed to the cells representing the molecule.

The lattice cell charges can be calculated by integrating the charge density in the effective hamiltonian over the extent of a lattice cell. The effective hamiltonian for the Coulombic system is

$$F_C = \frac{1}{2V}\sum_{\mathbf{k}}\sum_{i,j=+,-}\frac{\hat{\rho}_i(\mathbf{k})\hat{\rho}_j(-\mathbf{k})}{n_i n_j}\cdot\frac{4\pi z_i z_j}{k^2}$$

from which we see the charge density fields are $\left\{\frac{\hat{\rho}_+(\mathbf{k})z_+}{n_+}, \frac{\hat{\rho}_-(\mathbf{k})z_-}{n_-}\right\}$ or their real space counterparts $\{\rho_+(\mathbf{r})z_+/n_+, \rho_-(\mathbf{r})z_-/n_-\}$, or in surfactant terms, $\{\rho_H(\mathbf{r})z_H/n_H, \rho_T(\mathbf{r})z_T/n_T\}$. Integrating these gives the charge on any lattice cell. If we are integrating over a region that is occupied by the positive (head group) charge density, then the lattice cell's charge is

$$q_H = \int_{\delta^3} d\boldsymbol{r}\rho_H(\mathbf{r})z_H/n_H = \frac{z_H}{n_H}\int_{\delta^3} d\boldsymbol{r}\rho_H(\mathbf{r}) = \frac{z_H}{n_H}\,.$$

This result says that the charges we established before through the electrostatic analogy are a total charge on all head group cells (z_H) and a total charge on all tail group cells (z_T). To apply the electrostatic analogy to a lattice model, we must evenly distribute these total charges across all head group cells and the same for tail group cells. Since the two total charges are equal and opposite, charge neutrality and hence also stoichiometry is preserved in this procedure of distributing the charges.

4.3.3 Size and density of isomorphic charges

Comparing F_S and F_C determines the charge isomorphism:

$$z_+ = -z_- = \sqrt{\frac{3k_BT}{4\pi\rho\Delta^2}}$$

which in reduced units of k_BT and lattice parameter δ is:

$$|z| = \sqrt{\frac{3k_BT}{4\pi\rho\Delta^2}} = \sqrt{\frac{3k_BT\delta}{4\pi\left(\rho\delta^3\right)\left(\Delta/\delta\right)^2}} = \sqrt{\frac{3k_BT\delta}{4\pi\bar{\rho}\bar{\Delta}^2}}$$

where I defined dimensionless quantities $\bar{\rho} = \rho\delta^3$ and $\bar{\Delta} = \Delta/\delta$. Using SDS at its critical micelle concentration (8.2mM) as a model system, the reduced density and length are

$$\begin{aligned}\bar{\rho} &= 8.2 \times 6.02 \times 10^{23}\text{molecules/m}^3 \times \left(1.68 \times 10^{-10}\text{m}\right)^3 \\ &= 2.3 \times 10^{-5}\text{molecules} \\ \bar{\Delta} &= 5 \times 10^{-10} / \left(1.68 \times 10^{-10}\text{m}\right) = 3\end{aligned}$$

This means in our system of reduced units where we have effectively $k_B T = 1$ and $\delta = 1$, the isomorphic charges should be

$$z_H = -z_T = \sqrt{\frac{3}{4\pi \times 2.3 \times 10^{-5} \times 3^2}} = 33.7\,.$$

And the number of head or tail particles that should be placed into the $50\delta \times 50\delta \times 50\delta$ box is determined by the molar volume, $v_m = 50\mathring{\text{A}}^2 \times 1\text{nm} = 500\mathring{\text{A}}^3 \approxeq 105\,\delta^3$ per molecule, and the density of molecules $\rho = 8.2\text{mol/m}^3$:

$$\begin{aligned}N_{\text{surf}} &= \rho V \\ &= \left(8.2 \times 6.02 \times 10^{23}\text{molecules/m}^3\right) \times \left(50 \times 1.68 \times 10^{-10}\text{m}\right)^3 \\ &= 2.9 \text{ molecules} \\ V_{\text{surf}} &= N_{\text{surf}} \times v_m \\ &= 2.9 \text{ molecules} \times \left(105\,\delta^3/\text{molecule}\right) \\ &= 300\,\delta^3\end{aligned}$$

Inspecting SDS's molecular structure in Fig. 4.3, we see that roughly 80% of this volume is occupied by the tail atoms while the rest 20% is occupied by the head group. Given that each molecule occupies 105 of $1 \times 1 \times 1$ cells, we have $n_H = 105 \times 20\% \approxeq 20$ sites. Since the tail groups site on $2 \times 2 \times 2$ blocks. For the entire system containing 3 molecules, the number of tail and head groups should be set to

$$\begin{aligned}N_{\text{tail}} &= 240\,\delta^3/8\delta^3 = 30 \text{ blocks} \\ N_{\text{head}} &= 60\,\delta^3/1\delta^3 = 60 \text{ sites}\end{aligned}$$

SDS's structure also implies that the charges on the head and tail groups are distributed as follows

$$\begin{aligned}z_H/n_H &= 33.7/20 = 1.7 \text{ per site} \\ z_T/n_T &= -33.7/10 = -3.4 \text{ per block} = -0.425 \text{ per site}\,.\end{aligned}$$

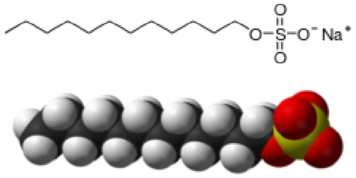

Figure 4.3: Sodium dodecyl sulfate (SDS).

4.4 Electrostatic analogy for a mixture of multiple amphiphilic species

As presented by Wu et al. in [26], for every surfactant species, the type-type intramolecular correlation matrix is found by summing correlations between groups of atoms (indexed by α, γ) where the groups belong to the particular types of interest (indexed by i, j), i.e. hydrophilic or hydrophobic:

$$\hat{\omega}_{ij}(k) = \sum_{\alpha \in i} \sum_{\gamma \in j} \hat{\omega}_{\alpha\gamma}(k)\,. \tag{4.1}$$

In a mixture of multiple surfactant species, we can continue to view all hydrophilic groups as one type and all hydrophobic groups as another and explore what the correlations between these are like. This means that the type-type correlation matrix will contain contributions from multiple intramolecular group-group correlation matrices,

$$\hat{\omega}_{ij}(k) = \sum_{m=1}^{N_s} \hat{\omega}_{ij}^{(m)}(k) = \sum_{m=1}^{N_s} \sum_{\alpha \in i} \sum_{\gamma \in j} \hat{\omega}_{\alpha\gamma}^{(m)}(k)\,, \tag{4.2}$$

where N_s is the total number of surfactant species present. In the following, I will work out what this sum means for the effective interactions in the simplest case of two surfactant species (labeled 1 and 2) in solution, and they have the same number of hydrophobic/hydrophilic groups ($n_A^{(1)} = n_A^{(2)} = n_A$, $n_B^{(1)} = n_B^{(2)} = n_B$). Suppose we have two surfactant species with overall density $\bar{\rho}_1$ and $\bar{\rho}_2$, the type-type intramolecular density correlation function is

$$\begin{aligned} \rho\hat{\omega}_{ij}(k) &= \bar{\rho}_1 \hat{\omega}_{ij}^{(1)}(k) + \bar{\rho}_2 \hat{\omega}_{ij}^{(2)}(k) \\ &= \bar{\rho}_1 \left[n_i n_j - k^2 \left(\Delta_{ij}^{(1)} \right)^2 /6 \right] + \bar{\rho}_2 \left[n_i n_j - k^2 \left(\Delta_{ij}^{(2)} \right)^2 /6 \right]\,. \end{aligned} \tag{4.3}$$

The determinant of this 2 × 2 matrix to $\mathcal{O}(k^2)$ is

$$\begin{aligned}\det\rho\hat{\omega}(k) &= \left[\bar{\rho}_1^2\Delta_1^2 + \bar{\rho}_2^2\Delta_2^2 + \bar{\rho}_1\bar{\rho}_2\left(\Delta_1^2 + \Delta_2^2\right)\right] n_A^2 n_B^2 k^2/3 \\ &= \rho^2\left[x_1^2\Delta_1^2 + x_2^2\Delta_2^2 + x_1 x_2(\Delta_1^2 + \Delta_2^2)\right] n_A^2 n_B^2 k^2/3 \\ &= \rho^2\left(x_1\Delta_1^2 + x_2\Delta_2^2\right) n_A^2 n_B^2 k^2/3\,, \end{aligned} \tag{4.4}$$

where $\Delta_m^2 = \left(\Delta_{AB}^{(m)}\right)^2/n_A n_B - \left(\Delta_{AA}^{(m)}\right)^2/n_A^2 - \left(\Delta_{BB}^{(m)}\right)^2/n_B^2$ is the average separation between the hydrophobic and hydrophilic groups in species m. I also defined $x_1 = \bar{\rho}_1/\rho$ and $x_2 = 1 - x_1 = \bar{\rho}_2/\rho$ to be the mole fractions of species 1 and 2, and $\rho = \bar{\rho}_1 + \bar{\rho}_2$ is the total density of surfactants. I will further define a subscript-free $\Delta^2 = x_1\Delta_1^2 + x_2\Delta_2^2$ to be the mole fraction weighted average separation between hydrophobic and hydrophilic groups. With these simplification in place, the inverse of the type-type correlation matrix condenses back into the familiar form

$$\rho^{-1}\hat{\omega}^{-1}(k) = \frac{1}{\rho k^2\Delta^2}\begin{bmatrix} 1/n_A^2 & -1/n_A n_B \\ -1/n_A n_B & 1/n_B^2 \end{bmatrix}. \tag{4.5}$$

4.4.1 Wu's approach to estimating the critical micelle concentration

This subsection follows the derivation presented by Maibaum et al.[47] and Wu et al. for a mixture of multiple surfactants. The critical micelle concentration is the concentration at which the free energy per unit volume of a solution of monomeric surfactants becomes higher than that of a surfactant assembly. Therefore, we need to find the free energies of both states as a function of concentration.

The soluble state has

$$F_{\text{soluble}}/L^3 = n\epsilon\rho\,. \tag{4.6}$$

where n is the number of hydrophobic or hydrophilic groups in each molecule (assuming that the surfactant has an equal number of either), ϵ is the energetic cost of moving a hydrophobic particle into the aqueous phase, and ρ is the density of surfactant molecules of any kind.

The micellar state has a contribution from the surface tension of the hydrophobic aggregate, and a charge contribution that mimics the head-tail connectivity,

$$F_{\text{micelle}} = E_{\text{surface}} + E_{\text{charge}}\,. \tag{4.7}$$

Assuming the entire system contains N micelles with uniform size R,

$$E_{\text{surface}} \sim N\sigma R^2 \tag{4.8}$$

$$E_{\text{charge}} \sim NR^5\left(\frac{z}{\Delta^3}\right)^2 \tag{4.9}$$

where the charge energy is calculated by applying Coulomb's law to a negatively charged sphere wrapped in a positively charged shell, both with charge density z/Δ^3. The total number of micelles can also be expressed in terms of the molecular parameters

$$N \sim \rho L^3 \Delta^3 / R^3 \tag{4.10}$$

After collecting everything into the free energy expression, we have the free energy density as a function of micelle size

$$F_{\text{micelle}}(R)/L^3 = \frac{\rho \Delta^3 \sigma}{R} + \frac{\rho z^2 R^2}{\Delta^3} \tag{4.11}$$

The system will always relax down to an optimal R^* that minimizes the free energy

$$R^* \sim \Delta^2 \sigma^{1/3} z^{-2/3} \tag{4.12}$$

and the isomorphic charges $z \sim (\beta \rho \Delta^2)^{-1/2}$. Substituting these into $F_{\text{micelle}}(R)$, the micellar state's free energy per unit volume can be written as

$$F_{\text{micelle}}/L^3 = \rho^{2/3} \Delta^{1/3} \sigma^{2/3} \beta^{-1/3}\,. \tag{4.13}$$

F_{soluble} scales with ρ, while F_{micelle} scales with $\rho^{2/3}$. The critical micelle concentration ρ_{cmc} is the concentration beyond which F_{soluble} exceeds F_{micelle},

$$\rho_{\text{cmc}} = \Delta \sigma^2 / \beta n \epsilon\,. \tag{4.14}$$

In a mixture of two surfactant species, Δ can be found by the mixing rule

$$\Delta = \sqrt{x_1 \Delta_1^2 + x_2 \Delta_2^2} \geq x_1 \Delta_1 + x_2 \Delta_2\,, \tag{4.15}$$

which means the cmc of a mixture will always be greater than or equal to a linear interpolation between two pure cmc's.

4.4.2 Alternative approach using the law of mass action

An alternative derivation starts with the law of mass action and also includes the monomeric surfactant's electrostatic contribution. Part of this derivation was presented in [47] for the one-amphiphile system.

Consider an equilibrium between soluble surfactants and those assembled into a monodisperse solution of micelles. I label the soluble surfactants “1” and the micelles formed from n surfactants “n”. The chemical potential of the soluble state is

$$\mu_1 = k_B T \ln\left(\frac{\rho_1}{\rho^{(0)}}\right) + \epsilon - \frac{z^2}{\Delta}\,, \tag{4.16}$$

where ρ_1 is the density of monomer surfactants, $\rho^{(0)}$ is the standard state density, and ϵ is the free energetic cost from solvating a hydrophobic tail in water as well as the free energetic gain from solvating a hydrophilic head in water.

The last term is the contribution from the electrostatic analogy of surfactant stoichiometry: Δ is the average head-tail separation, and z, the isomorphic charge, is related to the molecular density ρ by $z \sim (\beta\rho\Delta^2)^{-1/2}$. In principle, ρ should be the sum total of all surfactant molecules present in the system, aggregated or not,

$$\rho = \sum_{i=1}^{\infty} \rho_i = \rho_1 + \rho_n \,, \tag{4.17}$$

but for the purpose of finding the critical micelle concentration (cmc), most surfactants are soluble monomers when only a small, negligible number of micelles have formed. So I will approximate $\rho \approx \rho_1$. This simplifies the monomer chemical potential down to

$$\mu_1 = k_B T \ln\left(\frac{\rho_1}{\rho^{(0)}}\right) + \epsilon - \frac{1}{\beta\rho_1\Delta^3} \,. \tag{4.18}$$

Now turn to the chemical potential of a micelle. The aggregation number n (number of surfactants assembled into the micelle) is taken to be universal throughout the system:

$$\mu_n = k_B T \ln\left(\frac{\rho_n}{\rho^{(0)}}\right) + \gamma R^2 - R^5 \left(\frac{z}{\Delta^3}\right)^2 \,, \tag{4.19}$$

where ρ_n is the density of micelles, γ is the surface tension between tail groups and water, and R is the radius of the micelle. Again, the last term comes is the analogous Coulomb energy for a micelle, which is taken as a negatively charged core surrounded by a positively charged shell. If we assume that the micelle is simply made of individual surfactants packed together with no volume change, the micelle radius can be related to the surfactant head-tail separation by

$$R^3 \sim n\Delta^3 \,, \tag{4.20}$$

so μ_n can be rewritten in terms of n,

$$\beta\mu_n = \ln\left(\frac{\rho_n}{\rho^{(0)}}\right) + \beta\gamma\Delta^2 n^{2/3} - \frac{n^{5/3}}{\rho_1\Delta^3} \,. \tag{4.21}$$

At equilibrium, the associated micelle has the same chemical potential as the dissociated individual surfactants, $n\mu_1 = \mu_n$. This lets us relate the micelle density to monomer density,

$$\rho_n = \rho^{(0)} \exp\left[-\beta n \left(\gamma\Delta n^{-1/3} + \frac{1}{\beta\rho_1\Delta^3} + \frac{1 + n^{2/3}}{\beta\rho_1\Delta^3} - \epsilon - k_B T \ln\frac{\rho_1}{\rho^{(0)}}\right)\right] \,. \tag{4.22}$$

The first term in the exponent comes from the surface tension of the micelle, the second term is the effective Coulomb interaction in a monomer surfactant, the third term is the effective Coulomb interaction in a micelle, the fourth term is the energetic penalty of dissolving fatty tails, and the last term is the translational entropy of monomers.

As ρ_1 increases, the last term becomes more negative, and eventually, at the CMC, it begins to overwhelm the sum of all four terms that come before. This flips the sign of the exponent and suddenly increases ρ_n. Because the sum in the exponent is preceded by a factor of the aggregation number n, the greater the aggregation number, the more drastic this sign change will be. This is why micelles suddenly become favorable at the critical micelle concentration, even though micelle formation isn't a real phase transition.

The density of monomer when micelles first become favorable, ρ_{cmc}, can be found by setting the exponent in the previous expression to zero:

$$\ln \frac{\rho_{cmc}}{\rho^{(0)}} = -\beta\epsilon + \frac{1+n^{2/3}}{\rho_{cmc}\Delta^3} + \beta\gamma\Delta^2 n^{-1/3} \tag{4.23}$$

I rearranged Eq. 4.23 and used the Newton-Raphson method to solve

$$x = -\beta\epsilon + \beta\gamma\Delta^2 n^{-1/3} + \frac{1+n^{2/3}}{\rho^{(0)}\Delta^3} e^{-x}\,, \tag{4.24}$$

where $x = \ln \rho_{cmc}/\rho^{(0)}$. I set γ to the surface tension in lattice units, and $T = 1$. The aggregation number n is set to 50, which is a typical value for ionic small-molecule surfactants near room temperature and across a range of ionic strength. The resulting $\rho_{cmc}(\Delta)$ is a monotonically decreasing and slightly convex function. These two properties can now be used to predict how the mixed solution's CMC compares to that of a linear interpolation between two pure solution's CMCs. First, use the result for Δ_{mix} in subsection 4.4,

$$\rho_{cmc}\left(\Delta_{mix}\right) = \rho_{cmc}\left(\left[x_1\Delta_1^2 + x_2\Delta_2^2\right]^{1/2}\right)\,. \tag{4.25}$$

The argument on the right-hand side is greater than the mole-fraction weighted mean

$$\left[x_1\Delta_1^2 + x_2\Delta_2^2\right]^{1/2} \geq x_1\Delta_1 + x_2\Delta_2 \tag{4.26}$$

and since $\rho_{cmc}(\Delta)$ monotonically decreases, a larger argument value corresponds to a smaller function return,

$$\rho_{cmc}\left(\left[x_1\Delta_1^2 + x_2\Delta_2^2\right]^{1/2}\right) \leq \rho_{cmc}\left(x_1\Delta_1 + x_2\Delta_2\right)\,. \tag{4.27}$$

To further breakdown the right-hand side of the inequality, use Jensen's inequality (Convex transformation of a mean is less than or equal to the mean applied after convex transformation):

$$\rho_{cmc}\left(x_1\Delta_1 + x_2\Delta_2\right) \leq x_1\rho_{cmc}(\Delta_1) + x_2\rho_{cmc}(\Delta_2)\,. \tag{4.28}$$

Combining all these inequalities, we arrive at

$$\rho_{cmc}\left(\Delta_{mix}\right) \leq x_1\rho_{cmc}(\Delta_1) + x_2\rho_{cmc}(\Delta_2)\,, \tag{4.29}$$

which says that the mixed solution's CMC is always less than or equal to the mole-fraction-weighted average of the CMC's of the pure solutions. This confirms literature observations of surfactant mixtures having lower CMC than their pure components[48]. The extent that mixing lowers CMC depends on the size of the convex curvature of $\rho_{cmc}(\Delta)$, which further depends on ϵ, γ, n and T. The more convex it is, the more the CMC is lowered upon mixing.

