



SAPIENZA UNIVERSITÀ DI ROMA  
DOTTORATO DI RICERCA IN FISICA  
SCUOLA DI DOTTORATO “VITO VOLTERRA”

# **Self-assembly driven formation of nematic liquid crystals: theory and simulations**

Thesis submitted to obtain the degree of  
"Dottore di Ricerca" - Doctor of Philosophy in Physics  
October 2014

**Thesis advisors**  
Prof. Francesco Sciortino  
Dr. Cristiano De Michele

**Candidate**  
Thuy K. Nguyen

# Contents

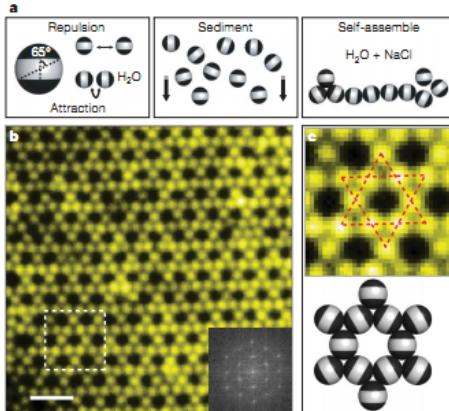
<b>Contents</b>	<b>2</b>
<b>Introduction</b>	<b>3</b>
<b>1 Backgrounds</b>	<b>7</b>
1.1 Liquid crystalline phases. . . . .	7
1.2 DNA duplexes . . . . .	11
<b>2 Theories</b>	<b>14</b>
2.1 Onsager theory for isotropic-nematic phase transition in systems of long hard rods . . . . .	14
2.2 Wertheim theory for isotropic self-assembling systems . . . . .	17
2.3 De Michele et. al. theory for isotropic-nematic phase transition in self-assembly system . . . . .	18
2.3.1 Isotropic phase . . . . .	21
2.3.2 Nematic phase . . . . .	22
2.3.3 Phase coexistence . . . . .	23
<b>3 Simulation techniques</b>	<b>24</b>
3.1 Basic Monte Carlo techniques . . . . .	24
3.1.1 A brief introduction to Monte Carlo method and Metropolis sampling . . . . .	24
3.1.2 Monte Carlo simulation in canonical ensemble (NVT) . . . . .	26
3.1.3 Monte Carlo simulation in isobaric-isothermal ensemble (NPT)	27
3.1.4 Monte Carlo simulation in grand canonical ensemble (GCMC)	28
3.1.5 Periodic boundary condition and reduced units in simulation	29
3.2 Scheme to compute phase coexistence lines . . . . .	29
3.2.1 Determine coexistence packing fraction using successive umbrella sampling (SUS) . . . . .	29
3.2.2 Kofke integration . . . . .	33
3.2.3 Reducing finite size effect . . . . .	37
3.3 Some other details of the simulations . . . . .	37
3.3.1 Linked list . . . . .	37
3.3.2 Creating an initial configuration at high packing fraction . . . . .	38
3.4 Computation of parameters required by theory . . . . .	38
3.4.1 Bonding volume . . . . .	38
3.4.2 Excluded volume and parameters . . . . .	39

3.4.3	Persistence length . . . . .	40
<b>4</b>	<b>Simulation results and comparison to theory</b>	<b>41</b>
4.1	Hard cylinders with 2 spherical patches at bases . . . . .	41
4.1.1	Model . . . . .	41
4.1.2	Checking overlap between 2 hard cylinders . . . . .	42
4.1.3	Computed excluded volume, persistence length and bond-ing volume . . . . .	44
4.1.4	Simulation results . . . . .	47
4.2	Bent cylinder . . . . .	54
4.2.1	Model . . . . .	54
4.2.2	Simulation results of isotropic-nematic phase transition . . .	57
4.3	Spheres . . . . .	61
4.3.1	Model . . . . .	61
4.3.2	Computed excluded volume, persistence length and bond-ing volume . . . . .	62
4.3.3	Simulation results . . . . .	65
<b>Conclusion</b>		<b>73</b>
<b>Bibliography</b>		<b>79</b>
<b>List of publications</b>		<b>80</b>