4.5 Fluctuation-free estimate of the relative stabilities of assemblies

The electrostatic analogy allows us to calculate the "stoichiometric entropy" of amphiphilic assemblies. By evaluating the sum of this effective electrostatic energy and the lattice energies, we can estimate the relative stabilities among assembled structures. Although this approach neglects all fluctuations in the shapes and sizes of surfaces and structures, it sketches out the parts of parameter space we should look to find certain structures (e.g. micelles, bilayers) before we use more sophisticated methods to investigate the effect of these fluctuations. It is also useful to check whether the trends we see in this back-of-the-envelope calculation agree with our qualitative understanding of how variations in the amphiphile's molecular structure influence its preferences for different assembled structures.

4.5.1 Stoichiometry

The asymmetry factor σ denotes the ratio of the number of heads to the number of tails per amphiphile molecule:

$$\sigma = \frac{n_h}{n_t}$$

For amphiphiles with long tails and small head groups, $\sigma < 1$.

The charge densities from head (ρ_h) and tail (ρ_t) groups are related to the isomorphic charge z by

$$\rho_h = \frac{z}{n_h} \qquad \rho_t = -\frac{z}{n_t}$$

and their ratio is inverse to the ratio of the relative total number of heads (N_h) and tails (N_t) by stoichiometry:

$$\sigma = -\frac{\rho_t}{\rho_h} = \frac{N_h}{N_t}$$

Assuming these head and tail particles can assemble into a planar bilayer (2 head layers and 2 tail layers) with each head layer thickness d_h and tail layer thickness d_t. Again these layer thicknesses have to obey stoichiometry,

$$\sigma = \frac{d_h}{d_t}\,.$$

They can alternatively assemble into N_{mic} of spherical micelles of uniform inner radius r and outer radius $R = (1{+}\sigma)^{1/3}r$. Since the total amount of material in this system is $N_t = 2d_tL^2$ and $N_h = 2d_hL^2$, we know the number of micelles given their size r and R:

$$N_{\text{mic}} = \frac{2d_tL^2}{(4\pi/3)r^3} = \frac{3d_hL^2}{2\pi r^3\sigma}$$

4.5.2 Planar bilayer

Applying Gauss's Law to a patch of a flat infinite bilayer of area $L \times L$ and thickness $t = 2d_h + 2d_t$, the effective electrostatic energy per unit area is

$$U_{\text{elec}}^{\text{bilayer}} = -\frac{4\pi}{3}\rho_t\rho_h d_t d_h(d_h + d_t)L^2 = \frac{4\pi(\sigma+1)z^2d_h^3L^2}{3\sigma n_h^2\delta^6}$$

where $\rho_t = -z/n_t\delta^3$ and $\rho_h = +z/n_h\delta^3$ are the charge densities found from distributing the isomorphic charge z evenly onto the amphiphile's n_h head particles and n_t tail particles.

I can also count the number of nearest-neighbor pairs and get the Ising energy

$$U_{\text{ising}}^{\text{bilayer}} = -3\epsilon\left(\frac{L}{\delta}\right)^3 + 4\epsilon\left(\frac{L}{\delta}\right)^2$$

where the lattice energy is $\epsilon = -\frac{1}{4}\left(2\epsilon_{\text{ow}} - \epsilon_{\text{oo}} - \epsilon_{\text{ww}}\right)$, and δ is the lattice spacing. This part of the energy is independent of d_h and σ because the interface area between heads and tails remain constant for bilayers of various thickness. Combining these two contributions, we have

$$U^{\text{bilayer}} = U_{\text{elec}}^{\text{bilayer}} + U_{\text{ising}}^{\text{bilayer}} = \frac{4\pi(\sigma+1)z^2d_h^3L^2}{3\sigma n_h^2\delta^6} - 3\epsilon\left(\frac{L}{\delta}\right)^3 + 4\epsilon\left(\frac{L}{\delta}\right)^2\,.$$

4.5.3 Spherical micelles

Repeating the procedure above for a solution of N_{mic} micelles, the total electrostatic energy is

$$U_{\text{elec}}^{\text{micelle}}(r) = \frac{4\pi f(\sigma) z^2 r^2 d_h L^2}{3\sigma n_h^2 \delta^6}$$

where $f(\sigma)$ is a monotonically increasing function of σ:

$$f(\sigma) = (\sigma+1)^2 + \frac{1}{5}\sigma^2 - \frac{9}{5}(\sigma+1)^{5/3} + \sigma + \frac{4}{5}\,.$$

Similarly, counting the nearest-neighbor pairs gives

$$U_{\text{ising}}^{\text{micelle}}(r) = -3\epsilon\left(\frac{L}{\delta}\right)^3 + \left(\frac{12d_h}{r}\right)\epsilon\left(\frac{L}{\delta}\right)^2\,.$$

An estimate for the average micelle size can be found by minimizing $U_{\text{elec}}^{\text{micelle}}(r) + U_{\text{ising}}^{\text{micelle}}(r)$ with respect to the inner radius r. This gives the optimal inner and outer radius as

$$\begin{aligned} r^* &= \left[\frac{9}{2\pi f(\sigma)}\right]^{1/3}\left(\frac{n_h^2\epsilon\delta^4}{z^2}\right)^{1/3} \\ R^* &= (1+\sigma)^{1/3} r^* \end{aligned}$$

Substituting these into the expressions for the effective electrostatic energy and total ising coupling,

$$U^{\text{micelle}} = U_{\text{elec}}^{\text{micelle}}(r^*) + U_{\text{ising}}^{\text{micelle}}(r^*) = \frac{D\left[f(\sigma)\right]^{1/3} z^{2/3}\epsilon^{2/3} d_h L^2}{\sigma n_h^{2/3}\delta^{10/3}} - 3\epsilon\left(\frac{L}{\delta}\right)^3$$

where the numerical constant $D = \frac{4\pi}{3}\left(\frac{9}{2\pi}\right)^{2/3} + 12\left(\frac{2\pi}{9}\right)^{1/3} \simeq 16.0$.

4.5.4 Optimal thickness

The energetic difference between a planar bilayer and a collection of micelles is

$$\begin{aligned} \Delta U &= U_{\text{elec}}^{\text{bilayer}} + U_{\text{ising}}^{\text{bilayer}} - U_{\text{elec}}^{\text{micelle}}(r^*) - U_{\text{ising}}^{\text{micelle}}(r^*) \\ &= \frac{4\pi(\sigma+1) z^2 d_h^3 L^2}{3\sigma n_h^2 \delta^6} - \frac{D\left[f(\sigma)\right]^{1/3} z^{2/3}\epsilon^{2/3} d_h L^2}{\sigma n_h^{2/3}\delta^{10/3}} + 4\epsilon\left(\frac{L}{\delta}\right)^2\,. \end{aligned}$$

This expression lets us determine the optimal bilayer thickness $t^* = 2d_h^* + 2d_t^*$ with respect to our choice of Ising coupling strength ϵ and isomorphic charge z, as well as the amphiphiles head and tail numbers, n_h and $n_t = n_h/\sigma$. Minimizing ΔU with respect to d_h gives the optimal thickness

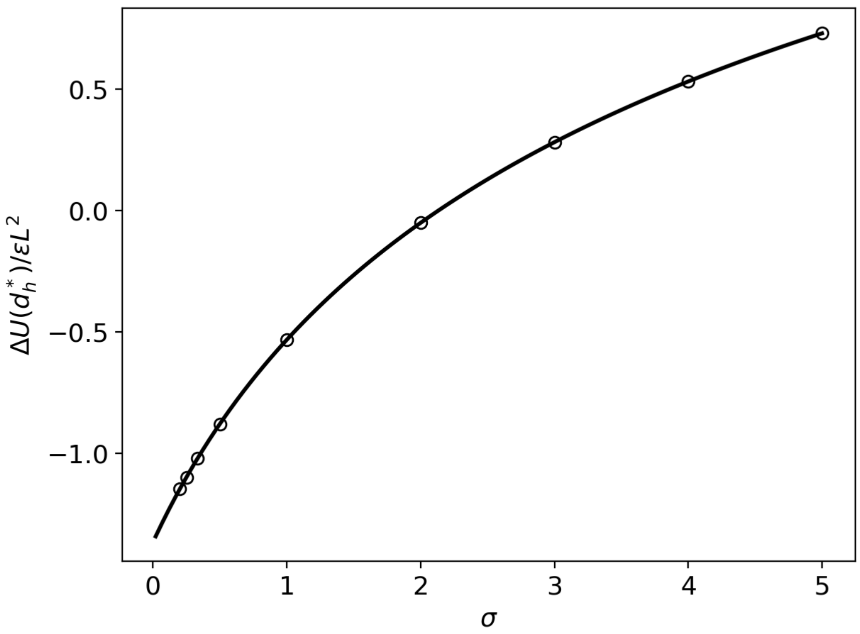

Figure 4.4: Energetic stability of bilayer relative to micelles as a function of head-tail asymmetry. Integral σ values are marked by circles.

$$
\begin{aligned}
d_h^* &= \left(\frac{D}{4\pi}\right)^{1/2} \frac{[f(\sigma)]^{1/6}}{(\sigma+1)^{1/2}} \left(\frac{n_h^2 \epsilon \delta^4}{z^2}\right)^{1/3} , \\
t^* &= 2 d_h^* \left(\frac{\sigma+1}{\sigma}\right) .
\end{aligned}
$$

At this optimal thickness, the stability of bilayer relative to that of a micelle is

$$
\Delta U(d_h^*)/L^2 = \left[4 - \frac{D^{3/2}}{3(\pi)^{1/2}} \left(\frac{f(\sigma)}{(\sigma+1)\sigma^2}\right)^{1/2}\right] \frac{\epsilon}{\delta^2}
$$

where the combination of numerical constants and σ in the prefactor is again a monotonically increasing function of σ, as shown in Fig. 4.4.

4.5.5 Rules for constructing a stable bilayer

With the equations worked out above, we can draw out some rules for constructing a stable bilayer:

1. The **Ising coupling** $\epsilon = -\frac{1}{4}\left(2\epsilon_{\mathrm{ow}} - \epsilon_{\mathrm{oo}} - \epsilon_{\mathrm{ww}}\right)$ is the only relevant lattice energy constant in determining $\Delta U(d_h^*)$. Because $\Delta U(d_h^*)$ turns out to scale proportionately

to ϵ, the choice of ϵ alone does not alter the relative energetic stability between micelles and bilayers. Under weak $\beta\epsilon$ the amphiphilic tail particles dissociate and no assembly will be observed, whereas under strong $\beta\epsilon$ the tail-head and tail-water interfaces lose the correct large-scale height fluctuations and become smooth. So we need an intermediate ϵ that allows for the correct height fluctuations while it is still "holding the assembly together".

2. The **asymmetry factor** σ is a measure of the relative number of head and tail groups in the amphiphile. As illustrated in Fig 4.4, small σ (small n_h, large n_t) favors bilayers while large σ (large n_h, small n_t) favors micelles. This trend is in agreement with the intuition that amphiphiles with large heads and small tails will assemble into micelles, while those with small head and large tail assemble into bilayers. When σ is small enough to favor bilayers ($\Delta U(d_h^*) < 0$ for $\sigma \lesssim 2$), increasing the strength of Ising couple ϵ will proportionally increase the energetic preference for bilayers. Furthermore, in addition to determining the overall energetic preference between micelles and bilayers, variations in σ also have some more subtle effects on the geometries of these structures: when we have a large σ that favors micelles, a larger σ gives smaller inner radius r^*; when we have a small σ that favors bilayers, a smaller σ gives thinner d_h^* but thicker d_t^*.

3. The **isomorphic charge density** on head particles $z/n_h\delta^3$ (or on tail particles $-z/n_t\delta^3 = -z\sigma/n_h\delta^3$) are determined by the amphiphile's average head-tail separation Δ, and the number of heads and tails n_h and n_t. These constants do not enter the equation for $\Delta U(d_h^*)$, which means they do not determine whether bilayers or micelles are more stable, at least from this limited, purely energetic perspective. However they do appear in the expressions for the size of bilayers d_h^*, d_t^* and the size of micelles r^*, R^*, which all scale $\sim n_h^{2/3} z^{-2/3}$. Since we know the isomorphic charge is related to temperature β, molecular density ρ and head-tail separation Δ by $z \sim (\beta\rho\Delta^2)^{-1/2}$, we know that decreasing the temperature, increasing the amphiphile size, or increasing its molecular density can all reduce z, consequently making r^*, R^*, d_h^* and d_t^* larger. Increasing n_h and n_t while holding their ratio σ constant have the same effect.

4. The calculations in this section neglect various entropic contributions to the free energy difference between bilayer and micelle configurations, including the surface fluctuations, translational entropy, and size fluctuations. Though it is difficult to know whether surface and size fluctuations favor one configuration or the other, it is reasonable to guess that the translational entropy strongly favors the micelles over the bilayers. Therefore for a bilayer to be stable, the Ising couplings and isomorphic charges must give rise to a strongly energetically favorable bilayer. In other words, simply having $\Delta U(d_h^*)/L^2 < 0$ is not enough - the energetic advantage for bilayers must be strong enough to outweigh the entropic preference for micelles. At the same

time, we could also reduce the translational entropy of micelles by simply looking at systems at high molecular densities.

4.5.6 Planar lamellae

In the sections above I calculated the relative energetic stability of a planar lipid bilayer against that of a solution of micelles. It is easy to extend this calculation to a lamellar phase in which planar bilayers on stacked on top of one another.

Previously during the calculations of the bilayer, the thicknesses of head and tail layers were fixed because the system was not constructed to be periodic in all three dimensions. But now the lamellar phase is periodic along all three axes, so we can fix the amphiphile molecular density at ρ_m instead. The number of head particles and tail particles in a box of dimensions (L_x, L_y, L_z) is then $N_h = \rho_m L_x L_y L_z n_h$ and $N_t = \rho_m L_x L_y L_z n_t$ respectively. Note that I am writing (L_x, L_y, L_z) instead of L^3 because this way it will be easier to check the extensivity of the energies later on.

Consider the case when these heads and tails assemble into n identical pieces of bilayers stacked along the z-axis. The thicknesses of the layers are then inversely proportional to n:

$$\begin{aligned} d_h &= \rho_m \delta^3 L_z n_h / 2n \\ d_t &= \rho_m \delta^3 L_z n_t / 2n \end{aligned}$$

We can now derive the total fluctuation-free energy of the system by considering the contributions from Ising couplings and electrostatics. The Ising part of the energy is simply the cohesive lattice energy of the whole water box, subtracted by the cost of having hydrophobe-hydrophile interfaces:

$$U_{\text{ising}}^{\text{lamella}}(n) = -3\epsilon \frac{L_x L_y L_z}{\delta^3} + 4n\epsilon \frac{L_x L_y}{\delta^2}$$

As we increase n, more interfaces are created and so $U_{\text{ising}}^{\text{lamella}}$ rises linearly with n.

For the effective electrostatic energy, the extension from our calculation for the planar bilayer case is again simple because each bilayer is overall charge-neutral. By Gauss's Law, there is no field in the space between the bilayers, so the effective electric field from one bilayer will not feel that of its neighbors. So the total electrostatic energy is that of a single bilayer multiplied by the number of bilayers n:

$$\begin{aligned} U_{\text{elec}}^{\text{lamella}}(n) &= \frac{4\pi(\sigma+1) z^2 d_h^3 L_x L_y n}{3\sigma n_h^2 \delta^6} \\ &= \frac{\pi}{6} \cdot \frac{(n_h + n_t) z^2 \rho_m^3 \delta^3 L_x L_y L_z^3}{n^2} \end{aligned}$$

where on the second line I made the change of variable from d_h to ρ_m. The electrostatics is proportional to $1/n^2$. It favors a larger n because in that case the opposite charges are better mixed and the system is more locally charge-neutral. This preference for larger n opposes the Ising coupling's preference for a smaller n, which reduces the cost of creating interfaces.

Since the electrostatics and Ising energies push the preferred n in opposite directions, it is possible to derive a preferred number of bilayers n^* at fixed molecular density ρ_m, temperature β and molecular structure $\{z, n_h, n_t\}$ by minimizing $U^{\text{lamella}}_{\text{ising}}(n) + U^{\text{lamella}}_{\text{elec}}(n)$ with respect to n. Minimization yields

$$n^* = \left(\frac{\pi}{12} \frac{z^2(n_h + n_t)\delta^5}{\epsilon} \right)^{1/3} \rho_m L_z$$

This is the optimal number of bilayers inside the periodic box of size (L_x, L_y, L_z), with the surface normals aligned with the z-axis. The proportionality to L_z is expected, as doubling the number of periodic images along z-axis should double the number of bilayers. The proportionality to ρ_m is also expected under our assumption that the system is dilute enough that the bilayers do not interact with one another. This assumption is likely going to break down for high concentrations or lipids that have low bending rigidities, which allow greater height fluctuations along z-axis.

Even entropy from individual bilayer's height fluctuations is ignored for the current calculation, as is the micelle phase's entropy of translation, because solving for optimal configuration parameters, such as the optimal micellar radii or the optimal number of bilayers, becomes much less tractable analytically when these extra entropic terms are incorporated into the free energy.

Substituting n^* into $U^{\text{lamella}}_{\text{ising}} + U^{\text{lamella}}_{\text{elec}}$, we find the optimal lamellar phase energy to be

$$\begin{aligned} U^{\text{lamella}}(n^*) &= U^{\text{lamella}}_{\text{ising}}(n^*) + U^{\text{lamella}}_{\text{elec}}(n^*) \\ &= \frac{\epsilon L_x L_y L_z}{\delta^3} \left\{ -3 + \rho_m \delta^3 \left[\frac{18\pi(n_h + n_t)z^2}{\delta\epsilon} \right]^{1/3} \right\} . \end{aligned}$$

Comparing this with the micellar phase's energy at the optimal inner radius r^*

$$U^{\text{micelle}}(r^*) = \frac{\epsilon L_x L_y L_z}{\delta^3} \left\{ -3 + \rho_m \delta^3 \frac{D}{2\sigma} \left(\frac{f(\sigma) n_h z^2}{\delta\epsilon} \right)^{1/3} \right\} ,$$

we find the energy density of the lamellar phase relative to micelles,

$$\begin{aligned} \Delta u &= \left[U^{\text{lamella}}(n^*) - U^{\text{micelle}}(r^*) \right] / L_x L_y L_z \\ &= \rho_m (n_h z^2 \epsilon^2/\delta)^{1/3} \left\{ [18\pi(\sigma + 1)/\sigma]^{1/3} - [f(\sigma)]^{1/3} D/2\sigma \right\} . \end{aligned}$$

The constants and functions of σ inside the curly bracket are negative for all σ values, and are monotonically increasing with σ, as shown in Fig. 4.5. That is, the longer the amphiphile's tail, the more stable the lamellar phase is. The linear dependency on ρ_m indicates that the lamellar phase is favored at higher densities.