# Introduction

Self-assembly is defined as the spontaneous formation of a complex structure starting by elementary components (monomers). The resulting structure is encoded in the interaction potential between the monomers. When this process is reversible, it is also called supra-molecular polymerization or equilibrium polymerization. Reversible aggregation or equilibrium polymerization occurs when the energy scale of the monomer-monomer interaction is such that the thermal fluctuations can compete with bonding. In this case bonds between monomers can form and break and thermal equilibrium can be reached within the experimental observation time scale. Self-assembly is ubiquitous in nature and therefore has long been considered in life science as crucial to understand living structures. The self-assembly of amphipathic phospholipids into lipid membranes [1], the folding of polypeptide chains into proteins [2, 3], the formation of DNA double helix from two complimentary oligonucleotides, to name a few, are examples of self-assembly in nature. Even the term self-assembly derives from an appreciation of the capacity of viruses to spontaneously reconstitute themselves from their molecular components [4].

As self-assembly is a self-correction process based on reversible, non-covalent or weak covalent bonds, where the final assembly structure is coded in the shape and interactions of the component monomers, it has a great potential as a method to fabricate new materials from bottom-up with high accuracy [5]. In fact, it has been extensively exploited in molecular synthesis. However recently much attention has been given to self-assembly in colloidal systems, i.e. particles of size of order  $1\mu m - 1nm$  dispersed in a solution of much smaller particles, for many advantages they offer. First, while with molecular synthesis the building block are fixed, it is possible to synthesize colloids with desired shape, patch patterns and patch geometry to obtain the targeted assemblies of interest. For example recently Granick et. al. [6] have been successful in fabricating triblock Janus particles, i.e. spherical colloids with electrostatic repulsion in the middle, hydrophobic attraction at the poles. These colloids self-assemble into an open Kagome lattice, different from the closed-packed crystal usually expected from hard spheres (figure 1). Figure 2 illustrates some of the shapes and patch patterns of colloid particles that have been synthesised. Second, the large size of colloids makes it more convenient to follow the self-assembly process in experiments, and to employ visible light to investigate the systems. The third advantage, which is a consequence of the first one, is that colloidal systems also make it easier to study self-assembly theoretically and numerically using primitive models. Primitive models are very simple models, usually consist of only a hard core and a square-well potential, but they offer the possibility to check theoretical hypotheses against exact numerical data and lead to an understanding of the common, universal properties of self-assembly



**Figure 1 | Colloidal kagome lattice after equilibration.** **a**, Triblock Janus spheres hydrophobic on the poles (black, with an opening angle of 65°) and charged in the equator section (white), are allowed to sediment in deionized water. Then NaCl is added to screen electrostatic repulsion, allowing self-assembly by short-range hydrophobic attraction. **b**, Fluorescence image of a colloidal kagome lattice (main image) and its fast Fourier transform image (bottom right). Scale bar is 4  $\mu\text{m}$ . The top panel in **c** shows an enlarged view of the dashed white rectangle in **b**. Dotted red lines in **c** highlight two staggered triangles. The bottom panel in **c** shows a schematic illustration of particle orientations.

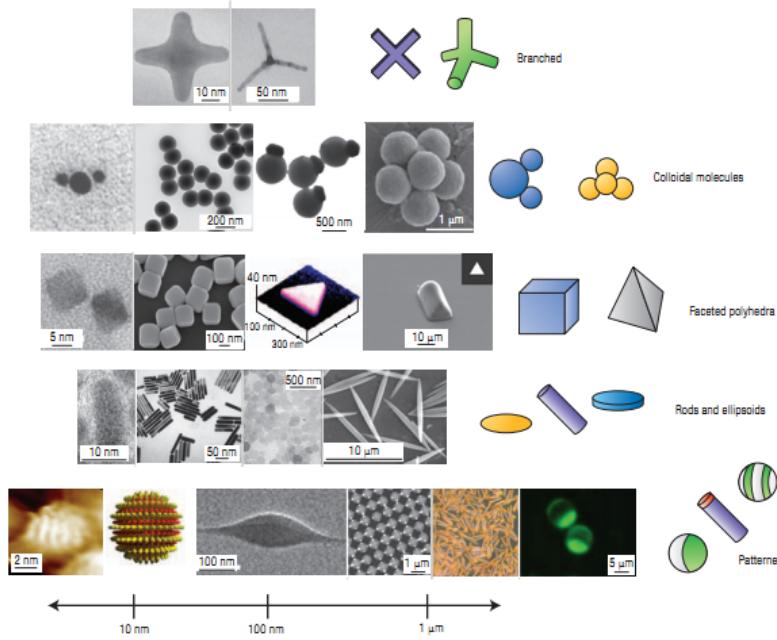
Figure 1: The triblock Janus particles and Kagome lattice produced by Granick et. al. Images and caption reproduced from Ref. [6]

processes behind the diversity of the geometry and nature of the building blocks and their patches [8].

An interesting case of self-assembly is when the interactions between monomers favour the formation of linear chains. If the chains are sufficiently rigid, nematic order may emerge in these systems. Examples of such systems include filamentous viruses such as tobacco mosaic virus, *fd* and *pf1* [9, 10, 11], protein and peptide fibers and fibrils [12, 13, 14], cylindric micelles [15], DNA duplexes consist of hundreds of base pairs [16, 17]. Self-assembly and nematic order can then couple and promote the formation of longer chains. In some cases, self-assembly can even induce nematic order in systems where the anisotropy of the monomers is not sufficient to form nematic order, as in short DNA duplexes with end-to-end stacking interaction [18] and hard spheres with short-ranged attractive, directional interaction [19].

Despite the relevance of the phenomenon, there have been very few studies focusing on isotropic-nematic phase transition in self-assembling systems. In particular a precise computation of the phase boundary in those system is still lacking. Among the most extensive works on the subject are the one by Lü and Kindt [19], which studies self-assembly of sticky hard spheres which form rather rigid chains using grand canonical ensemble simulations, and that by Kuriabova et. al. [20], which approximately determines the phase boundary of sticky hard cylinders by NPT simulations and which also observes higher liquid crystal orders, namely columnar liquid crystal, columnar crystal and cubatic-like phases.

Regarding theoretical study, on one hand, the isotropic-nematic phase transition in non-self-assembling long hard rods is well described by the classical theory developed by Onsager [21] and Wertheim theory works extremely well with isotropic self-assembling systems at low density [22]. On the other hand, although there has been much effort to study theoretically the coupling between self-assembly and isotropic-nematic phase transition [23, 24, 19, 20], the success is still somewhat moderate. As suggested by Ref. [24], a successful theory has to take into account both the polydispersity and flexibility of chains in self-assembling systems. In addition, proper approximation of steric interaction is also important as the nematic phase only forms in high packing fraction.



**Figure 1** Representative examples of recently synthesized anisotropic particle building blocks. The particles are classified in rows by anisotropy type and increase in size from left to right according to the approximate scale at the bottom. From left to right, top to bottom: branched particles include gold<sup>31</sup> and CdTe<sup>71</sup> tetrapods. DNA-linked gold nanocrystals<sup>32</sup> (the small and large nanocrystals are 5 nm and 10 nm respectively), silica dumb-bells<sup>33</sup>, asymmetric dimers<sup>34</sup> and fused clusters<sup>35</sup> form colloidal molecules. PbSe<sup>36</sup> and silver cubes<sup>10</sup> as well as gold<sup>37</sup> and polymer triangular prisms<sup>38</sup> are examples of faceted particles. Rods and ellipsoids of composition CdSe<sup>75</sup>, gold<sup>39</sup>, gibbsite<sup>4</sup> and polymer latex<sup>40</sup> are shown. Examples of patterned particles include striped spheres<sup>77</sup>, biphasic rods<sup>14</sup>, patchy spheres with ‘valence’<sup>34</sup>, Au–Pt nanorods<sup>78</sup> (the rod diameters are of the order of 200–300 nm) and Janus spheres<sup>33</sup>. Images reprinted with permission from the references as indicated. Copyright, as appropriate, AAAS, ACS, RSC, Wiley-VCH.