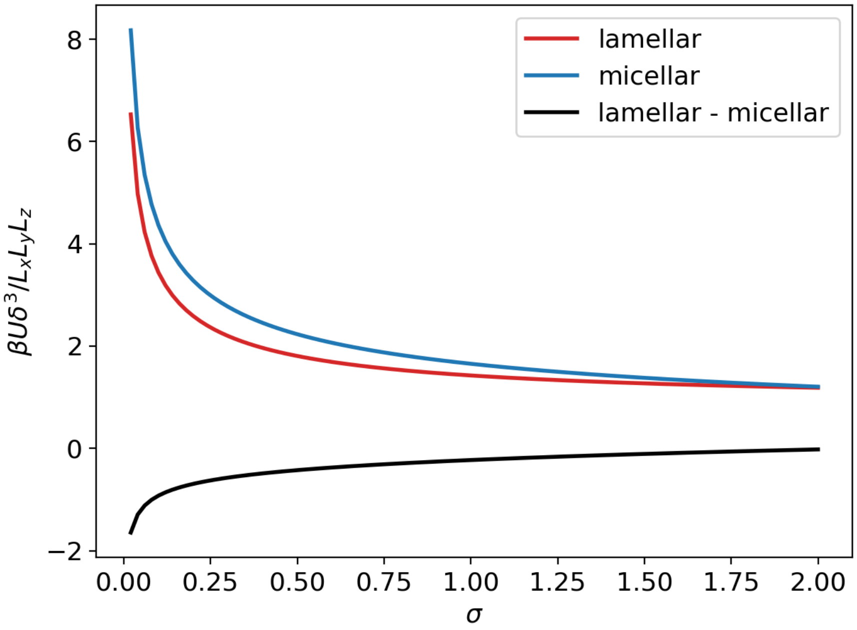

Figure 4.5: Energy density differences between the lamellar phase and the micellar phase.

These expressions for the optimal bilayer number and the energy density difference are useful as a guideline for setting up a numerical simulation with a good initial configuration of the lamellar phase that would stay relatively stable against decomposing into spherical micelles.

The usefulness of this approach is limited to a molecular density ρ_m that allows for both phases to exist. This result cannot apply to the high-density limit as mentioned previously, due to the expected strong inter-bilayer interactions. At very low density, this calculation also does not accommodate for the fact that there are not enough molecules in the system to make complete sheets of bilayers for the lamellar phase, therefore at that point, we need only consider the micelle phase and the soluble phase.

4.5.7 Dependency on lattice spacing choice

To make a lattice representation of the effective charges derived through the electrostatic analogy, we have to choose a lattice spacing δ. An amphiphile molecule has head group volume v_h and tail group volume v_t, which are represented by $n_h = v_h/\delta^3$ head particles and $n_t = v_t/\delta^3$ tail particles, each of volume δ^3 on the lattice.

The electrostatic analogy says that the head particles should carry $+z$ charge in total, while the tail particles should carry $-z$ charge in total, making the whole molecule overall charge-neutral. This implies that every head particle has a charge density of $+z/n_h\delta^3 = +z/v_h$, and every tail particle has a charge density of $-z/n_t\delta^3 = -z/v_t$. For the strength of Ising coupling as we change the lattice spacing, we have been emphasizing that the surface tension should be kept at a constant. A fluctuation-free, purely energetic surface tension match will require us to choose $\epsilon \propto \delta^2\gamma$, while a complete picture incorporating the capillary wave entropy will require us to choose $\epsilon \sim \delta\sqrt{k_BT\gamma}$.

The physical phenomenon we are interested in is the relative stability of the lamellar phase against spherical micelles. From a physical standpoint, the energy density difference should be invariant of the lattice spacing, because the lattice spacing is a fictional variable we use to put the configurations on a lattice and by doing so facilitate the free energy calculation. Let's inspect the expression below and see if the invariance holds:

$$\Delta u = \rho_m (n_h z^2 \epsilon^2/\delta)^{1/3} \left\{ [18\pi(\sigma+1)/\sigma]^{1/3} - [f(\sigma)]^{1/3} D/2\sigma \right\}$$

As we scale the lattice spacing by a multiple of $\delta' = m\delta$, if the amphiphile molecule was previously represented by $(n_h + n_t)$ lattice cells, then it is now only represented by $(n'_h + n'_t) = (n_h/m^3 + n_t/m^3)$ lattice cells. The new ising couple is $\epsilon' = m\epsilon$. The temperature β, molecular density ρ_m, effective charge valency z, and asymmetry factor σ do not change. So the new energy density is

$$\begin{aligned}\Delta u' &= \rho_m (n'_h z^2 \epsilon'^2/\delta')^{1/3} \left\{ [18\pi(\sigma+1)/\sigma]^{1/3} - [f(\sigma)]^{1/3} D/2\sigma \right\} \\ &= m^{-2/3} \rho_m (n_h z^2 \epsilon^2/\delta)^{1/3} \left\{ [18\pi(\sigma+1)/\sigma]^{1/3} - [f(\sigma)]^{1/3} D/2\sigma \right\}\end{aligned}$$

The residual scaling factor of $m^{-2/3}$ shows that the energy density is not lattice-spacing independent. If we were to make the lattice coupling scale in such a way that Δu is invariant to changes in δ, we should have chosen $\epsilon' = m^2\epsilon$ instead. Adding in entropy of translation and height fluctuations will change this result.

4.6 Representation of spontaneous curvature

In the previous section, we derived the relative stabilities of different amphiphilic assemblies inside the framework of the charge-frustrated Ising model. But in the literature, an amphiphilic molecule's spontaneous curvature is known to be one of the most significant factors that determine such stabilities.

Most notably, Israelachivili et al.[49] quantified the spontaneous curvature into the packing parameter, which is a ratio of head and tail group sizes. Amphiphiles with large head groups and small tail groups have concave spontaneous curvatures that stabilize oil-in-water micelles, while amphiphiles with small head groups and large tail groups have convex spontaneous curvatures that stabilize water-in-oil reverse micelles.

The charge-frustrated Ising model is concerned with only the distances among hydrophobic and hydrophilic groups within the molecule; the spontaneous curvature due to varying tail angles or head sizes is not explicitly included in the model. The goal of this chapter is to derive a higher-order correction to the charge-frustrated Ising model to produce asymmetry in the surfactant head and tail interactions so that a spontaneous curvature may emerge in the assemblies of layers.

4.6.1 Higher-order corrections to $\hat{\omega}_{ij}(k)$ in 3D

Following the previous formulation, I can write the density correlations to a higher order in k:

$$\hat{\omega}_{ij}(k) = n_i n_j - \frac{1}{6}k^2\Delta_{ij}^2 + \frac{1}{120}k^4\Phi_{ij}^4 - \frac{1}{5040}k^6\Omega_{ij}^6 + \mathcal{O}(k^8)$$

where I defined distances:

$$\Delta_{ij}^2 = \sum_{\alpha\in i}\sum_{\gamma\in j}\left\langle\left|\vec{r}^{(\alpha)} - \vec{r}^{(\gamma)}\right|^2\right\rangle$$

$$\Phi_{ij}^4 = \sum_{\alpha\in i}\sum_{\gamma\in j}\left\langle\left|\vec{r}^{(\alpha)} - \vec{r}^{(\gamma)}\right|^4\right\rangle$$

$$\Omega_{ij}^6 = \sum_{\alpha\in i}\sum_{\gamma\in j}\left\langle\left|\vec{r}^{(\alpha)} - \vec{r}^{(\gamma)}\right|^6\right\rangle$$

In matrix form, the density correlations are:

$$\hat{\omega}(k) = \begin{bmatrix} n_A^2 - \frac{1}{6}k^2\Delta_{AA}^2 + \frac{1}{120}k^4\Phi_{AA}^4 - \frac{1}{5040}k^6\Omega_{AA}^6 & n_A n_B - \frac{1}{6}k^2\Delta_{AB}^2 + \frac{1}{120}k^4\Phi_{AB}^4 - \frac{1}{5040}k^6\Omega_{AB}^6 \\ n_A n_B - \frac{1}{6}k^2\Delta_{AB}^2 + \frac{1}{120}k^4\Phi_{AB}^4 - \frac{1}{5040}k^6\Omega_{AB}^6 & n_B^2 - \frac{1}{6}k^2\Delta_{BB}^2 + \frac{1}{120}k^4\Phi_{BB}^4 - \frac{1}{5040}k^6\Omega_{BB}^6 \end{bmatrix}$$

To invert the density correlation matrix, first find determinant to $\mathcal{O}(k^4)$:

$$\det\hat{\omega}(k) = n_A^2 n_B^2\left(\frac{1}{3}k^2\Delta^2 + \frac{1}{12}k^4\Phi^4 + \frac{1}{720}k^6\Omega^6\right) + \mathcal{O}(k^8)$$

where the distances are:

$$\Delta^2 = \Delta_{AB}^2/n_A n_B - \Delta_{AA}^2/2n_A^2 - \Delta_{BB}^2/2n_B^2$$

$$\Phi^4 = \frac{1}{3}\left(\Delta_{AA}^2\Delta_{BB}^2 - \Delta_{AB}^4\right)/n_A^2 n_B^2 - \frac{1}{5}\left(\Phi_{AB}^4/n_A n_B - \Phi_{AA}^4/2n_A^2 - \Phi_{BB}^4/2n_B^2\right)$$

$$\Omega^6 = \left(2\Delta_{AB}^2\Phi_{AB}^4 - \Delta_{AA}^2\Phi_{BB}^4 - \Delta_{BB}^2\Phi_{AA}^4\right)/n_A^2 n_B^2 + \frac{1}{7}\left(2\Omega_{AB}^6/n_A n_B - \Omega_{BB}^6/n_B^2 - \Omega_{AA}^6/n_A^2\right)$$

In the following I will truncate the density correlation matrix at $\mathcal{O}(k^4)$. Consider first the AA-th element of the inverse (effective interaction) matrix. By the rules of inverting a 2×2 matrix:

$$\begin{aligned}
\hat{\omega}^{-1}_{AA}(k) &= \frac{\hat{\omega}_{BB}(k)}{\det \hat{\omega}(k)} \\
&= \frac{n_B^2 - \frac{1}{6}k^2\Delta^2_{BB} + \frac{1}{120}k^4\Phi^4_{BB} + \mathcal{O}(k^6)}{n_A^2 n_B^2 \left(\frac{1}{3}k^2\Delta^2 + \frac{1}{12}k^4\Phi^4 + \frac{1}{720}k^6\Omega^6\right) + \mathcal{O}(k^8)} \\
&= k^{-2}\left[n_B^2 - \frac{1}{6}k^2\Delta^2_{BB} + \frac{1}{120}k^4\Phi^4_{BB} + \mathcal{O}(k^6)\right] \times \\
&\qquad \left[n_A^2 n_B^2 \left(\frac{1}{3}\Delta^2 + \frac{1}{12}k^2\Phi^4 + \frac{1}{720}k^4\Omega^6\right) + \mathcal{O}(k^6)\right]^{-1} \\
&= k^{-2}\left[n_B^2 - \frac{1}{6}k^2\Delta^2_{BB} + \frac{1}{120}k^4\Phi^4_{BB} + \mathcal{O}(k^6)\right] \times \\
&\qquad \left[\frac{3}{n_A^2 n_B^2 \Delta^2} - \frac{3\Phi^4}{4n_A^2 n_B^2 \Delta^4}k^2 + \frac{1}{16 n_A^2 n_B^2}\left(\frac{3\Phi^8}{\Delta^6} - \frac{\Omega^6}{5\Delta^4}\right)k^4 + \mathcal{O}(k^6)\right]
\end{aligned}$$

Carry out the multiplication:

$$\begin{aligned}
\hat{\omega}^{-1}_{AA}(k) &= \frac{3/n_A^2}{k^2\Delta^2} - \left(\frac{3\Phi^4}{4\Delta^4 n_A^2} + \frac{\Delta^2_{BB}}{2n_A^2 n_B^2 \Delta^2}\right) + \\
&\qquad \left\{\frac{1}{16 n_A^2}\left(\frac{3\Phi^8}{\Delta^6} - \frac{\Omega^6}{5\Delta^4}\right) + \frac{\Phi^4\Delta^2_{BB}}{8n_A^2 n_B^2 \Delta^4} + \frac{\Phi^4_{BB}}{40 n_A^2 n_B^2 \Delta^2}\right\} k^2 + \mathcal{O}(k^4) \\
&= \frac{1}{n_A^2}\left(a_{AA}k^{-2} - b_{AA} + c_{AA}k^2 + \mathcal{O}(k^4)\right)
\end{aligned}$$

This can be written into a format that will be nicer for inversion back to real space later:

$$\begin{aligned}
\hat{\omega}^{-1}_{AA}(k) &= \frac{1}{n_A^2}\left(a_{AA}k^{-2} - b_{AA} + c_{AA}k^2 + \mathcal{O}(k^4)\right) \\
&= \frac{1}{n_A^2}\left(a_{AA}k^{-2} + \frac{-b_{AA}}{1 + (c_{AA}/b_{AA})k^2} + \mathcal{O}(k^4)\right)
\end{aligned}$$

where on the second line, I combined the zeroth and second order term to a form that retains the accuracy of the power series up to k^2. The constants are defined as:

$$\begin{aligned}
a_{AA} =& \frac{3}{\Delta^2} \\
b_{AA} =& \frac{3\Phi^4}{4\Delta^4} + \frac{\Delta_{BB}^2}{2n_B^2\Delta^2} \\
c_{AA} =& \frac{1}{16}\left(\frac{3\Phi^8}{\Delta^6} - \frac{\Omega^6}{5\Delta^4}\right) + \frac{\Phi^4\Delta_{BB}^2}{8n_B^2\Delta^4} + \frac{\Phi_{BB}^4}{40n_B^2\Delta^2}
\end{aligned}$$

As k is small, we can drop the $\mathcal{O}(k^4)$ terms in later calculations. Repeating this procedure for the other three elements, this gives the cleaned up effective correlation matrix (accurate up to $\mathcal{O}(k^2)$):

$$\hat{\omega}^{-1}(k) = \frac{1}{k^2}\begin{bmatrix} a_{AA}/n_A^2 & a_{AB}/n_An_B \\ a_{AB}/n_An_B & a_{BB}/n_B^2 \end{bmatrix} + \begin{bmatrix} \frac{1}{n_A^2}\cdot\frac{-b_{AA}}{1+k^2c_{AA}/b_{AA}} & \frac{1}{n_An_B}\cdot\frac{-b_{AB}}{1+k^2c_{AB}/b_{AB}} \\ \frac{1}{n_An_B}\cdot\frac{-b_{AB}}{1+k^2c_{AB}/b_{AB}} & \frac{1}{n_B^2}\cdot\frac{-b_{BB}}{1+k^2c_{BB}/b_{BB}} \end{bmatrix} + \mathcal{O}(k^4)$$

So on top of the head-head repulsion in the original formulation, they experience an additional screened electrostatic attraction by an amount proportionally to tail-tail rms distance. The invariance of head-head and tail-tail interactions with respect to head-head/tail-tail distances is broken.

In the expression above I defined more coefficients for the diagonal elements:

$$\begin{aligned}
a_{BB} =& \frac{3}{\Delta^2} \\
b_{BB} =& \frac{3\Phi^4}{4\Delta^4} + \frac{\Delta_{AA}^2}{2n_A^2\Delta^2} \\
c_{BB} =& \frac{1}{16}\left(\frac{3\Phi^8}{\Delta^6} - \frac{\Omega^6}{5\Delta^4}\right) + \frac{\Phi^4\Delta_{AA}^2}{8n_A^2\Delta^4} + \frac{\Phi_{AA}^4}{40n_A^2\Delta^2}
\end{aligned}$$

and off-diagonal coefficients carry opposite signs:

$$\begin{aligned}
a_{AB} =& -\frac{3}{\Delta^2} \\
b_{AB} =& -\left(\frac{3\Phi^4}{4\Delta^4} + \frac{\Delta_{AB}^2}{2n_An_B\Delta^2}\right) \\
c_{AB} =& -\frac{1}{16}\left(\frac{3\Phi^8}{\Delta^6} - \frac{\Omega^6}{5\Delta^4}\right) - \frac{\Phi^4\Delta_{AB}^2}{8n_An_B\Delta^4} - \frac{\Phi_{AB}^4}{40n_An_B\Delta^2}
\end{aligned}$$

4.6.2 Effects of higher order corrections on AB$_2$

In the $\mathcal{O}(k^{-2})$ model, the analogous charge densities interact with the same strength independent of how the head and tail groups are arranged spatially within the molecule.

To see how the higher-order terms introduce these elements, which are necessary for a non-zero spontaneous curvature, we can apply the expressions derived above to a simple molecule, B-A-B, with fixed bond length $r_{\mathrm{AB}} = \ell$ and tunable bond angle $\theta_{\mathrm{BAB}} = \theta$. While ℓ sets the length scale for interactions in this system, larger θ gives the molecule more steric bulk.

The results are shown in Fig.4.6 and Fig.4.7. θ does not change the relative ratio of strengths of the $\mathcal{O}(k^{-2})$ effective interactions. From this, we again see that the $\mathcal{O}(k^{-2})$ interactions represent the spatial constraint that overall the head and tail groups are separated, as explained in the Electrostatic Analogy paper.

The $\mathcal{O}(k^0)$ interactions, however, are capable of describing in more detail how the groups are arranged relative to each other, and the asymmetric changes to the molecule's shape when the arrangement changes. For $\mathcal{O}(k^0)$, the changes in these strengths follow the expected trend up to $\theta \approx \pi/2$: As θ widens from 0 to $\pi/2$, the surfactant tail groups are farther away from each other, and the graph shows that the tail-tail repulsion gets stronger while the head-head repulsion gets weaker. The behavior beyond $\theta \approx \pi/2$ opposes the trend before and instead of frustrating the phase separation of head and tail groups, it now facilitates such separation.

The $\mathcal{O}(k^2)$ interaction strengths follow the same trend up to about $\theta \approx \pi/2$. But for inverting the Fourier transform the sign of b_{ij}/c_{ij}, i.e. the ratio of the $\mathcal{O}(k^0)$ and $\mathcal{O}(k^2)$ strengths, will determine whether the real-space potential is of the from $e^{-\epsilon r}/r$ (exponential) or $\cos(\epsilon r)/r$ (oscillatory). On the plot of b_{ij}/c_{ij} we see all three types of interaction strengths change sign - they all switch between exponential and oscillatory at some point when θ becomes large. Note that for small θ values, when the tails are relatively close to each other and the molecule has the shape of a conventional surfactant, the head-head and tail-tail interactions start as oscillatory. Furthermore, $b_{ij}/c_{ij} < 0$ has consequences in the types of density fluctuations that are prohibited, as discussed below.