Figure 2: Some examples of the shapes and patch patterns of recently synthesized colloids. Images and caption reproduced from Ref. [7]

This thesis is another attempt to contribute to the understanding of the phenomenon. Our aim is to exploit the available techniques of simulation to establish a procedure to determine the isotropic-nematic phase boundary as well as coexistence properties of self-assembling systems with high precision and within reasonable time frame, overcoming the obstacle of metastability usually presents in “traditional” simulation methods like NPT ensemble or grand canonical ensemble simulations. The availability of such a method is important to provide the data needed to check against proposed, or to be proposed, theoretical frameworks on the coupling of nematic order and self-assembly. The outline of the thesis is as follows. First we will give a brief introduction on liquid crystal and DNA duplexes, which motivate our study and coarse-grained model. In chapter 2 we represent Onsager theory for isotropic-nematic transition in hard rods and Wertheim theory for isotropic phase of self-assembling system. These two theories serve as the basis for the theory proposed by De Michele et. al. [25] for isotropic-nematic phase transition in linear self-assembly systems. A great advantage of the theory is that it is parameter-free, i.e. all the parameters required for the theory can be computed by supplementary simulations instead of resorting to guesses. The theory predicts a behaviour of self-assembling systems at nematic coexistence which is quantitatively different from non self-assembling systems, namely the nematic packing fraction at phase boundary does not monotonically relates to the average aspect ratio of the chains. The

---

simulation techniques to be applied in this study are described in chapter 3. The simulation methods described in chapter 3 will be applied to different models to assess the results produced by the theory and validate the prediction, the details are reported in chapter 4. We have been successful in computing numerically, with high precision, the isotropic-nematic coexistent lines of those models over a large range of packing fraction, as well as properties of systems at phase boundary such as chain length distribution and average aspect ratio of chains, validating the predictions and assumptions made by the theory. We also studied the dependence of the phase boundary on the bending angles of short DNA dodecamers, modelling as hard bent cylinders.

# Chapter 1

## Backgrounds

### 1.1 Liquid crystalline phases.

All of us are well acquainted with 3 matter phases: gas, liquid and crystal. A system in gas and isotropic liquid phases possesses no long range order, whether it be positional or orientational. On the other hand, crystals exhibit 3D long range positional order.

However if the particles of a system possess enough anisotropy, the system can exhibit some intermediate order between isotropic liquid and crystal. These are liquid crystalline phases. Liquid crystalline phases exhibit many mechanical properties of liquid, but at the same time they have some optical properties of crystal, and hence the name.

There are many different liquid crystalline phases and the classification of them is usually based on their order and structure. Three major classes are nematic, smectic and columnar phases. Below is a brief description of the phases and some of their most common variants/subclasses.

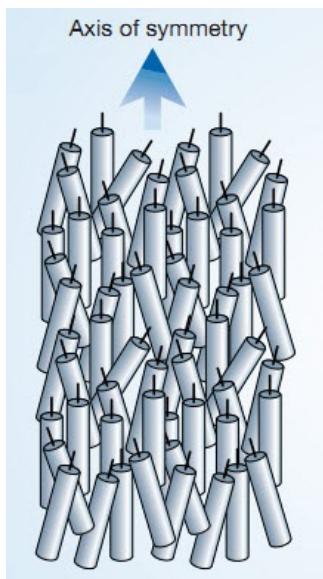
**Nematics and cholesterics** In a nematic phase, the system possesses no long-range positional order, but the particles tend to align along a common axis (figure 1.1a), called the *director*, which is represented by a unit vector  $\mathbf{n}$ . The states of director  $\mathbf{n}$  and  $-\mathbf{n}$  are indistinguishable, even if the 2 ends of the particles are different (for example when the particles are electrical dipoles).