4.6.3 Prohibited density fluctuations

In the original $\mathcal{O}(k^0)$ electrostatic analogy, the free energy in reciprocal space is:

$$F_G = \frac{k_B T}{2V\rho} \sum_{\boldsymbol{k}} \frac{3}{k^2 \Delta^2} \left(\frac{|\hat{\rho}_A(\boldsymbol{k})|^2}{n_A^2} + \frac{|\hat{\rho}_B(\boldsymbol{k})|^2}{n_B^2} - \frac{\hat{\rho}_A(\boldsymbol{k})\hat{\rho}_B(-\boldsymbol{k})}{n_A n_B} - \frac{\hat{\rho}_A(-\boldsymbol{k})\hat{\rho}_B(\boldsymbol{k})}{n_A n_B} \right)$$

where $\rho_i(\boldsymbol{r}) = (1/V)\sum_{\mathbf{k}} e^{i\mathbf{k}.\mathbf{r}}\hat{\rho}_i(\mathbf{k})$, and $\hat{\rho}_i(\mathbf{k})$ is dimensionless. The $1/k^2$ prefactor becomes infinite for the term $\boldsymbol{k} \to 0$, unless the densities in the bracket become 0:

$$\lim_{\boldsymbol{k} \to 0} \left(\frac{|\hat{\rho}_A(\boldsymbol{k})|^2}{n_A^2} + \frac{|\hat{\rho}_B(\boldsymbol{k})|^2}{n_B^2} - \frac{\hat{\rho}_A(\boldsymbol{k})\hat{\rho}_B(-\boldsymbol{k})}{n_A n_B} - \frac{\hat{\rho}_A(-\boldsymbol{k})\hat{\rho}_B(\boldsymbol{k})}{n_A n_B} \right) = 0$$

Rearranging this,

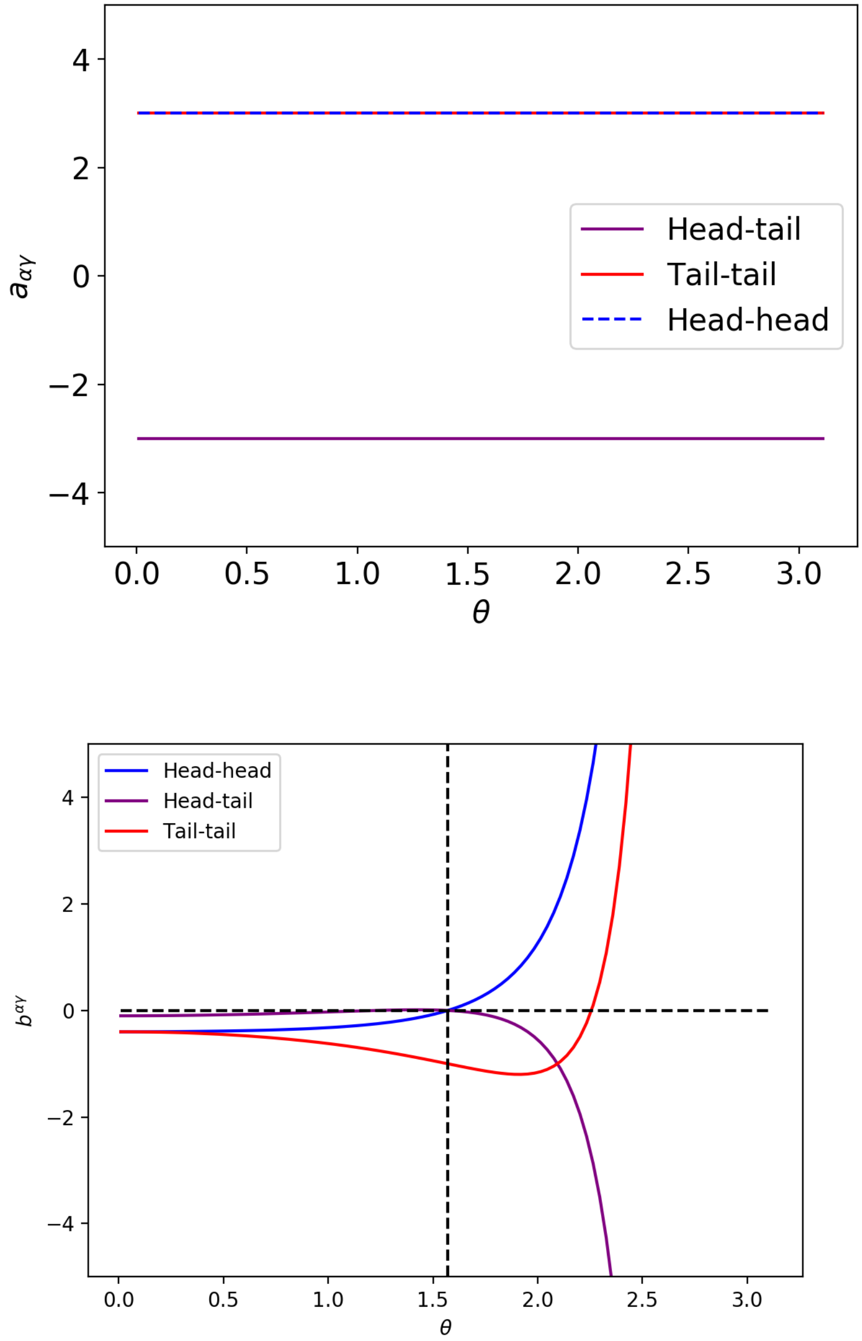

Figure 4.6: Coefficients for higher order corrections as a function of the inter-tail angle θ.

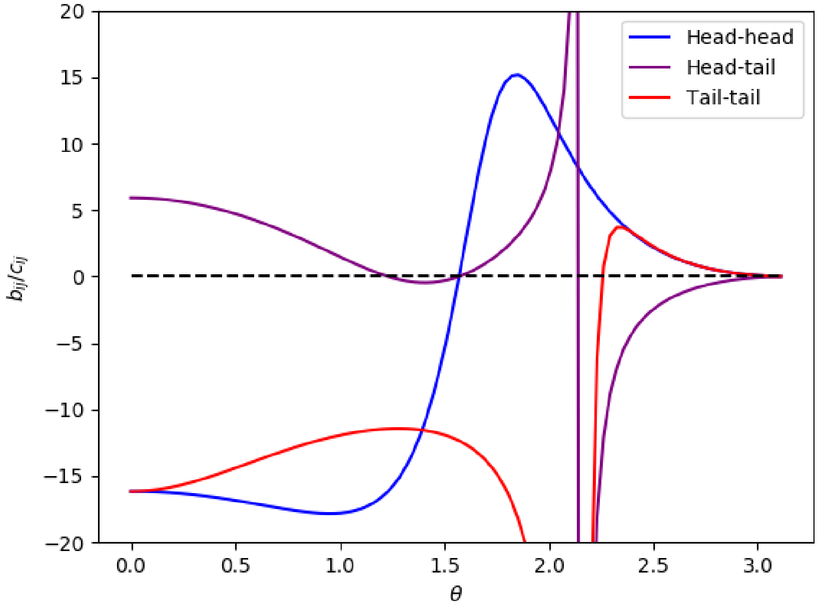

Figure 4.7: The ratio b_{ij}/c_{ij} as a function of the inter-tail angle θ.

$$\left(\frac{\hat{\rho}_A(0)}{n_A} - \frac{\hat{\rho}_B(0)}{n_B}\right)^2 = 0$$

where $\hat{\rho}_i(0) = V\sum_{\boldsymbol{r}} \rho_i(\boldsymbol{r}) = N_i$ is the total number of group i in the system. This equation, therefore, implies that the system must be overall charge neutral to avoid having prohibitively large free energy values. Translated into the language of amphiphiles, this constraint is equivalent to saying that stoichiometry of head and tail groups must be preserved in the field theory model.

Extending this to a free energy expression with higher-order k terms, if we are simply adding on the $\mathcal{O}(k^0)$ and $\mathcal{O}(k^2)$ terms, these terms do not go to infinity as $k \to 0$. However, if we decide to combine these two terms into the $\frac{1}{1+k^2}$ form, a new constraint emerges. The additional terms in the free energy expression are:

$$F_G' = \frac{k_B T}{2V\rho} \sum_{\boldsymbol{k}} \sum_{i,j=A,B} \frac{\hat{\rho}_i(\boldsymbol{k})\hat{\rho}_i(-\boldsymbol{k})}{n_i n_j} \left(\frac{-b_{ij}}{1 + k^2 c_{ij}/b_{ij}}\right)$$

When $c_{ij}/b_{ij} \geq 0$, no singular values are present; but as we saw in the section above, $c_{ij}/b_{ij} < 0$ is seen even in the simple AB$_2$ surfactant. This means that the $k = \sqrt{-b_{ij}/c_{ij}}$ terms must be eliminated from the sum by having $\hat{\rho}_i(\boldsymbol{k})\hat{\rho}_j(-\boldsymbol{k}) = 0$ at $|\boldsymbol{k}| = \sqrt{-b_{ij}/c_{ij}}$.

4.6.4 Real space effective interactions $\hat{\omega}_{ij}^{-1}(r)$

After finding the low-k expansion to $\mathcal{O}(k^2)$, the next step is to transform the following expression back to real space long-range effective interactions:

$$\begin{aligned}\hat{\omega}_{ij}^{-1}(k) &= \frac{1}{n_i n_j}\left(a_{ij}k^{-2} + \frac{-b_{ij}}{1+k^2 c_{ij}/b_{ij}} + \mathcal{O}(k^4)\right)\\ &= \frac{1}{n_i n_j}\left(a_{ij}k^{-2} + \frac{-b_{ij}^2/c_{ij}}{k^2+b_{ij}/c_{ij}} + \mathcal{O}(k^4)\right)\end{aligned}$$

For the case of $b_{ij}/c_{ij} \geq 0$, we have the following known result:

$$f_\epsilon(\vec{r}) = \frac{e^{-\epsilon r}}{4\pi r} \quad \Leftrightarrow \quad g_\epsilon(\vec{k}) = \frac{1}{k^2+\epsilon^2}$$

This means with $\epsilon = \sqrt{b_{ij}/c_{ij}}$,

$$\mathcal{F}^{-1}\left[\frac{-b_{ij}^2/c_{ij}}{k^2+b_{ij}/c_{ij}}\right] = \left(\frac{-b_{ij}^2/c_{ij}}{4\pi}\right)\mathcal{F}^{-1}\left[\frac{4\pi}{k^2+b_{ij}/c_{ij}}\right] = \frac{-b_{ij}^2 e^{-r\sqrt{b_{ij}/c_{ij}}}}{4\pi r c_{ij}}$$

And with $\epsilon \to 0$,

$$\mathcal{F}^{-1}\left[\frac{a_{ij}}{k^2}\right] = \frac{a_{ij}}{4\pi r}$$

Therefore in real space the effective interaction matrix is:

$$\omega_{ij}^{-1}(r) = \frac{a_{ij}}{4\pi r n_i n_j} - \frac{b_{ij}^2}{4\pi r n_i n_j c_{ij}} e^{-r\sqrt{b_{ij}/c_{ij}}}, \qquad b_{ij}/c_{ij} \geq 0$$

For the case of $b_{ij}/c_{ij} < 0$, contour integration gives

$$f_\epsilon(\vec{r}) = \frac{\cos(\epsilon r)}{4\pi r} \quad \Leftrightarrow \quad g_\epsilon(\vec{k}) = \frac{1}{k^2-\epsilon^2}$$

And so the real space effective interaction matrix is:

$$\omega_{ij}^{-1}(r) = \frac{a_{ij}}{4\pi r n_i n_j} - \frac{b_{ij}^2}{4\pi r n_i n_j c_{ij}} \cos\left(r\sqrt{-b_{ij}/c_{ij}}\right), \qquad b_{ij}/c_{ij} < 0$$

4.6.5 Lattice model

Landau-Ginzburg form of the effective Hamiltonian

Without changing the k-expansion, the additional terms in the free energy can be put into an Ising-like form:

$$\begin{aligned}
\beta F_G^{'} &= \frac{1}{2\rho V}\sum_{i,j=A,B}\sum_{\boldsymbol{k}}\frac{\hat{\rho}_i(\boldsymbol{k})}{n_i}(-b_{ij}+c_{ij}k^2)\frac{\hat{\rho}_j(-\boldsymbol{k})}{n_j}\\
&= \frac{1}{2\rho V^2}\sum_{i,j=A,B}\sum_{\boldsymbol{k}}\sum_{\boldsymbol{k}'}V\delta_{\boldsymbol{k},-\boldsymbol{k}'}\frac{\hat{\rho}_i(\boldsymbol{k})}{n_i}(-b_{ij}+c_{ij}k^2)\frac{\hat{\rho}_j(\boldsymbol{k}')}{n_j}
\end{aligned}$$

To turn the sum over k into integrals over r, use $\int_V d\boldsymbol{r}e^{i(\boldsymbol{k}+\boldsymbol{k}').\boldsymbol{r}} = V\delta_{\boldsymbol{k},-\boldsymbol{k}'}$:

$$\begin{aligned}
\beta F_G^{'} &= \frac{1}{2\rho V^2}\sum_{i,j=A,B}\int d\boldsymbol{r}\sum_{\boldsymbol{k}}\sum_{\boldsymbol{k}'}e^{i\boldsymbol{k}.\boldsymbol{r}}\frac{\hat{\rho}_i(\boldsymbol{k})}{n_i}(-b_{ij}+c_{ij}k'^2)e^{i\boldsymbol{k}'.\boldsymbol{r}}\frac{\hat{\rho}_j(\boldsymbol{k}')}{n_j}\\
&= \frac{1}{2\rho}\sum_{i,j=A,B}\int d\boldsymbol{r}\left(\frac{1}{V}\sum_{\boldsymbol{k}}e^{i\boldsymbol{k}.\boldsymbol{r}}\frac{\hat{\rho}_i(\boldsymbol{k})}{n_i}\right)(-b_{ij}+c_{ij}k'^2)\left(\frac{1}{V}\sum_{\boldsymbol{k}'}e^{i\boldsymbol{k}'.\boldsymbol{r}}\frac{\hat{\rho}_j(\boldsymbol{k}')}{n_j}\right)\\
&= \frac{1}{2\rho}\sum_{i,j=A,B}\int d\boldsymbol{r}\left(\frac{1}{V}\sum_{\boldsymbol{k}}e^{i\boldsymbol{k}.\boldsymbol{r}}\frac{\hat{\rho}_i(\boldsymbol{k})}{n_i}\right)(-b_{ij}-c_{ij}\nabla^2)\left(\frac{1}{V}\sum_{\boldsymbol{k}'}e^{i\boldsymbol{k}'.\boldsymbol{r}}\frac{\hat{\rho}_j(\boldsymbol{k}')}{n_j}\right)\\
&= \frac{1}{2\rho}\sum_{i,j=A,B}\int d\boldsymbol{r}\int d\boldsymbol{r}'\frac{\rho_i(\boldsymbol{r}')}{n_i}\delta(\boldsymbol{r}-\boldsymbol{r}')(-b_{ij}-c_{ij}\nabla^2)\frac{\rho_j(\boldsymbol{r})}{n_j}\\
&= \frac{1}{2\rho}\sum_{i,j=A,B}\int d\boldsymbol{r}\left[c_{ij}\nabla\frac{\rho_i(\boldsymbol{r})}{n_i}.\nabla\frac{\rho_j(\boldsymbol{r})}{n_j}-b_{ij}\frac{\rho_i(\boldsymbol{r})\rho_j(\boldsymbol{r})}{n_in_j}\right]
\end{aligned}$$

where I neglected the electrostatic term. The full effective hamiltonian should be:

$$\beta F_G = \frac{1}{2\rho}\sum_{i,j=A,B}\int d\mathbf{r}\int d\mathbf{r}'\frac{a_{ij}\rho_i(\mathbf{r})\rho_j(\mathbf{r})}{4\pi n_in_j\left|\mathbf{r}-\mathbf{r}'\right|}+\frac{1}{2\rho}\sum_{i,j=A,B}\int d\mathbf{r}\left[c_{ij}\nabla\frac{\rho_i(\mathbf{r})}{n_i}.\nabla\frac{\rho_j(\mathbf{r})}{n_j}-b_{ij}\frac{\rho_i(\mathbf{r})\rho_j(\mathbf{r})}{n_in_j}\right]$$

Hubbard-Stratonovich transformation

To see how this result can be mapped back into interactions in a lattice model, we use the Hubbard-Stratonovich transformation.[50] Doing the transformation backward from an effective field-theoretic hamiltonian back to a lattice model hamiltonian is more difficult, so I will start with a lattice model hamiltonian with unknown coefficients, and compare combinations of these coefficients with the coefficients of the field theory hamiltonian above at the end of the forward transformation:

$$-\beta\mathcal{H}[\{n_i^\alpha\}] = \frac{1}{2}\sum_{ij}\sum_{\alpha\gamma} n_i^\alpha K_{ij}^{\alpha\gamma} n_j^\gamma + \sum_{i,\alpha} \beta\mu^\alpha n_i^\alpha$$

where $i, j \in \{1, ..., N\}$ are indices for lattice sites; $\alpha, \gamma \in \{o, w, h, t\}$ are the particle types; $n_i^\alpha \in \{0, 1\}$ are occupation variables;$K_{ij}^{\alpha\gamma} = \beta J_{ij}^{\alpha\gamma}$ is the temperature-adjusted coupling between sites; and μ^α is the chemical potential associated with particle type α. The transformation takes the partition function for this system and applies the result of Gaussian functional integrals:

$$\begin{aligned}
\mathcal{Z} &= \sum_{\{n_i^o\}} \cdots \sum_{\{n_i^t\}} \exp\left[\frac{1}{2}\sum_{ij}\sum_{\alpha\gamma} n_i^\alpha K_{ij}^{\alpha\gamma} n_j^\gamma + \sum_{i,\alpha} \beta\mu^\alpha n_i^\alpha\right] \\
&= \sum_{\{n_i^o\}} \cdots \sum_{\{n_i^t\}} \int \mathcal{D}\left[\phi\right] \exp\left[-\frac{1}{2}\sum_{ij}\sum_{\alpha\gamma} \phi_i^\alpha \left(K^{-1}\right)_{ij}^{\alpha\gamma} \phi_j^\gamma + \sum_{i,\alpha} \left(\phi_i^\alpha + \beta\mu^\alpha\right) n_i^\alpha\right] \\
&= \int \mathcal{D}\left[\phi\right] \exp\left[-\frac{1}{2}\sum_{ij}\sum_{\alpha\gamma} \phi_i^\alpha \left(K^{-1}\right)_{ij}^{\alpha\gamma} \phi_j^\gamma\right] \times \sum_{\{n_i^o\}} \cdots \sum_{\{n_i^t\}} \exp\left[\sum_{i,\alpha} \left(\phi_i^\alpha + \beta\mu^\alpha\right) n_i^\alpha\right]
\end{aligned} \tag{4.30}$$

where the inverse couplings are defined by $\sum_{\xi k} K_{ik}^{\alpha\xi} \left(K^{-1}\right)_{kj}^{\xi\gamma} = \delta_{ij}\delta_{\alpha\gamma}$. Since the occupation variables are now uncoupled, they can be traced out. I will split the trace into two parts: a trace over all possible particle types at each site i, and a trace over all sites in the system.

$$\begin{aligned}
\sum_{\{n_i^o\}} \cdots \sum_{\{n_i^t\}} \exp\left[\sum_{i,\alpha} \left(\phi_i^\alpha + \beta\mu^\alpha\right) n_i^\alpha\right] &= \sum_{\{n_i^o\}} \cdots \sum_{\{n_i^t\}} \prod_{i=1}^{N} \exp\left[\sum_{\alpha} \left(\phi_i^\alpha + \beta\mu^\alpha\right) n_i^\alpha\right] \\
&= \prod_{i=1}^{N} \sum_{n_i^o}\sum_{n_i^w}\sum_{n_i^h}\sum_{n_i^t} \exp\left[\sum_{\alpha} \left(\phi_i^\alpha + \beta\mu^\alpha\right) n_i^\alpha\right] \\
&= \prod_{i=1}^{N} \sum_{\alpha} \exp\left[\left(\phi_i^\alpha + \beta\mu^\alpha\right)\right]
\end{aligned}$$

where the last step uses the fact that only one type of particle can occupy a site at a time (incompressibility constraint). Incorporating this result back to the partition function:

$$\mathcal{Z} = \int \mathcal{D}\left[\phi\right] \exp\left[-\frac{1}{2}\sum_{ij}\sum_{\alpha\gamma} \phi_i^\alpha \left(K^{-1}\right)_{ij}^{\alpha\gamma} \phi_j^\gamma + \sum_{i} \ln\left\{\sum_{\alpha} \exp\left[\left(\phi_i^\alpha + \beta\mu^\alpha\right)\right]\right\}\right]$$

where $\int \mathcal{D}\left[\phi\right]$ includes the appropriate normalization.

Saddle point approximation

Denoting the exponent as $-\beta\mathcal{H}_{\text{eff}}\left[\{\phi_i^\alpha\}\right]$, we now expand it in a power series of $\{\phi_i^\alpha\}$ around the saddle point. First, find the saddle point $\left\{\bar{\phi}_i^\alpha\right\}$:

$$\frac{\partial}{\partial\phi_k^\xi}\left(-\beta\mathcal{H}_{\text{eff}}\left[\{\phi_i^\alpha\}\right]\right) = -\sum_{i,\alpha}\left(K^{-1}\right)_{ki}^{\xi\alpha}\phi_i^\alpha + \frac{e^{\phi_k^\xi+\beta\mu^\xi}}{\sum_\alpha e^{\phi_k^\alpha+\beta\mu^\alpha}}$$

This derivative vanishes at the saddle point, giving a self-consistent equation for the field variables:

$$\sum_{i,\alpha}\left(K^{-1}\right)_{ki}^{\xi\alpha}\bar{\phi}_i^\alpha = \frac{\exp\left[\bar{\phi}_k^\xi+\beta\mu^\xi\right]}{\sum_\alpha\exp\left[\bar{\phi}_k^\alpha+\beta\mu^\alpha\right]} \quad\Leftrightarrow\quad \bar{\phi}_i^\alpha = \sum_{k,\xi}K_{ik}^{\alpha\xi}\frac{\exp\left[\bar{\phi}_k^\xi+\beta\mu^\xi\right]}{\sum_\alpha\exp\left[\bar{\phi}_k^\alpha+\beta\mu^\alpha\right]}$$

On the right hand side, the field variables always appear with the chemical potential. This observation helps us determine the solution through a derivative of the partition function with respect to the chemical potential:

$$\begin{aligned}\frac{\partial\ln\mathcal{Z}}{\partial\beta\mu^\xi} &= \left\langle\sum_i\frac{\partial}{\partial\beta\mu^\xi}\ln\left\{\sum_\alpha\exp\left[\left(\phi_i^\alpha+\beta\mu^\alpha\right)\right]\right\}\right\rangle\\ &= \left\langle\sum_i\frac{\exp\left[\left(\phi_i^\xi+\beta\mu^\xi\right)\right]}{\sum_\alpha\exp\left[\left(\phi_i^\alpha+\beta\mu^\alpha\right)\right]}\right\rangle\\ &= N\left\langle\frac{\exp\left[\left(\phi_1^\xi+\beta\mu^\xi\right)\right]}{\sum_\alpha\exp\left[\left(\phi_1^\alpha+\beta\mu^\alpha\right)\right]}\right\rangle\end{aligned}$$

But we also know, from inspecting the paritition function of the binary occupation variables, that this is equivalent to:

$$\frac{\partial\ln\mathcal{Z}}{\partial\beta\mu^\xi} = N\left\langle n_1^\xi\right\rangle_{\{n\}} = N\left\langle\frac{\exp\left[\left(\phi_1^\xi+\beta\mu^\xi\right)\right]}{\sum_\alpha\exp\left[\left(\phi_1^\alpha+\beta\mu^\alpha\right)\right]}\right\rangle_{\{\phi\}}$$

Now substitute this result back to the self-consistent equation, taking ensemble averages on both sides:

$$\left\langle\bar{\phi}_1^\alpha\right\rangle = \sum_{k,\xi}K_{1k}^{\alpha\xi}\left\langle\frac{\exp\left[\bar{\phi}_k^\xi+\beta\mu^\xi\right]}{\sum_\alpha\exp\left[\bar{\phi}_k^\alpha+\beta\mu^\alpha\right]}\right\rangle_{\{\phi\}} = \sum_{k,\xi}K_{1k}^{\alpha\xi}\left\langle n_1^\xi\right\rangle_{\{n\}}$$

Here $\left\langle\bar{\phi}_1^\alpha\right\rangle$ is a proxy for the effective field around any certain site, and the system is homogeneous so the effective field should have no location-dependence. Now we write $\bar{\phi}^\alpha = \left\langle\bar{\phi}_1^\alpha\right\rangle$ as the saddle point, and $\bar{n}^\xi = \left\langle n_1^\xi\right\rangle$ the average density of type ξ in the system.

The saddle point is therefore related to the average densities via a simple transformation by the occupation variables' coupling matrix:

$$\bar{\phi}^{\alpha} = \sum_{k,\xi} K_{1k}^{\alpha\xi} \bar{n}^{\xi}$$

Note this implies that the field variable $\phi^{\alpha}(\boldsymbol{r})$ is the continuum equivalent to the effective fields acting on the spins.

Next, to expand the effective hamiltonian around the saddle point we need the second derivative:

$$\frac{\partial^2}{\partial \phi_\ell^\eta \partial \phi_k^\xi} \left(-\beta \mathcal{H}_{\text{eff}} \left[\{\phi_i^\alpha\}\right]\right) = -\left(K^{-1}\right)_{k\ell}^{\xi\eta} + \frac{\delta_{k\ell}\delta_{\xi\eta} e^{\phi_k^\xi + \beta\mu^\xi}}{\sum_\alpha e^{\phi_k^\alpha + \beta\mu^\alpha}} - \frac{\delta_{k\ell} e^{\left(\phi_k^\xi + \phi_k^\eta\right) + \beta\left(\mu^\xi + \mu^\eta\right)}}{\left(\sum_\alpha e^{\phi_k^\alpha + \beta\mu^\alpha}\right)^2}$$

I will use the notation $\delta\phi_k^\xi = \phi_k^\xi - \bar{\phi}^\xi$. Writing out the expansion to second order,

$$-\beta \mathcal{H}_{\text{eff}} \left[\{\phi_i^\alpha\}\right] \simeq \text{const.} - \frac{1}{2} \sum_{\xi\eta} \sum_{k\ell} \delta\phi_k^\xi \left(K^{-1}\right)_{k\ell}^{\xi\eta} \delta\phi_\ell^\eta \tag{4.31}$$

$$+ \frac{1}{2} \sum_{\xi\eta} \sum_{k} \delta\phi_k^\xi \delta\phi_k^\eta \left(\delta_{\xi\eta} \bar{n}^\xi - \bar{n}^\xi \bar{n}^\eta\right) \tag{4.32}$$

Gradient expansion

I will now assume that the interactions $(K^{-1})_{k\ell}^{\xi\eta}$ are short ranged:

$$\left(K^{-1}\right)_{k\ell}^{\xi\eta} = \begin{cases} (K^{-1})_{\text{nn}}^{\xi\eta} & \text{if } \ell = k + a \\ 0 & \text{otherwise} \end{cases}$$

where the subscript "nn" denotes nearest-neighbors, and a is the index change when moving from one site to its nearest-neighbor site. Under this assumption, the second term in $-\beta\mathcal{H}_{\text{eff}}$ can be further broken down into a gradient expansion:

$$\begin{aligned} \sum_{k\ell} \delta\phi_k^\xi \left(K^{-1}\right)_{k\ell}^{\xi\eta} \delta\phi_\ell^\eta &= \sum_k \sum_a \left(K^{-1}\right)_{\text{nn}}^{\xi\eta} \delta\phi_k^\xi \delta\phi_{k+a}^\eta \\ &= \sum_k \left(K^{-1}\right)_{\text{nn}}^{\xi\eta} \delta\phi_k^\xi \sum_a \left(\delta\phi_k^\eta + \vec{a}\cdot\nabla\phi_k^\eta + \frac{1}{2}\left(\vec{a}\cdot\nabla\right)^2 \phi_k^\eta + \dots\right) \end{aligned}$$

The first-order term will cancel as we carry out the sum over all lattice vector directions. Now assuming the underlying lattice is cubic with lattice vector length ℓ and coordination number z, we can further simply the second-order term:

$$\sum_{k\ell} \delta\phi_k^\xi \left(K^{-1}\right)_{k\ell}^{\xi\eta} \delta\phi_\ell^\eta = \sum_k \left(K^{-1}\right)_{\text{nn}}^{\xi\eta} \delta\phi_k^\xi \left(z\delta\phi_k^\eta + \ell^2\nabla^2\phi_k^\eta + \ldots\right)$$

Substituting this back into $-\beta\mathcal{H}_{\text{eff}}$ and take the continuum limit of site index k:

$$\begin{aligned} -\beta\mathcal{H}_{\text{eff}}\left[\{\phi_i^\alpha\}\right] = \text{const.} - \frac{1}{2\ell^3}\sum_{\xi\eta}\int d\boldsymbol{r}\,&\left\{a^2\left(K^{-1}\right)_{\text{nn}}^{\xi\eta}\nabla\phi^\eta(\boldsymbol{r})\cdot\nabla\phi^\xi(\boldsymbol{r})\right.\\ &\left.\ldots - \left[z\left(K^{-1}\right)_{\text{nn}}^{\xi\eta} + \bar{n}^\xi\bar{n}^\eta - \delta_{\xi\eta}\bar{n}^\xi\right]\phi^\xi(\boldsymbol{r})\phi^\eta(\boldsymbol{r})\right\} \end{aligned}$$

This is now the same form as the effective hamiltonian in section 4.6.5 if we establish the mapping $\phi^\alpha(\boldsymbol{r}) = \rho_\alpha(\boldsymbol{r})/n_\alpha$. Comparing the coefficients of the gradient expansion we have

$$\begin{aligned} \frac{c_{\alpha\gamma}}{\rho\ell^6} &= \left(K^{-1}\right)_{\text{nn}}^{\alpha\gamma}/\ell \\ \frac{b_{\alpha\gamma}}{\rho\ell^6} &= -\left[z\left(K^{-1}\right)_{\text{nn}}^{\alpha\gamma} + \bar{n}^\alpha\bar{n}^\gamma - \delta_{\alpha\gamma}\bar{n}^\alpha\right]/\ell^3 \end{aligned}$$

Alternative transformation

The expressions that relate the lattice model's coupling matrix to our field theory hamiltonian's coefficients involve the inverse of $K_{ij}^{\alpha\gamma}$, which is difficult to compute or interpret. We can eliminate occurrences of K^{-1} by applying a linear transformation to our field variables $\{\phi_i^\alpha\}$, because such transformations will not change the Gaussian integral's outcome in Eq.4.30. The saddle point then directly corresponds to the average density, $\bar{\sigma}^\alpha = \bar{n}^\alpha$, which means that the field variable $\sigma^\alpha(\boldsymbol{r})$ is the continuum equivalent to the discrete spin variables.

Applying the transformation $\phi_i^\alpha = \sum_{m,\lambda} K_{im}^{\alpha\lambda}\sigma_m^\lambda$ to Eq.4.31 yields

$$\begin{aligned} -\beta\mathcal{H}_{\text{eff}}\left[\{\phi_i^\alpha\}\right] &\simeq \frac{1}{2}\sum_{\alpha\gamma}\left(\sum_i \phi_i^\alpha\phi_i^\gamma\left(\delta_{\alpha\gamma}\bar{n}^\alpha - \bar{n}^\alpha\bar{n}^\gamma\right) - \sum_{ij}\phi_i^\alpha\left(K^{-1}\right)_{ij}^{\alpha\gamma}\phi_j^\gamma\right) \\ &= \frac{1}{2}\sum_{\alpha\gamma}\left(\sum_i\sum_{m,\lambda}K_{im}^{\alpha\lambda}\sigma_m^\lambda\sum_{n,\xi}K_{in}^{\gamma\xi}\sigma_n^\xi\left(\delta_{\alpha\gamma}\bar{n}^\alpha - \bar{n}^\alpha\bar{n}^\gamma\right) - \sum_{ij}\sum_{m,\lambda}K_{im}^{\alpha\lambda}\sigma_m^\lambda\left(K^{-1}\right)_{ij}^{\alpha\gamma}\sum_{n,\xi}K_{jn}^{\gamma\xi}\sigma_n^\xi\right) \\ &= \frac{1}{2}\sum_{\lambda\xi}\sum_{mn}\sigma_m^\lambda\sigma_n^\xi\left[\sum_{\alpha\gamma}\sum_i K_{im}^{\alpha\lambda}K_{in}^{\gamma\xi}\left(\delta_{\alpha\gamma}\bar{n}^\alpha - \bar{n}^\alpha\bar{n}^\gamma\right) - K_{mn}^{\lambda\xi}\right] \end{aligned}$$

Let's now again consider the simplest step forward by assuming that the coupling matrix only contains nearest-neighbor interactions. The first term now reads

$$
\begin{aligned}
&\frac{1}{2}\sum_{\alpha\gamma\lambda\xi}\sum_{i}\sum_{a,a'}\sigma_{i+a}^{\lambda}\sigma_{i+a'}^{\xi}K_{i,i+a}^{\alpha\lambda}K_{i,i+a'}^{\gamma\xi}\left(\delta_{\alpha\gamma}\bar{n}^{\alpha}-\bar{n}^{\alpha}\bar{n}^{\gamma}\right)\\
=&\frac{1}{2}\sum_{\alpha\gamma\lambda\xi}\sum_{i}\left(\sum_{a}\sigma_{i+a}^{\lambda}\right)\left(\sum_{a'}\sigma_{i+a'}^{\xi}\right)K_{\text{nn}}^{\alpha\lambda}K_{\text{nn}}^{\gamma\xi}\left(\delta_{\alpha\gamma}\bar{n}^{\alpha}-\bar{n}^{\alpha}\bar{n}^{\gamma}\right)\\
=&\frac{1}{2}\sum_{\alpha\gamma\lambda\xi}\sum_{i}\left(z\sigma_{i}^{\lambda}+\ell^{2}\nabla^{2}\sigma_{i}^{\lambda}+\ldots\right)\left(z\sigma_{i}^{\xi}+\ell^{2}\nabla^{2}\sigma_{i}^{\xi}+\ldots\right)K_{\text{nn}}^{\alpha\lambda}K_{\text{nn}}^{\gamma\xi}\left(\delta_{\alpha\gamma}\bar{n}^{\alpha}-\bar{n}^{\alpha}\bar{n}^{\gamma}\right)\\
\simeq&\frac{1}{2}\sum_{\lambda\xi}\sum_{i}\left(z^{2}\sigma_{i}^{\lambda}\sigma_{i}^{\xi}+2z\ell^{2}\sigma_{i}^{\xi}\nabla^{2}\sigma_{i}^{\lambda}\right)\left[\sum_{\alpha\gamma}K_{\text{nn}}^{\alpha\lambda}K_{\text{nn}}^{\gamma\xi}\left(\delta_{\alpha\gamma}\bar{n}^{\alpha}-\bar{n}^{\alpha}\bar{n}^{\gamma}\right)\right]
\end{aligned}
$$

Taking the continuum limit $\sum_i \to \frac{1}{\ell^3}\int d\boldsymbol{r}$:

$$
-\frac{1}{2\ell^{3}}\sum_{\lambda\xi}\int d\boldsymbol{r}\chi_{\text{nn}}^{\lambda\xi}\left(-z^{2}\sigma^{\lambda}(\boldsymbol{r})\sigma^{\xi}(\boldsymbol{r})+2z\ell^{2}\nabla\sigma^{\xi}(\boldsymbol{r})\cdot\nabla\sigma^{\lambda}(\boldsymbol{r})\right)
$$

where $\chi_{\text{nn}}^{\lambda\xi}=\sum_{\alpha\gamma}K_{\text{nn}}^{\alpha\lambda}K_{\text{nn}}^{\gamma\xi}\left(\delta_{\alpha\gamma}\bar{n}^{\alpha}-\bar{n}^{\alpha}\bar{n}^{\gamma}\right)$. Under a similar gradient expansion, the second term in the effective hamiltonian becomes

$$
\begin{aligned}
&-\frac{1}{2}\sum_{\lambda\xi}\sum_{m}K_{\text{nn}}^{\lambda\xi}\sigma_{m}^{\lambda}\left(z\sigma_{m}^{\xi}+\ell^{2}\nabla^{2}\sigma_{m}^{\xi}+\ldots\right)\\
\to&-\frac{1}{2\ell^{3}}\sum_{\lambda\xi}\int d\boldsymbol{r}K_{\text{nn}}^{\lambda\xi}\left(z\sigma^{\lambda}(\boldsymbol{r})\sigma^{\xi}(\boldsymbol{r})-\ell^{2}\nabla\sigma^{\xi}(\boldsymbol{r})\cdot\nabla\sigma^{\lambda}(\boldsymbol{r})\right)
\end{aligned}
$$

Combining all terms in the effective hamiltonian

$$
\begin{aligned}
\beta\mathcal{H}_{\text{eff}}\left[\sigma^{\alpha}(\boldsymbol{r})\right]=\frac{1}{2\ell^{3}}\sum_{\lambda\xi}\int d\boldsymbol{r}&\left(2z\ell^{2}\chi_{\text{nn}}^{\lambda\xi}-\ell^{2}K_{\text{nn}}^{\lambda\xi}\right)\nabla\sigma^{\xi}(\boldsymbol{r})\cdot\nabla\sigma^{\lambda}(\boldsymbol{r})\\
&\ldots+\left(zK_{\text{nn}}^{\lambda\xi}-z^{2}\chi_{\text{nn}}^{\lambda\xi}\right)\sigma^{\lambda}(\boldsymbol{r})\sigma^{\xi}(\boldsymbol{r})
\end{aligned}
$$

Comparing to the coefficients derived in section 4.6.5, we find

$$
\begin{aligned}
\frac{c_{\alpha\gamma}}{2\rho\ell^{6}}&=\frac{2z\ell^{2}\chi_{\text{nn}}^{\lambda\xi}-\ell^{2}K_{\text{nn}}^{\lambda\xi}}{2\ell^{3}}\\
-\frac{b_{\alpha\gamma}}{2\rho\ell^{6}}&=\frac{zK_{\text{nn}}^{\lambda\xi}-z^{2}\chi_{\text{nn}}^{\lambda\xi}}{2\ell^{3}}
\end{aligned}
$$

So the nearest-neighbor coupling matrices can be written as linear combinations of the structural coefficients:

$$
\begin{aligned}
z\chi_{\mathrm{nn}}^{\alpha\gamma} = z\sum_{\lambda\xi} K_{\mathrm{nn}}^{\alpha\lambda} K_{\mathrm{nn}}^{\gamma\xi}\left(\delta_{\lambda\xi}\bar{n}^{\lambda} - \bar{n}^{\lambda}\bar{n}^{\xi}\right) &= \frac{1}{\rho\ell^5 z}\left(zc_{\alpha\gamma} - \ell^2 b_{\alpha\gamma}\right) \\
K_{\mathrm{nn}}^{\alpha\gamma} &= \frac{1}{\rho\ell^5 z}\left(zc_{\alpha\gamma} - 2\ell^2 b_{\alpha\gamma}\right)
\end{aligned}
$$

From the second equation, we can find four elements of the matrix $K_{\mathrm{nn}}^{\alpha\gamma}$ from the molecules structural coefficients $\{b_{\alpha\gamma}, c_{\alpha\gamma}\}$ for $\alpha, \gamma \in \{h, t\}$. But to fully solve the first equation, we need the rest 12 elements of $K_{\mathrm{nn}}^{\alpha\gamma}$, which describe the surfactant head/tail-solvent (oil/water) interactions and solvent-solvent interactions. Since the intramolecular pair correlation function $\hat{\omega}(k)$ only gives information about the interactions among surfactant head and/or tail groups, these 12 real-space elements cannot be uniquely determined but can instead freely chosen. Combining with the long-ranged interactions,

$$
\begin{aligned}
\beta\mathcal{H}_{\mathrm{eff}} &= \frac{1}{2}\sum_{\alpha\gamma}\left(\sum_{i\neq j} n_i^{\alpha} V_{ij}^{\alpha\gamma} n_j^{\gamma} - \sum_{\langle ij\rangle} n_i^{\alpha} K_{\mathrm{nn}}^{\alpha\gamma} n_j^{\gamma}\right) - \beta\sum_{i,\alpha}\mu^{\alpha} n_i^{\alpha} \\
V_{ij}^{\alpha\gamma} &= \frac{a_{\alpha\gamma}}{4\pi r_{ij}\rho} \\
K_{\mathrm{nn}}^{\alpha\gamma} &= \frac{1}{\rho\ell^5 z}\left(zc_{\alpha\gamma} - 2\ell^2 b_{\alpha\gamma}\right) .
\end{aligned}
$$

Bibliography

[1] Elizabeth C. Griffith, Rebecca J. Rapf, Richard K. Shoemaker, Barry K. Carpenter, and Veronica Vaida. Photoinitiated synthesis of self-assembled vesicles. *Journal of the American Chemical Society*, 136(10):3784–3787, 2014.