If the particles do not possess mirror symmetry, the system can also exhibit cholesteric phase. The cholesteric phase is similar to nematic phase in that there is only long-range orientational order in the system. However it differs from the nematic phase in that the director varies regularly in space, as can see in figure 1.1b. In the cholesteric phase there exists a helical axis, along which the director rotates when one moves along the axis (in the figure the helical axis is  $\mathbf{z}$ ). The distance along the helical axis over which  $\mathbf{n}$  rotates a full circle is called the *pitch* of the cholesteric. However, because the 2 states of director  $\mathbf{n}$  and  $-\mathbf{n}$  are indistinguishable, the periodicity of the cholesteric phase is only half of the pitch.

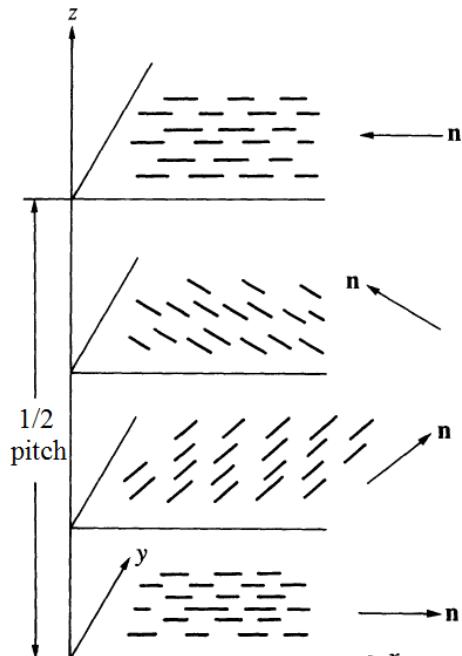
Recently, a new variant of nematic phase has been observed in experiments on bent core particles [26, 27] although the theoretical prediction of its existence was made more than 40 years ago [28]. In this variant, the system possesses orientational

## 1.1. Liquid crystalline phases.

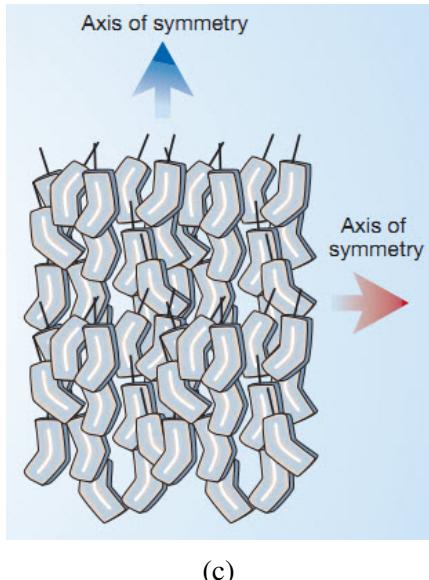
order both along a long axis and another short axis (see figure 1.1c).



(a)



(b)



(c)

- (a) Uniaxial nematic phase. Picture adapted from Ref. [29].
- (b) Change of nematic director in cholesteric phase. Picture adapted from Ref. [30].
- (c) Biaxial nematic phase. Picture adapted from Ref. [29].

Figure 1.1: Variants of the nematic phases.

**Smectics** In smectics the particles apart from having orientational order also organize in layers, which can slide over one another. Particles in the same layer do not exhibit any positional order, so each layer is a 2D liquid. The thickness of the layers can range from approximate the length of the particles to thousands of Ångströms.

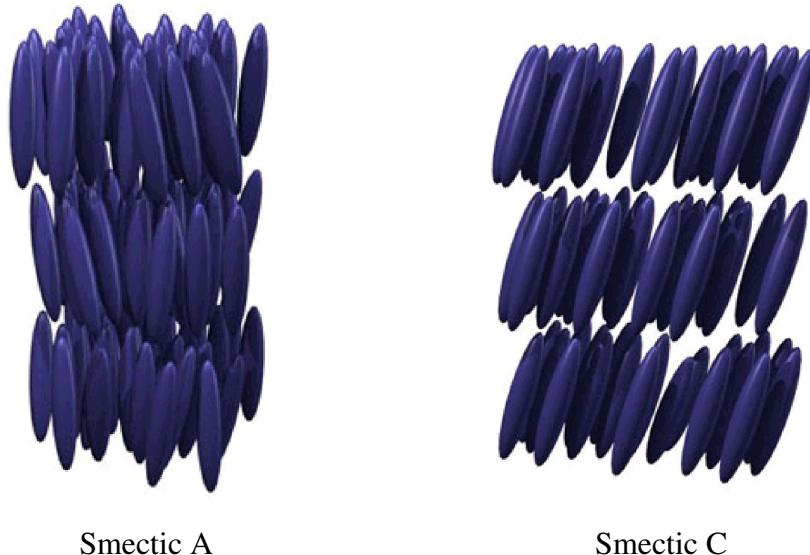


Figure 1.2: Variants of the smectic phases. Images reproduced from [31].

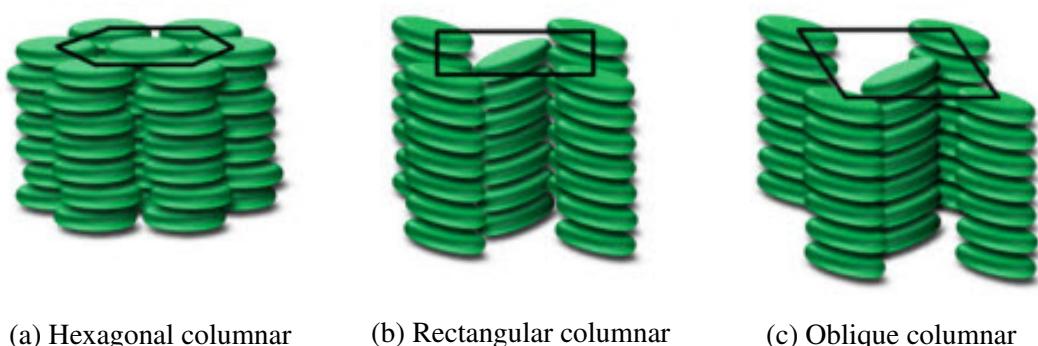


Figure 1.3: Variants of the columnar phases.

Similarly to nematics, there are also many variants of smectics. The most common of them are smectic A and smectic C (figure 1.2). In smectic A, the director of each layer is normal to the planes of the layers. On the other hand, in smectic C the director is tilted.

**Columnars** In columnars the particles stack into columns. In each column the particles exhibit no long range positional order, while the columns themselves pack into 2D crystal. The most common variant among the columnars is the hexagonal phase, in which the 2D crystal formed by the columns is a hexagonal lattice, as can be seen in figure 1.3a. In some other variants the particles within each column can be tilted with respect to the column axis, as in the rectangular columnar 1.3b and oblique columnar phase 1.3c.

## Order parameter of the nematic phase

In this study we limit ourselves only in uniaxial nematic phase. Denote the unit vector indicating the directions of particles in the systems  $\mathbf{u}_i$ . As discussed above, in the nematic phase there is no preference in the arrangement of the 2 ends of

### 1.1. Liquid crystalline phases.

---

single particles. Therefore the average of  $\mathbf{u}_i$  should be vanish. Thus, the nematic order should be a tensor quadratic in  $\mathbf{u}_i$ :

$$\underline{\mathbf{Q}} = \langle \mathbf{u}_i \otimes \mathbf{u}_i \rangle - \frac{1}{3} \mathbf{I} \quad (1.1)$$

In uniaxial phases the average can be rewritten via the director  $\mathbf{n}$  as:

$$\underline{\mathbf{Q}} = S \left( \mathbf{n} \otimes \mathbf{n} - \frac{1}{3} \mathbf{I} \right) \quad (1.2)$$

This tensor  $\underline{\mathbf{Q}}$  has 3 eigenvalues:

$$q_1 = q_2 = -\frac{1}{3}S \quad q_3 = \frac{2}{3}S$$

where

$$S = \frac{1}{2} [3 \langle (\mathbf{u}_i \cdot \mathbf{n}) \rangle] \quad (1.3)$$

The scalar quantity  $S$  measures the level of alignment of particles, and can be expressed via the angle distribution  $f(\theta)$ , where  $\cos \theta = \mathbf{u}_i \cdot \mathbf{n}$ , as

$$S = \frac{1}{2} \langle 3 \cos^2 \theta - 1 \rangle = 2\pi \int_0^\pi P_2(\cos \theta) f(\theta) \sin \theta d\theta \quad (1.4)$$

where  $P_2(\cos \theta)$  is the second order Legendre polynomial. It can be shown easily that in the isotropic phase, where particles have no preferential orientations,  $S = 0$  and when the orientations of particles are perfectly ordered,  $S = 1$ .