[2] Douglas G. Dalgleish and Milena Corredig. The structure of the casein micelle of milk and its changes during processing. *Annual Review of Food Science and Technology*, 3(1):449–467, 2012.

[3] R. Tuinier and C. G. De Kruif. Stability of casein micelles in milk. *Journal of Chemical Physics*, 117(3):1290–1295, 2002.

[4] Horst Robenek, Oliver Hofnagel, Insa Buers, et al. Butyrophilin controls milk fat globule secretion. *Proceedings of the National Academy of Sciences of the United States of America*, 103(27):10385–10390, 2006.

[5] Abdou Rachid Thiam and Lionel Forêt. The physics of lipid droplet nucleation, growth and budding. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, 1861(8):715–722, aug 2016.

[6] Kalthoum Ben M'barek, Dalila Ajjaji, Aymeric Chorlay, et al. ER Membrane Phospholipids and Surface Tension Control Cellular Lipid Droplet Formation. *Developmental Cell*, 41(6):591–604.e7, 2017.

[7] J. Zanghellini, F. Wodlei, and H.H. von Grünberg. Phospholipid demixing and the birth of a lipid droplet. *Journal of Theoretical Biology*, 264(3):952–961, jun 2010.

[8] Himanshu Khandelia, Lars Duelund, Kirsi I. Pakkanen, and John H. Ipsen. Triglyceride blisters in lipid bilayers: Implications for lipid droplet biogenesis and the mobile lipid signal in cancer cell membranes. *PLoS ONE*, 5(9):1–8, 2010.

[9] Ahanjit Bhattacharya, Henrike Niederholtmeyer, Kira A. Podolsky, et al. Lipid sponge droplets as programmable synthetic organelles. *Proceedings of the National Academy of Sciences of the United States of America*, 117(31):18206–18215, aug 2020.

[10] David Julian McClements. Emulsion design to improve the delivery of functional lipophilic components. *Annual Review of Food Science and Technology*, 1(1):241–269, 2010.

[11] Alice Verchère, Isabelle Broutin, and Martin Picard. Reconstitution of Membrane Proteins in Liposomes. In *Methods in Molecular Biology*, volume 1635, pages 259–282. 2017.

[12] Michael D. Buschmann, Manuel J. Carrasco, Suman Alishetty, et al. Nanomaterial delivery systems for mrna vaccines. *Vaccines*, 9(1):1–30, 2021.

[13] Yulia Eygeris, Siddharth Patel, Antony Jozic, Gaurav Sahay, and Gaurav Sahay. Deconvoluting Lipid Nanoparticle Structure for Messenger RNA Delivery. *Nano Letters*, 20(6):4543–4549, 2020.

[14] Jayesh A. Kulkarni, Maria M. Darjuan, Joanne E. Mercer, et al. On the Formation and Morphology of Lipid Nanoparticles Containing Ionizable Cationic Lipids and siRNA. *ACS Nano*, 12(5):4787–4795, 2018.

[15] Jayesh A. Kulkarni, Sam Chen, and Yuen Yi C. Tam. Scalable Production of Lipid Nanoparticles Containing Amphotericin B. *Langmuir*, 37(24):7312–7319, 2021.

[16] Jayesh A. Kulkarni, Dominik Witzigmann, Sarah B. Thomson, et al. The current landscape of nucleic acid therapeutics. *Nature Nanotechnology*, 16(6):630–643, 2021.

[17] Greg Van Anders, Daphne Klotsa, Andrew S. Karas, Paul M. Dodd, and Sharon C. Glotzer. Digital Alchemy for Materials Design: Colloids and Beyond. *ACS Nano*, 9(10):9542–9553, 2015.

[18] E. Gorter and F Grendel. On bimolecular layers of lipoids on the chromocytes of the blood. *Journal of Experimental Medicine*, 41(4):439–444, 1925.

[19] W Helfrich. Elastic Properties of Lipid Bilayers: Theory and Possible Experiments. *Zeitschrift fur Naturforschung - Section C Journal of Biosciences*, 28(11-12):693–703, dec 1973.

[20] Udo Seifert, Karin Berndl, and Reinhard Lipowsky. Shape transformations of vesicles: Phase diagram for spontaneous- curvature and bilayer-coupling models. *Physical Review A*, 44(2):1182–1202, 1991.

[21] Udo Seifert and Reinhard Lipowsky. Adhesion of vesicles. *Physical Review A*, 42(8):4768–4771, 1990.

[22] S A Safran, P A Pincus, David Andelman, and F. C. MacKintosh. Stability and phase behavior of mixed surfactant vesicles. *Physical Review A*, 43(2):1071–1078, 1991.

[23] Zachary A Levine, Richard M Venable, Max C Watson, et al. Determination of Biomembrane Bending Moduli in Fully Atomistic Simulations. *Journal of the American Chemical Society*, 136(39):13582–13585, oct 2014.

[24] Jacob N. Israelachvili, D John Mitchell, and Barry W Ninham. Theory of self-assembly of hydrocarbon amphiphiles into micelles and bilayers, 1976.

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

[26] David Wu, David Chandler, and Berend Smit. Electrostatic analogy for surfactant assemblies. *Journal of Physical Chemistry*, 96(10):4077–4083, 1992.

[27] Hyung June Woo, Carlo Carraro, and David Chandler. Quantitative molecular interpretation of curvature elasticity of saturated surfactant monolayers. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, 53(1):4–7, 1996.

[28] Michael W Deem and David Chandler. Charge-frustrated model of bicontinuous phases. *Physical Review E*, 49(5):4268–4275, may 1994.

[29] Rebecca J Rapf, Russell J Perkins, Haishen Yang, et al. Photochemical Synthesis of Oligomeric Amphiphiles from Alkyl Oxoacids in Aqueous Environments. *Journal of the American Chemical Society*, 139(20):6946–6959, may 2017.

[30] Shirley W.I. Siu, Kristyna Pluhackova, and Rainer A. Böckmann. Optimization of the OPLS-AA force field for long hydrocarbons. *Journal of Chemical Theory and Computation*, 8(4):1459–1470, apr 2012.

[31] Mark James Abraham, Teemu Murtola, Roland Schulz, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1-2:19–25, sep 2015.

[32] Promoting transparency and reproducibility in enhanced molecular simulations. *Nature Methods*, 16(8):670–673, aug 2019.

[33] Gareth A. Tribello, Massimiliano Bonomi, Davide Branduardi, Carlo Camilloni, and Giovanni Bussi. PLUMED 2: New feathers for an old bird. *Computer Physics Communications*, 185(2):604–613, feb 2014.

[34] D. P. Tieleman, D. van der Spoel, and H. J. C. Berendsen. Molecular Dynamics Simulations of Dodecylphosphocholine Micelles at Three Different Aggregate Sizes: Micellar Structure and Chain Relaxation. *The Journal of Physical Chemistry B*, 104(27):6380–6388, jul 2000.

[35] G.M. Torrie and J.P. Valleau. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *Journal of Computational Physics*, 23(2):187–199, feb 1977.

[36] Pieter Rein Ten Wolde, Maria J. Ruiz-Montero, and Daan Frenkel. Numerical calculation of the rate of homogeneous gas-liquid nucleation in a Lennard-Jones system. *Journal of Chemical Physics*, 110(3):1591–1599, 1999.

[37] Pieter Rein Ten Wolde and Daan Frenkel. Homogeneous nucleation and the Ostwald step rule. *Physical Chemistry Chemical Physics*, 1(9):2191–2196, 1999.

[38] Pieter Rein Ten Wolde and Daan Frenkel. Enhancement of protein crystal nucleation by critical density fluctuations. *Science*, 277(5334):1975–1978, 1997.

[39] Pavel L. Krapivsky, Sidney Redner, and Eli Ben-Naim. *A kinetic view of statistical physics*. Cambridge University Press, Cambridge, jan 2011.

[40] Victor K. La Mer. Nucleation in Phase Transitions. *Industrial Engineering Chemistry*, 44(6):1270–1277, 1952.

[41] Victor K. LaMer and Robert H. Dinegar. Theory, Production and Mechanism of Formation of Monodispersed Hydrosols. *Journal of the American Chemical Society*, 72(11):4847–4854, nov 1950.

[42] Christopher B. Whitehead, Saim Özkar, and Richard G. Finke. LaMer's 1950 Model for Particle Formation of Instantaneous Nucleation and Diffusion-Controlled Growth: A Historical Look at the Model's Origins, Assumptions, Equations, and Underlying Sulfur Sol Formation Kinetics Data. *Chemistry of Materials*, 31(18):7116–7132, sep 2019.

[43] Stephen G. Brush. History of the Lenz-Ising Model. *Reviews of Modern Physics*, 39(4):883–893, oct 1967.

[44] Hyung-June Woo, Carlo Carraro, and David Chandler. Assembly of extended interfaces and micelles: charge frustrated models of amphiphilic mixtures. *Faraday Discussions*, 104:183, 1996.

[45] Suriyanarayanan Vaikuntanathan and Phillip L Geissler. Putting Water on a Lattice: The Importance of Long Wavelength Density Fluctuations in Theories of Hydrophobic and Interfacial Phenomena. *Physical Review Letters*, 112(2):020603, jan 2014.

[46] Peter Debye and Erich Hückel. The theory of electrolytes. I. Freezing point depression and related phenomena. *Physikalische Zeitschrift*, 24(1):185–206, 1923.

[47] Lutz Maibaum, Aaron R Dinner, and David Chandler. Micelle Formation and the Hydrophobic Effect. *The Journal of Physical Chemistry B*, 108(21):6778–6781, may 2004.

[48] Gohsuke Sugihara, Shigemi Nagadome, Se Woung Oh, and Jeong So Ko. A review of recent studies on aqueous binary mixed surfactant systems. Technical Report 2, 2008.

[49] Jacob N. Israelachvili. Refinement of the fluid-mosaic model of membrane structure. *BBA - Biomembranes*, 469(2):221–225, 1977.

[50] J. E. Hirsch. Discrete Hubbard-Stratonovich transformation for fermion lattice models. *Physical Review B*, 28(7):4059–4061, oct 1983.

Appendix A

Computational techniques for the charge-frustrated Ising model

A.1 Monte Carlo moves

A.1.1 Bulk equilibrium

Detailed balance requires that the flow between old configuration o and new configuration n obey:

$$\begin{gathered} K(o \to n) = K(n \to o) \\ \mathcal{N}(o)\alpha(o \to n)\mathrm{acc}(o \to n) = \mathcal{N}(n)\alpha(n \to o)\mathrm{acc}(n \to o) \end{gathered}$$

where K is the flow of probability between the system's configurations, $\mathcal{N}$ is the probability of observing the system in a given configuration for the ensemble of interest, α is the probability of generating and proposing a configuration from the current configuration, and acc is the probability of accepting the proposed changes.

When detailed balance is satisfied, the system is guaranteed to evolve towards equilibrium. In the following, I use detailed balance to verify the validity of the generation and acceptance criteria of various moves used in the Monte Carlo simulations of the canonical ensemble,

$$\frac{\mathcal{N}(o)}{\mathcal{N}(n)} = \frac{\exp(-\beta U(o))}{\exp(-\beta U(n))} .$$

A.1.2 Spin swaps

One spin is randomly selected from the total of N spins. A neighbor of this spin is randomly selected among its $(3^d - 1)$ nearest and second-nearest neighbors. A swap between these two spins is proposed. Hence the generation probability is

$$\alpha(o \to n) = \frac{1}{N(3^d - 1)} = \alpha(n \to o)\,.$$

Given the generation probability, detailed balance requires the simple metropolis rule:

$$\frac{\mathrm{acc}(o \to n)}{\mathrm{acc}(n \to o)} = \frac{\mathcal{N}(n)\alpha(n \to o)}{\mathcal{N}(o)\alpha(o \to n)} = \exp\left(-\beta\left[U(n) - U(o)\right]\right)$$

which in practice can be programmed by setting the acceptance criterion as:

$$\mathrm{acc}(o \to n) = \min\left\{1, \exp\left(-\beta\left[U(n) - U(o)\right]\right)\right\}\,.$$

A.2 Minimum box size

To make sure this system will form micelles once isomorphic charges are included, the charge-free oil-water binary lattice gas must exhibit phase separation, because the inclusion of charges will not increase the propensity of aggregation but will only work against it. Without phase separation in the charge-free system, it is not possible to form micelles (mesoscopic aggegates).

This mean that with the given set of energy parameters ϵ_{oo}, ϵ_{ww} and ϵ_{ow}, the charge-free binary oil-water system must be below the critical temperature $T < T_c$, and the overall volume fraction ϕ_{oil} must be between the boundaries of coexistence, ϕ_{OR} and ϕ_{WR} . Here I denote the volume fraction of oil in the oil-rich phase ϕ_{OR}, and the volume fraction of oil in the water-rich phase ϕ_{WR}. When the fraction of system volume f is occupied by the oil-rich phase, the three volume fractions are related by

$$\phi_{\mathrm{oil}} V = \phi_{\mathrm{OR}} f V + \phi_{\mathrm{WR}}(1 - f)V$$

and by approximate symmetry in a half oil half water system,

$$\phi_{\mathrm{OR}} + \phi_{\mathrm{WR}} \approx 1\,.$$

Setting $\phi_{\mathrm{WR}} < \phi_{\mathrm{oil}} < \phi_{\mathrm{OR}}$, we expect a oil-rich slab of thickness d and cross-sectional area L^2, so the volume fraction accounting equation is now

$$\phi_{\mathrm{oil}} L^3 = \phi_{\mathrm{OR}} L^2 d + \phi_{\mathrm{WR}} L^2 (1 - d)$$

which can be rearranged and simplied using the approximate oil-water symmetry to

$$d = L \times \frac{\phi_{\mathrm{oil}} - \phi_{\mathrm{WR}}}{\phi_{\mathrm{OR}} - \phi_{\mathrm{WR}}} \approx L \times \frac{\phi_{\mathrm{oil}} - \phi_{\mathrm{WR}}}{1 - 2\phi_{\mathrm{WR}}}$$

So this tells us the thickness of the oil-rich slab we should expect in a system of box size L and overall oil volume fraction ϕ_{oil}. The other volume fractions, ϕ_{OR} and ϕ_{WR}, are boundaries of the coexistence region. They depend only on temperature and the cohesion energies, which are held fixed for our purposes. Another way to think about this is a minimal box size, L_{min}, below which no phase-separated slab is possible. Since our oil particles are $2 \times 2 \times 2$, for such a slab to exist we must have thickness of at least $d_{\text{min}} = 2$, so

$$L_{\text{min}} \approx 2 \times \frac{1 - 2\phi_{\text{WR}}}{\phi_{\text{oil}} - \phi_{\text{WR}}}$$

is smaller when ϕ_{oil} is large, and increases asymptotically if ϕ_{oil} goes close to its lower limit, ϕ_{WR}. This is in line with intution: the system has a greater tendency towards slab formation when there is a high concentration of oils; and if there is little excess of oil compared to the soluble concentration of oil in the water-rich phase, there is little tendency to produce a separate oil-rich phase.

Similarly, this equation can only be used to estimate the minimum concentration required for slab formation in a certain box size,

$$\phi_{\text{oil}}^{\text{min}} = \frac{d_{\text{min}}}{L}(1 - 2\phi_{\text{WR}}) + \phi_{\text{WR}} = \frac{d_{\text{min}}}{L} + \left(1 - \frac{2d_{\text{min}}}{L}\right)\phi_{\text{WR}} \,.$$

Another consideration comes after satisfying these minimum concentration or minimum system size constraints. If there is only a very small volume of the oil-rich phase, it might not form a slab as we expected above, but instead form a sphere to minimize its surface area. Micelles might still be able to form in such sphere-forming systems, but it is confusing because we could mistake the formation of a spherical aggregate because of finite size effects with the formation of micelles in macroscopic systems because of the competition of real physical forces.

This explains why we need to know for a system of overall oil volume fraction $\phi_{\text{oil}} \geq \phi_{\text{oil}}^{\text{min}}$, what minimum system size is required for the formation of a slab instead of a sphere. The total volume of the oil-rich phase is

$$V_{\text{OR}} = L^3 \times \frac{\phi_{\text{oil}} - \phi_{\text{WR}}}{1 - 2\phi_{\text{WR}}}$$

At any slab thickness, this volume gives rise to a slab surface area $A_{\text{slab}} = 2L^2$. But the sphere's surface area depends on its volume,

$$A_{\text{sphere}} = 4\pi \left(\frac{3V_{\text{OR}}}{4\pi}\right)^{2/3} = (36\pi)^{1/3} \left(\frac{\phi_{\text{oil}} - \phi_{\text{WR}}}{1 - 2\phi_{\text{WR}}}\right)^{2/3} L^2$$

For $A_{\text{sphere}} \geq A_{\text{slab}}$, we need

$$\phi_{\text{oil}} \geq \frac{2}{9\pi}(1 - 2\phi_{\text{WR}}) + \phi_{\text{WR}}$$

which is another lower bound on ϕ_{oil} that may be smaller or larger than the previous lower bound, depending on the system size L.

To summarise, to observe an oil-rich slab with minimal thickness d_{min} sites, the choice of system size and oil concentration needs to satisfy two requirements:

1. There is enough amount of oil to form a slab

$$\phi_{\text{oil}} \geq \frac{d_{\text{min}}}{L} + \left(1 - \frac{2d_{\text{min}}}{L}\right)\phi_{\text{WR}}$$

2. The system prefers a slab over a sphere

$$\phi_{\text{oil}} \geq \frac{2}{9\pi}(1 - 2\phi_{\text{WR}}) + \phi_{\text{WR}}$$

For large L, it is possible that Requirement 1 is satisfied but Requirement 2 is not. For smaller L, satisfying Requirement 1 will automatically satisfy Requirement 2. For the current choice of parameters, the site-site interactions are $\epsilon_{\text{oo}}/T = 0.37$, $\epsilon_{\text{ww}}/T = 1.25$ and $\epsilon_{\text{ow}}/T = 0.55$, and $\phi_{\text{WR}} = 0.0017$ in a box of $L = 50$. This means we need $\phi_{\text{oil}} \geq 0.042$ for Requirement 1 and $\phi_{\text{oil}} \geq 0.072$ for Requirement 2. It is curious to see that increasing the box size L helps with relaxing the first requirement for oil concentration but not the second.

A.3 Measurements of interface properties

The absolute location of the interface will drift over time, because the Monte Carlo moves by design do not explicitly preserve total momentum. But the location of the interface relative to the center of mass of one of the two phases should stay constant, so all the particle coordinates will now be taken as relative to the center of mass, $\mathbf{r}_c = (x_c, y_c, z_c)$.

The interface is assumed to be perpendicular to the x-axis. The simulation box can be divided into N_s slabs, each with width L/N_s. The average density of oil particles in each slab is measured. Using the Gibbs' definition, the oil-water interface sits at the height where the oil density is half-way between the densities of oil in the two bulk phases that the interface divides. The bulk densities correspond to the edge points of the coexistence region on the $T - x_{\text{oil}}$ phase diagram at the given temperature.

A.4 Optimizing CPU time

A.4.1 Changes in energy

Since acceptance criteria of Monte Carlo simulations only require changes in the system total energy, ΔE, rather than the absolute total energy, we can cut down computing time by only looking at how the proposed coordinate changes affect the total energy of the

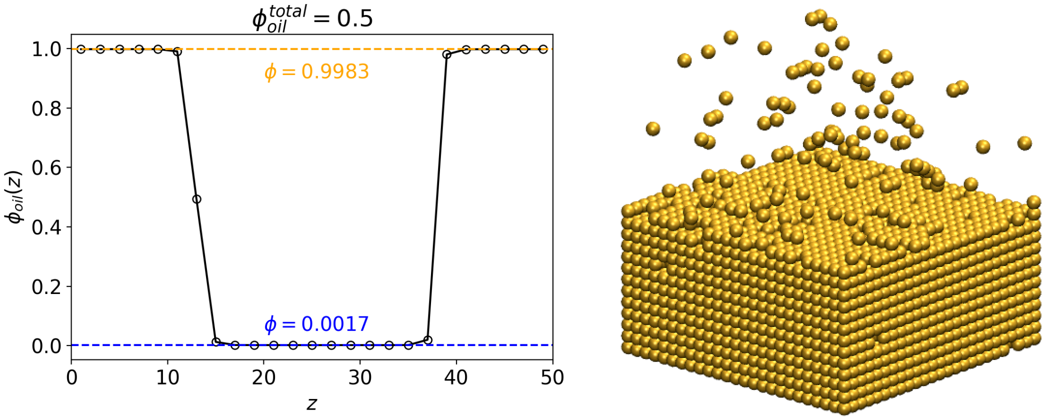

Figure A.1: When $T < T_c$, the densities of the oil-rich and water-rich phases give the edge of the coexistence region. Here I used a 50/50 mixture of oil and water in a $L = 50$ periodic box. The overall volume fraction of oil is 0.5, with the oil-rich phase $\phi_{\text{oil}} = 0.9983$ and the water-rich phase $\phi_{\text{oil}} = 0.0017$. This in turn means that in a $L = 50$ system, $N_{\text{oil}} = 0.0017 \times 50^3/8 = 27$ is the minimal number of oil blocks equired for droplets to form.

system, instead of computing the total energy across all coordinates twice to get the difference.

To make this clear, consider a proposed move that involves only n coordinates within a N-coordinate system, where $n \leq N$ is a subset of all the coordinates. I will denote the old coordinates as $\{r_1, r_2, ..., r_n, r_{n+1}, ..., r_N\}$ and the new coordinates $\{\bar{r}_1, \bar{r}_2, ..., \bar{r}_n, r_{n+1}, ..., r_N\}$. Assuming the total energy of the system can be decomposed into a sum of pair potentials among all coordinates,

$$\begin{aligned}
E_{\text{old}} &= \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} u(r_i, r_j) \\
&= \frac{1}{2}\left(\sum_{i=1}^{n} + \sum_{i=n+1}^{N}\right)\left(\sum_{j=1}^{n} + \sum_{j=n+1}^{N}\right) u(r_i, r_j) \\
&= \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} u(r_i, r_j) + \sum_{i=n+1}^{N}\sum_{j=1}^{n} u(r_i, r_j) + \frac{1}{2}\sum_{i=n+1}^{N}\sum_{j=n+1}^{N} u(r_i, r_j)
\end{aligned}$$

and the new total energy would be,

$$E_{\text{new}} = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} u(\bar{r}_i, \bar{r}_j) + \sum_{i=n+1}^{N}\sum_{j=1}^{n} u(r_i, \bar{r}_j) + \frac{1}{2}\sum_{i=n+1}^{N}\sum_{j=n+1}^{N} u(r_i, r_j)$$

from which we can see that the last term that accounts for the pair interactions among the unchanged coordinates remains the same after coordinate changes for the first n particles. Repeatedly calculating those interactions is therefore only a waste of time. Taking the difference between the two, we have

$$\begin{aligned}\Delta E &= E_{\text{new}} - E_{\text{old}} \\ &= \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left[u(\bar{r}_i, \bar{r}_j) - u(r_i, r_j)\right] + \sum_{i=n+1}^{N}\sum_{j=1}^{n}\left[u(r_i, \bar{r}_j) - u(r_i, r_j)\right]\end{aligned}$$

where the first term sums the $n(n-1)/2$ changes in the pair interactions among the first n particles, and the second term sums the $n(N-n)$ changes in the pair interactions between the first n particles and the rest of the system. So when a Monte Carlo move has been proposed, the calculation for energy changes should only include i) the pair potentials between moved coordinates, and ii) the pair potentials between moved coordinates and the rest of the system. In practice this can be implemented by a one-particle energy, $E_{1p}(i)$, that calculates the sum of the interactions between particle i and all $N-1$ other particles in the system (we can discard some of these depending on the cut-off distance), and a two-particle energy, $E_{2p}(i, j)$, that calculates the interaction between particle i and particle j. The energy that will contribute to ΔE is then

$$E_{\text{MC}} = \sum_{i=1}^{n} E_{1p}(i) - \frac{1}{2}\sum_{i=1}^{n}\sum_{j\neq i}^{n} E_{2p}(i, j)\,.$$

A.4.2 Ewald sum

The electrostatic interactions among charges are computationally costly to evaluate because the long-ranged Coulomb interactions extend way beyond the extent of periodic boxes, and our usual way of summing over pair interactions among all particles within the single periodic box at origin will fail to account for the interactions between these particles and those in the box's periodic images. The Ewald sum technique addresses this problem by splitting the electrostatic energy into a short-ranged part (summed in real space), a long-ranged part (summed in reciprocal space). The parameter that determines what is short-ranged and what is long-ranged is denoted $1/\sqrt{\alpha}$, a "screening" length.

For either sum, the energy will only be exact if we sum towards infinity which is not tractable computationally. But since the real space part is short-ranged, we can potentially cut off the range of interaction at $r = r_{\text{cut}}$ using methods like Verlet neighbor lists; on the other hand, since the reciprocal space part is long-ranged in real space and hence short-ranged in reciprocal space, we may cut off the reciprocal sum by some $k = k_{\text{cut}}$. In the following I discuss how the choices for all three can be made.

Patterns at intermediate scale

The interaction range cutoff is a tradeoff between computational efficiency and accurate observations: If we enforce a small cutoff by summing over fewer terms in the real and reciprocal space series, it is possible to greatly speed up the energy calculation. However, the small cutoff might also be too small for our energy calculations to be accurate, because the discarded terms still contribute significantly to the sum. The inaccurate energies will then cause our observations of the system's configurations to be inaccurate, too.

Of course, cutting off the energy calculation will always result in inaccurate configurations to some extent. So what kind of inaccruacy is acceptable, and what is not? Before deciding on a cutoff distance or wavenumber, it is important to first understand the scale of the phenomena we are interested in observing. If the cut-off gives inaccuracies on scales that are either far smaller or far greater than the scale of our observations, then we can accept such inaccuracies and enforce the cutoff.

For the charge-frustrated Ising model, we can find this relevant scale by examining the hamiltonian in fourier space:

$$\mathcal{H}\left[\{s_i\}_{i=1}^N\right] = -J\sum_{\langle i,j\rangle} s_i s_j + \frac{1}{2}\sum_i\sum_{j\neq i}\frac{s_i s_j Q_{ij}}{|\mathbf{r}_i - \mathbf{r}_j|}$$

where $\langle i, j\rangle$ means summation over all nearest-neighbor pairs of spins in the system, and $Q_{ij} \in \mathcal{R}^{N\times N}$ contains information about the size of charges,

$$Q_{ij} = \begin{cases} z^2/n_A^2 & \text{, if } s_i = s_j = 1; \\ z^2/n_A n_B & \text{, if } s_i = 1, s_j = -1; \\ z^2/n_B^2 & \text{, if } s_i = s_j = -1. \end{cases}$$

For simplicity we will assume $n_A = n_B = n$ for now; and $Q_{ij} = Q = 1/n^2$ is a constant scalar. The spin coupling and charge interactions can be summarized into one single interaction potential $u(r)$,

$$\mathcal{H}\left[\{s_i\}_{i=1}^N\right] = \frac{1}{2}\sum_i\sum_{j\neq i} s_i s_j u(\mathbf{r}_i - \mathbf{r}_j)$$

Now we can define a pair of Fourier transforms $\hat{s}_{\mathbf{k}} = \sum_j s_j e^{-i\mathbf{k}.\mathbf{r}_j}$ and $s_j = \frac{1}{N}\sum_{\mathbf{k}} \hat{s}_{\mathbf{k}} e^{i\mathbf{k}.\mathbf{r}_j}$, and their Delta-function $\delta_{mn} = \frac{1}{N}\sum_{\mathbf{k}} e^{i\mathbf{k}.(\mathbf{r}_m - \mathbf{r}_n)}$, so that the hamiltonian can be written as a sum in $\mathbf{k}$-space rather than $\mathbf{r}$ space:

$$\begin{aligned}
\mathcal{H}\left[\{\hat{s}_{\mathbf{k}}\}_{k_x,k_y,k_z=-\pi/a}^{\pi/a}\right] &= \frac{1}{2}\sum_i s_i\left[\sum_{j\neq i} s_j u(\mathbf{r}_i-\mathbf{r}_j)\right] \\
&= \frac{1}{2}\sum_i s_i\left[\frac{1}{N}\sum_{\mathbf{k}} \hat{s}_{\mathbf{k}}\hat{u}_{\mathbf{k}} e^{i\mathbf{k}.\mathbf{r}_i}\right] \\
&= \frac{1}{2N}\sum_{\mathbf{k}} \hat{s}_{\mathbf{k}}\hat{u}_{\mathbf{k}}\left[\sum_i s_i e^{i\mathbf{k}.\mathbf{r}_i}\right] \\
&= \frac{1}{2N}\sum_{\mathbf{k}} \hat{s}_{\mathbf{k}}\hat{u}_{\mathbf{k}}\hat{s}_{-\mathbf{k}} \\
&= \frac{1}{2N}\sum_{\mathbf{k}} \left|\hat{s}_{\mathbf{k}}\right|^2 \hat{u}_{\mathbf{k}}
\end{aligned}$$

where on the second line I applied the discrete version of convolution theorem. Note $u(\mathbf{r}) = -J\delta_{|\mathbf{r}|,a} + \frac{Q_{ij}}{|\mathbf{r}|}$ has the fourier transform,

$$\begin{aligned}
\hat{u}_{\mathbf{k}} &= a^3\sum_{j=1}^{N} u(\mathbf{r}_j)e^{-i\mathbf{k}.\mathbf{r}_j} \\
&= -Ja^3\sum_{j=1}^{N}\delta_{|\mathbf{r}_j|,a}e^{-i\mathbf{k}.\mathbf{r}_j} + a^3\sum_{j=1}^{N}\frac{Q}{|\mathbf{r}_j|}e^{-i\mathbf{k}.\mathbf{r}_j} \\
&= -Ja^3\sum_{m=x,y,z}\left(e^{ik_m a}+e^{-ik_m a}\right)+\frac{4\pi Q}{k^2} \\
&= -2Ja^3\sum_{m=x,y,z}\cos\left(k_m a\right)+\frac{4\pi Q}{k^2} \\
&\approxeq -2Ja^3\sum_{m=x,y,z}\left(1-\frac{1}{2}k_m^2a^2\right)+\frac{4\pi Q}{k^2} \\
&= -2Ja^3\left(3-\frac{1}{2}k^2a^2\right)+\frac{4\pi Q}{k^2} \\
&= Jk^2a^5+\frac{4\pi Q}{k^2}+\text{constants}
\end{aligned}$$

These two results let us understand how the pair-potential influences the system's structures and dynamics in reciprocal space. The first term in $\hat{u}_{\mathbf{k}}$ is the square gradient Ising coupling that rises with increasing wavenumber $\mathbf{k}$ (short-ranged and favoring patterns at large scales), whereas the second term is the surfactant-mimicking electrostatics that decreases with increasing $\mathbf{k}$ (long-ranged and favoring patterns at small scales). The net effect due to these two terms with opposite scaling trends with $\mathbf{k}$ is that a minimum of $\hat{u}_{\mathbf{k}}$, $|\mathbf{k}| = k^*$, exists at an intermediate scale. This also means that in real space, the corresponding wavelength $\lambda^* = 2\pi/k^*$ should encompass the interesting assembly phenomena that is otherwise not observed in a pure Ising or pure electrostatics system. Consequently, this is the scale of observation that we should take care to preserve when using cutoffs to improve computational effciency. In this particular system that has equal number of head and tail groups per surfactant, we can work out the characteristic wavenumber k^* exactly according to the charge valence z, the surfactant density ρ, and head-tail separation Δ,

$$k^* = \left(\frac{4\pi z^2}{Ja^5 n^2}\right)^{1/4} = \left(\frac{4\pi}{Ja^5 n^2} \times \frac{3k_BT}{4\pi\rho\Delta^2}\right)^{1/4} = \left(\frac{3k_BT}{Ja^5 n^2 \rho\Delta^2}\right)^{1/4}$$

with a corresponding characteristic wavelength,

$$\lambda^* = \frac{2\pi}{k^*} = 2\pi\left(\frac{Ja^5 n^2 \rho\Delta^2}{3k_BT}\right)^{1/4} .$$

Given the summation increment $\Delta k = \pi/L$, the number of wavevectors needed to sum up to this k^* in each direction is then

$$n_{\mathbf{k}} = k^*/\Delta k = \frac{L}{\pi}\left(\frac{3k_BT}{Ja^5 n^2 \rho\Delta^2}\right)^{1/4} .$$

Truncating the real-space sum

Now we look into the methods that improve computational speed in relation to the characteristic scale. Fundamentally, in all these methods the computational efficiency is gained by modifying the way the system's total energy is calculated, often by approximating the true hamiltonian using a truncated hamiltonian to reduce the number of terms that needs to be iterated over.

In order to decide how to implement the Ewald sum correctly so that the observed phenomena at the characteristic scale remains true to the original hamiltonian, we need to see analytically what the modified, approximate hamiltonian is and make a direct comparison with its unmodified counterpart. As demonstrated in the previous section, our system's exact total electrostatic energy can be written as

$$\begin{aligned}
E_{\text{elec}} &= \frac{1}{2N}\sum_{\mathbf{k}} |\hat{s}_{\mathbf{k}}|^2\, \hat{u}_{\mathbf{k}}^{\text{elec}} \\
&= \frac{1}{2N}\sum_{\mathbf{k}} \frac{4\pi Q}{k^2} |\hat{s}_{\mathbf{k}}|^2 \left[e^{-k^2/4\alpha} + \left(1 - e^{-k^2/4\alpha}\right)\right] \\
&= \frac{1}{2N}\sum_{\mathbf{k}} \frac{4\pi Q}{k^2} |\hat{s}_{\mathbf{k}}|^2\, e^{-k^2/4\alpha} + \frac{1}{2N}\sum_{\mathbf{k}} \frac{4\pi Q}{k^2} |\hat{s}_{\mathbf{k}}|^2 \left(1 - e^{-k^2/4\alpha}\right)
\end{aligned}$$

where the first term is a fast-decaying gaussian centered around $k = 0$, so it should be summed in reciprocal space; while the second term is slow-decaying in reciprocal space, but that in turn means it is fast-decaying in real space, so it is transformed back into real space,

$$\begin{aligned} E_{\text{elec}} &= E_{\text{reciprocal}} + E_{\text{real}} \\ &= \frac{1}{2N}\sum_{\mathbf{k}} \frac{4\pi Q}{k^2} \left|\hat{s}_{\mathbf{k}}\right|^2 e^{-k^2/4\alpha} + \frac{1}{2}\sum_i \sum_{j\neq i} \frac{q_i q_j \operatorname{erfc}(\sqrt{\alpha} r_{ij})}{r_{ij}} \,. \end{aligned}$$

Before going into more details about the first term, let's look at the second term. The complementary error function has a decay length of $1/\sqrt{\alpha}$, which means if we want to truncate the real-space sum to only include the charges that are in the periodic box at the origin, we should choose $1/\sqrt{\alpha} \ll L/2$, where $L/2$ is the largest possible distance without crossing into neighboring periodic box images. Rearranging this expression, we see that $\alpha \gg 4/L^2$ is the lower limit for choosing α if we want to truncate the real-space sum at one periodic box, $r_{\text{cut}} \leq L/2$.

However, in my experience the real-space sum is generally cheaper to calculate by implementing list of charges and neighbor lists, so this criterion can be relaxed if choosing an even larger screening length (smaller α) help speed up the reciprocal sum. We will just have to extend the sum into neighboring periodic box images.

In any case, r_{cut} should not need to be more than a few multiples of $1/\sqrt{\alpha}$, the decay length of the real-space pair potential, because beyond that the complementary error function essentially evaluates to zero.

In the source code I combine the real-space sum with other real-space interactions, such as the Ising near-neighbor coupling.