### Distinguish between uniaxial nematic and other liquid crystalline phases in simulation

Because our attention is on the transition between isotropic phase and uniaxial nematic phase, in our simulations sometimes we had to check that the systems we were simulating were not in some other liquid crystalline phases.

First, to check if the system is not biaxial, we considered the 3 eigenvalues  $q_i$  of tensor  $\underline{\mathbf{Q}}$ . Denote  $q_1$  and  $q_3$  the smallest and largest, respectively, eigenvalue. In the uniaxial phase, in thermodynamic limit, one has  $q_1 = q_2$ . Since our systems are finite, the equality does not hold, but the quantity which measures the biaxiality of the system:

$$\eta = \frac{q_2 - q_1}{q_3} \quad (1.5)$$

is very close to zero. On the other hand if  $\eta$  is remarkably different from zero, then the system is biaxial.

To distinguish between the nematic phase and smectics and columnars, one possibility is to calculate the 3D pair distribution function and analyse the positional correlation in the system. The 3D pair distribution function is defined as:

$$g(\mathbf{r}) = \frac{1}{\rho N} \left\langle \sum_{i=1}^N \sum_{j \neq i} \delta(\mathbf{r} - (\mathbf{r}_i - \mathbf{r}_j)) \right\rangle \quad (1.6)$$

where  $\delta(\mathbf{r})$  is the Dirac delta function.

## 1.2 DNA duplexes

This introduction on the structure, geometry and phase behaviour of DNA duplexes follows the review provided in Ref. [32].

DNA is a strand composed of repeated units called nucleotides. Each nucleotide consists of a sugar molecule (deoxyribose), a nitrogen base (nucleobase) and a phosphate group. There are 4 types of nucleotides, distinguished from each other by their nucleobases: adenine (A), guanine (G), cytosine (C), and thymine (T). The sugars and phosphate groups of different nucleotides alternately bond to each other by covalent bonds and form the backbone of the DNA strand. It is the sequence of the nucleotides that encodes the genetic information.

In nature, DNA strands usually do not stand alone but pair up with each other in the double helix shape. The interaction between nucleotides belong to the 2 strands are weak and allow the double helix to zip or unzip in the replication process in the cells. The pairing of nucleotides are complementary: A bonds with T and C bonds with G. The most common conformation of DNA double helix is B-DNA, in which the nucleobases are on average perpendicular to the axis of the helix. The periodicity of the helix is around 10 base pairs ( $\sim 3.4$  nm) while the diameter of the helix is around 2 nm. An illustration of B-DNA form is provided in figure 1.4.

There are 2 kinds of interaction between the nucleotides of the 2 strands of a dou-

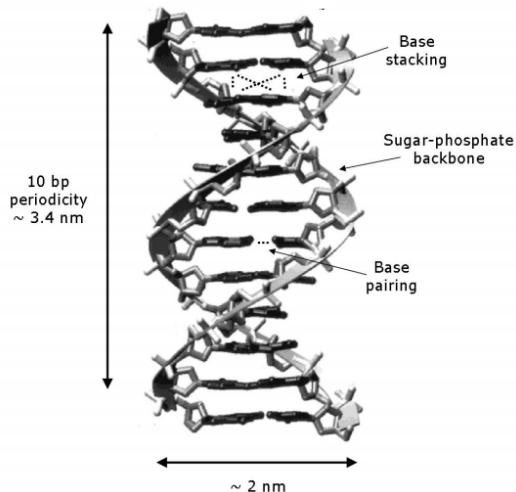


Figure 1.4: The double helix structure of DNA in its B-form, with its relevant dimensions and the main interactions. Image and caption reproduced from Ref. [32].