Truncating the reciprocal-space sum

Next, we return to the reciprocal sum which can be further split into a small-k series and a large-k series, preparing it for truncation,

$$\begin{aligned} E_{\text{reciprocal}} &= \frac{1}{2N}\sum_{\mathbf{k}} \frac{4\pi Q}{k^2} \left|\hat{s}_{\mathbf{k}}\right|^2 e^{-k^2/4\alpha} \\ &= \frac{1}{2N}\left(\sum_{|\mathbf{k}|<k_{\text{cut}}} + \sum_{|\mathbf{k}|\geq k_{\text{cut}}}\right) \frac{4\pi Q}{k^2} \left|\hat{s}_{\mathbf{k}}\right|^2 e^{-k^2/4\alpha} \end{aligned}$$

Combining this with the real-space sum we see the total electrostatic energy accounted for by the Ewald sum, and a large-k reciprocal part that is being tossed out,

$$\begin{aligned}
E_{\text{elec}} &= \frac{1}{2N}\left(\sum_{|\mathbf{k}|<k_{\text{cut}}} + \sum_{|\mathbf{k}|\geq k_{\text{cut}}}\right)\frac{4\pi Q}{k^2}\left|\hat{s}_{\mathbf{k}}\right|^2 e^{-k^2/4\alpha} + \frac{1}{2}\sum_i\sum_{j\neq i}\frac{Q\text{erfc}(\sqrt{\alpha}r_{ij})}{r_{ij}} \\
&= \left\{\frac{1}{2N}\sum_{|\mathbf{k}|<k_{\text{cut}}}\frac{4\pi Q}{k^2}\left|\hat{s}_{\mathbf{k}}\right|^2 e^{-k^2/4\alpha} + \frac{1}{2}\sum_i\sum_{j\neq i}\frac{Q\text{erfc}(\sqrt{\alpha}r_{ij})}{r_{ij}}\right\} + \frac{1}{2N}\sum_{|\mathbf{k}|\geq k_{\text{cut}}}\frac{4\pi Q}{k^2}\left|\hat{s}_{\mathbf{k}}\right|^2 e^{-k^2/4\alpha} \\
&= E_{\text{ewald}} + \frac{1}{2N}\sum_{|\mathbf{k}|\geq k_{\text{cut}}}\frac{4\pi Q}{k^2}\left|\hat{s}_{\mathbf{k}}\right|^2 e^{-k^2/4\alpha}
\end{aligned}$$

which can then be rearranged and decomposed into the electrostatic pair potential $\hat{u}^{\text{elec}}_{\mathbf{k}} = \frac{4\pi Q}{k^2}$,

$$\begin{aligned}
E_{\text{ewald}} &= E_{\text{elec}} - \frac{1}{2N}\sum_{|\mathbf{k}|\geq k_{\text{cut}}}\frac{4\pi Q}{k^2}\left|\hat{s}_{\mathbf{k}}\right|^2 e^{-k^2/4\alpha} \\
&= \frac{1}{2N}\sum_{\mathbf{k}}\left|\hat{s}_{\mathbf{k}}\right|^2\hat{u}^{\text{elec}}_{\mathbf{k}} - \frac{1}{2N}\sum_{|\mathbf{k}|\geq k_{\text{cut}}}\left|\hat{s}_{\mathbf{k}}\right|^2\hat{u}^{\text{elec}}_{\mathbf{k}}e^{-k^2/4\alpha} \\
&= \frac{1}{2N}\sum_{|\mathbf{k}|<k_{\text{cut}}}\left|\hat{s}_{\mathbf{k}}\right|^2\hat{u}^{\text{elec}}_{\mathbf{k}} + \frac{1}{2N}\sum_{|\mathbf{k}|\geq k_{\text{cut}}}\left|\hat{s}_{\mathbf{k}}\right|^2\hat{u}^{\text{elec}}_{\mathbf{k}}\left(1-e^{-k^2/4\alpha}\right) \\
&= \frac{1}{2N}\sum_{\mathbf{k}}\left|\hat{s}_{\mathbf{k}}\right|^2\tilde{\hat{u}}^{\text{elec}}_{\mathbf{k}}
\end{aligned}$$

where on the last line I summarized the terms using an Ewald pair potential,

$$\tilde{\hat{u}}^{\text{elec}}_{\mathbf{k}} = \begin{cases}\hat{u}^{\text{elec}}_{\mathbf{k}} & \text{for } |\mathbf{k}| < k_{\text{cut}}; \\ \hat{u}^{\text{elec}}_{\mathbf{k}}\left(1-e^{-k^2/4\alpha}\right) & \text{for } |\mathbf{k}| \geq k_{\text{cut}}\,.\end{cases}$$

and the error of the Ewald sum can be written as,

$$\Delta E = E_{\text{elec}} - E_{\text{ewald}} = \frac{1}{2N}\sum_{|\mathbf{k}|\geq k_{\text{cut}}}\left(\hat{u}^{\text{elec}}_{\mathbf{k}} - \tilde{\hat{u}}^{\text{elec}}_{\mathbf{k}}\right)\left|\hat{s}_{\mathbf{k}}\right|^2 = \frac{1}{2N}\sum_{|\mathbf{k}|\geq k_{\text{cut}}}\hat{u}^{\text{elec}}_{\mathbf{k}}\left|\hat{s}_{\mathbf{k}}\right|^2 e^{-k^2/4\alpha}$$

This rearrangement shows us how the choice of k_{cut} influences the accuracy of the Ewald sum at different scales.When $|\mathbf{k}| < k_{\text{cut}}$, the full electrostatic pair potential $\hat{u}^{\text{elec}}_{\mathbf{k}}$ is included into the sum, so with all else done correctly the configurations at wavenumbers smaller than k_{cut} are expected to be accurate. However at $|\mathbf{k}| \geq k_{\text{cut}}$, such accuracy is no longer guaranteed as $\hat{u}^{\text{elec}}_{\mathbf{k}}$ is now discounted by a factor of $\left(1-e^{-k^2/4\alpha}\right)$. The discrepancy scales with $e^{-k^2/4\alpha}$, which is the greatest at $|\mathbf{k}| = k_{\text{cut}}$ where $\tilde{\hat{u}}^{\text{elec}}_{\mathbf{k}}$ has a discontinuous step down from $\hat{u}^{\text{elec}}_{\mathbf{k}}$. After that, the discrepancy decays to 0 at $|\mathbf{k}| \to \infty$, with a rate $r \propto 1/\sqrt{\alpha}$ that scales inversely with the choice of α.

In essense, the accuracy of the reciprocal sum is controlled by both α and k_{cut}. A smaller α gives a smaller discontinuity at $|\mathbf{k}| = k_{\text{cut}}$ and faster decay of the error introduced by the k-series truncation. This contrasts with the need for larger α values from our discussion of the real-space sum truncation.

Comparing k_{cut} and k^*

Let's consider a gradual relaxation of our k_{cut} choice and what each case implies for the choice of α and r_{cut}.

If we were given infinite computational resources, then we are free to choose k_{cut} to be much greater than k^*, which in turn means that we can afford to have any choice α value without introducing any discrepancy between $\hat{u}_{\mathbf{k}}^{\text{elec}}$ and $\tilde{\hat{u}}_{\mathbf{k}}^{\text{elec}}$ across a large range of $\mathbf{k}$ values. A larger α means $r_{\text{cut}} \propto 1/\sqrt{\alpha}$ can be very small. Of course, including many terms in the k-series is expensive, and maintaining accuracy beyond $|\mathbf{k}| \geq k_{\text{cut}}$ is unnecessary.

If we choose a smaller k_{cut} close to but greater than k^*, the step down of $\hat{u}_{\mathbf{k}}^{\text{elec}} e^{-k^2/4\alpha}$ will happen beyond k^*, so no noticeable change at the characteristic length is expected. From the perspective of the reciprocal sum, we've saved some time, and we don't care about the size of α because it doesn't change the behavior at k^* whatsoever. Again in this case we should be free to pick a large α to speed up the real-space sum.

Up to this point we have looked at two very conservative choices of k_{cut}, but since the reciprocal sum is much more expensive than the real sum, it is worth asking whether we can cut the number of wavevectors even more. The answer is yes as long as we choose α more carefully, not only increasing α to reduce the number of real space neighbors, but also balancing that with the widening gap between $\hat{u}_{\mathbf{k}}^{\text{elec}}$ and $\tilde{\hat{u}}_{\mathbf{k}}^{\text{elec}}$ as α increases.

The α of our choice must not be so large that the error at $|\mathbf{k}| = k^*$ is much larger than $a^3 k_B T$; but it also must not be so small that the real-space sum becomes prohibitively expensive. This is again contrasts with the conventional way of converging Ewald sums, which calls for a balance between small k_{cut} and small α that maintains a converged energy $E_{\text{ewald}} = E_{\text{elec}}$. But our focus with α and k_{cut} is different from the conventional choice, because we care more about the value of the error gap at the a specific contribution to the total energy - the characteristic scale - than the error of the total energy ΔE. To reiterate, errors at wavenumbers far away from the characteristic k^* should not give arise to inaccurate configurations at the scale of our interest, so even if those errors are large they should not be of our concern.

As an example, let's evaluate this error for the case of equal and opposite head and tail charges:

$$\begin{aligned}
\left(\hat{u}^{\text{elec}}_{\mathbf{k}} - \tilde{\hat{u}}^{\text{elec}}_{\mathbf{k}}\right)\Big|_{|\mathbf{k}|=k^*} &= \frac{4\pi Q}{k^{*2}} e^{-k^{*2}/4\alpha} \\
&= 4\pi \frac{z^2}{n^2} \left(\frac{Ja^5n^2}{4\pi z^2}\right)^{1/2} \exp\left[-\left(\frac{4\pi z^2}{Ja^5n^2}\right)^{1/2} /4\alpha\right] \\
&= \left(\frac{4\pi Ja^5z^2}{n^2}\right)^{1/2} \exp\left[-\left(\frac{\pi z^2}{4J\alpha^2a^5n^2}\right)^{1/2}\right] \\
&= \frac{a^{5/2}z\sqrt{4\pi J}}{n} \exp\left(-\sqrt{\frac{\pi}{J}} \times \frac{z}{2\alpha a^{5/2}n}\right)
\end{aligned}$$

Previously I have been using values $\alpha = 1$, $z = 0.3$, $J = 1$, $L = 12$ and $n = a = 1$, which yields an error of $0.815 \times a^3k_BT$ and $k^* = 1.03$. This does not meet the small error criteria, so if I choose to include $n_{\mathbf{k}} = k^*/\Delta k = 3.94 \approx 4$ or fewer wavevectors, the configurations at the characteristic length scale will be inaccurate and different from what we should see based on the full hamiltonian.

In practice, the pair potentials $\tilde{\hat{u}}^{\text{elec}}_{\mathbf{k}}$ and $\hat{u}^{\text{elec}}_{\mathbf{k}}$ can be found analytically, from which we can then find an optimal combination of $\{k_{\text{cut}}, \alpha\}$ by surveying the computation times of a number of combinations that satisfy the requirement $\left(\hat{u}^{\text{elec}}_{\mathbf{k}} - \tilde{\hat{u}}^{\text{elec}}_{\mathbf{k}}\right)\Big|_{|\mathbf{k}|=k^*} \ll a^3k_BT$. Among these combinations, we should pick the least computationally intensive one. Naturally this choice is highly dependent on $\tau_{\text{reciprocal}}/\tau_{\text{real}}$, the ratio of evaluation time for a reciprocal term and that for a real term, and hence highly dependent on how the reciprocal and real term calculations have been optimized, for example by using a cosine look-up table as described in the next Section. Finally the real-space cutoff r_{cut} is easily determined as some multiple of $1/\sqrt{\alpha}$.

Changes in structural factors

Calculating the reciprocal sum requires us to compute structural factors of the screened charges, $\{\hat{\rho}_{\mathbf{k}}\}$ with $k_x = 2\pi n_x/L$, $n_x = -L/2, ..., -1, 0, 1, ..., L/2$. It is related to the charge densities $\{\rho_{\mathbf{r}}\}$ by

$$\hat{\rho}_{\mathbf{k}} = \sum_{m=1}^{N_q} \rho_{\mathbf{r}_m} e^{-i\mathbf{k}.\mathbf{r}_m} = \sum_{m=1}^{N_q} \left(\sum_{n=1}^{N_q} q_n \delta_{\mathbf{r}_m,\mathbf{r}_n}\right) e^{-i\mathbf{k}.\mathbf{r}_m} = \sum_{m=1}^{N_q} q_m \exp\left[-i\left(k_x x_m + k_y y_m + k_z z_m\right)\right]$$

and its modulus

$$
\begin{aligned}
\left|\hat{\rho}_{\mathbf{k}}\right|^2 &= \sum_{m=1}^{N_q}\sum_{n=1}^{N_q} q_m q_n \exp\left[i\left(k_x x_n + k_y y_n + k_z z_n\right) - i\left(k_x x_m + k_y y_m + k_z z_m\right)\right] \\
&= \sum_{m=1}^{N_q}\sum_{n=1}^{N_q} q_m q_n e^{ik_x(x_n - x_m)} e^{ik_y(y_n - y_m)} e^{ik_z(z_n - z_m)} \\
&= \sum_m q_m^2 + \left(\sum_{m>n} + \sum_{n>m}\right) q_m q_n e^{ik_x(x_n - x_m)} e^{ik_y(y_n - y_m)} e^{ik_z(z_n - z_m)} \\
&= \sum_m q_m^2 + \sum_{m>n} q_m q_n \left[e^{ik_x(x_n - x_m)} e^{ik_y(y_n - y_m)} e^{ik_z(z_n - z_m)} + e^{-ik_x(x_n - x_m)} e^{-ik_y(y_n - y_m)} e^{-ik_z(z_n - z_m)}\right] \\
&= \sum_{m=1}^{N_q} q_m^2 + 2\sum_{n=1}^{N_q}\sum_{m>n}^{N_q} q_m q_n \cos\left[k_x\left(x_n - x_m\right) + k_y\left(y_n - y_m\right) + k_z\left(z_n - z_m\right)\right]
\end{aligned}
$$

which contributes to the reciprocal sum as

$$
E_{\text{reciprocal}} = \frac{1}{2N}\sum_{\mathbf{k}\neq 0}\frac{4\pi}{k^2}\left|\hat{\rho}_{\mathbf{k}}\right|^2 e^{-k^2/4\alpha}
$$

Note the $\{k_x, k_y, k_z\} = \{0, 0, 0\}$ term is excluded (Frenkel & Smit page298). Every time a Monte Carlo move is proposed, we make a change to the charge densities $\{\rho_{\mathbf{r}}\}$ by moving the first q out of N_q charges, so that the coordinates changes from $\left\{\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_p, \mathbf{r}_{p+1}, ..., \mathbf{r}_{N_q}\right\}$ to $\left\{\bar{\mathbf{r}}_1, \bar{\mathbf{r}}_2, ..., \bar{\mathbf{r}}_p, \mathbf{r}_{p+1} ..., \mathbf{r}_{N_q}\right\}$. Similar to what is discussed in Section A.4.1, we don't need to fully re-calculate the modulus of the structural factor $\left|\hat{\rho}_{\mathbf{k}}\right|^2$ after every move. Because the summation over $\mathbf{k}$ is a linear operation, we can speed up the computation by looking at only the change to the structural factor modulus at each wavenumber $\mathbf{k}$. Untangling the sums and breaking them into changed and unchanged parts:

$$\begin{aligned}
|\hat{\rho}_{\mathbf{k}}|^2_{\text{old}} &= 2\sum_{n=1}^{N_q}\sum_{m>n}^{N_q} q_m q_n \cos\left[\mathbf{k}.(\mathbf{r}_m - \mathbf{r}_n)\right] \\
&= 2\left(\sum_{n=1}^{p} + \sum_{n=p+1}^{N_q}\right)\sum_{m=n+1}^{N_q} q_m q_n \cos\left[\mathbf{k}.(\mathbf{r}_m - \mathbf{r}_n)\right] \\
&= 2\left(\sum_{n=1}^{p}\sum_{m=n+1}^{N_q} + \sum_{n=p+1}^{N_q}\sum_{m=n+1}^{N_q}\right) q_m q_n \cos\left[\mathbf{k}.(\mathbf{r}_m - \mathbf{r}_n)\right] \\
&= 2\sum_{n=1}^{p}\sum_{m=n+1}^{p} q_m q_n \cos\left[\mathbf{k}.(\mathbf{r}_m - \mathbf{r}_n)\right] \\
&\qquad ... + 2\sum_{n=1}^{p}\sum_{m=p+1}^{N_q} q_m q_n \cos\left[\mathbf{k}.(\mathbf{r}_m - \mathbf{r}_n)\right] \\
&\qquad ... + 2\sum_{n=p+1}^{N_q}\sum_{m=n+1}^{N_q} q_m q_n \cos\left[\mathbf{k}.(\mathbf{r}_m - \mathbf{r}_n)\right]
\end{aligned}$$

and

$$\begin{aligned}
|\hat{\rho}_{\mathbf{k}}|^2_{\text{new}} &= 2\sum_{n=1}^{p}\sum_{m=n+1}^{p} q_m q_n \cos\left[\mathbf{k}.(\bar{\mathbf{r}}_m - \bar{\mathbf{r}}_n)\right] \\
&\qquad ... + 2\sum_{n=1}^{p}\sum_{m=p+1}^{N_q} q_m q_n \cos\left[\mathbf{k}.(\mathbf{r}_m - \bar{\mathbf{r}}_n)\right] \\
&\qquad ... + 2\sum_{n=p+1}^{N_q}\sum_{m=n+1}^{N_q} q_m q_n \cos\left[\mathbf{k}.(\mathbf{r}_m - \mathbf{r}_n)\right]
\end{aligned}$$

and the change is

$$\begin{aligned}
\Delta\,|\hat{\rho}_{\mathbf{k}}|^2 &= |\hat{\rho}_{\mathbf{k}}|^2_{\text{new}} - |\hat{\rho}_{\mathbf{k}}|^2_{\text{old}} \\
&= 2\sum_{n=1}^{p}\sum_{m=n+1}^{p} q_m q_n \left\{\cos\left[\mathbf{k}.(\bar{\mathbf{r}}_m - \bar{\mathbf{r}}_n)\right] - \cos\left[\mathbf{k}.(\mathbf{r}_m - \mathbf{r}_n)\right]\right\} \\
&\qquad ... + 2\sum_{n=1}^{p}\sum_{m=p+1}^{N_q} q_m q_n \left\{\cos\left[\mathbf{k}.(\mathbf{r}_m - \bar{\mathbf{r}}_n)\right] - \cos\left[\mathbf{k}.(\mathbf{r}_m - \mathbf{r}_n)\right]\right\}
\end{aligned}$$

where the first double sum is across all pairs among the charges that have been moved, and the second double sum is over all pairs of one moved charge and one un-moved charge.

In practice, computing the change in the structural factor involves repeated evaluation of cosines, which can be computatinally expensive, and in fact unnecessary for a lattice where the number of possible arguments for the cosine functions are finite. This will be the topic of the next section, where I explain how we can construct a pre-computed cosine look-up table to speed up this step.

Regarding data structures of the implementation, we can think of the change in the structural factor modulus as a matrix, $\Delta \left|\hat{\rho}_{\mathbf{k}}\right|^2 = \mathbf{S} \in \mathcal{R}^{(L+1)\times(L+1)\times(L+1)}$, where the central $\mathbf{k} = 0$ element can just be set to zero for easy evaluation. Once $\Delta \left|\hat{\rho}_{\mathbf{k}}\right|^2$ has been evaluated, the change in reciprocal energy is then easily computed by performing a $\mathbf{k}$-space sum,

$$\Delta E_{\text{reciprocal}} = \frac{1}{2N} \sum_{\mathbf{k}\neq 0}^{|\mathbf{k}|<k_{\text{cut}}} \frac{4\pi}{k^2} \Delta \left|\hat{\rho}_{\mathbf{k}}\right|^2 e^{-k^2/4\alpha} \,.$$

Neighbor lists

The spin interactions are nearest-neighbor only, so a suitable neighbor list for that is simply a $L \times L \times L$ array that contains the particle id numbers, indexed by their coordinates.

The real space part of the Ewald sum can go beyond the nearest neighbor, but due to the nature of its length-scale separation, the real space pair potential will never extend beyond a few lattice cells. This means we should not blindly evaluate this pair potential throughout the size of the system, but rather impose a cut-off distance, r_{cut}, that depends on the Ewald sum parameter α,

$$r_{\text{cut}} \propto \frac{1}{\sqrt{\alpha}} \,.$$