ble helix: base pairing and base stacking. Base pairing are hydrogen bonds. The free energy involves in base pairing is negligible compared to that in base stacking in C-G pair and even positive in A-T pair. This suggests that in fact the role of base pairing interaction is only to prevent the pairing of uncomplimentary nucleotides. On the other hand, base stacking are non covalent attractive interactions between adjacent nucleobases and plays the main role in holding the 2 strands together. The ultimate conclusion on the nature of base stacking forces is not yet reached, but hydrophobic effect is considered a plausible explanation.

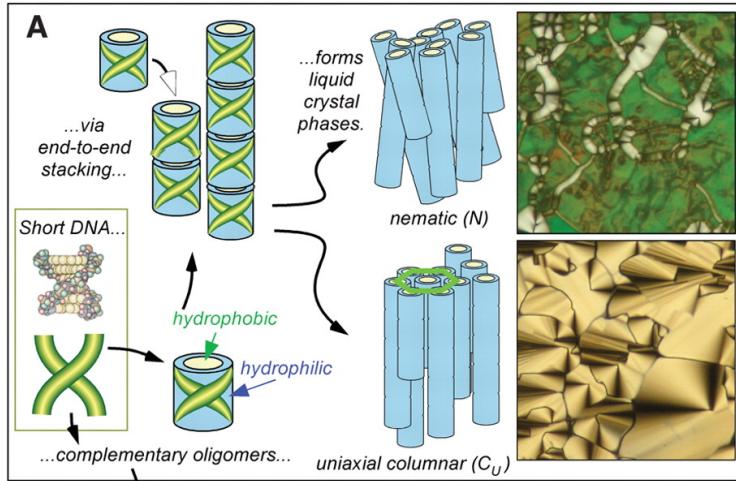
Long DNA, consisting of  $10^2 - 10^4$  base pairs, has been known for half a century to form liquid crystal phases *in vitro* and *in vivo*. It also has been shown that when

## 1.2. DNA duplexes

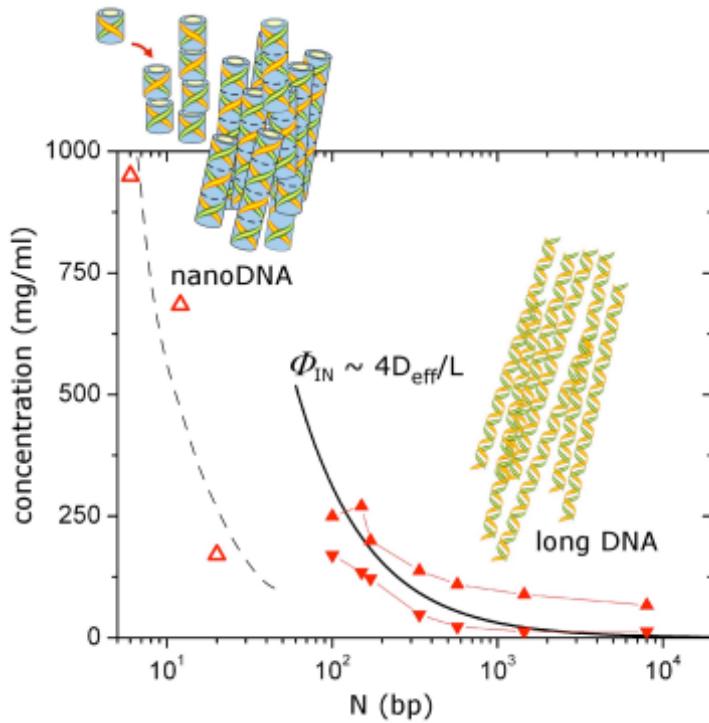
---

it comes to isotropic-nematic phase transition, long DNA can be well described as long hard rods: its nematic critical concentration is in good agreement with the prediction of Onsager theory if the effective diameter is adjusted to account for electrostatic repulsion (figure 1.5b).

However, a recent experiment by Nakata et. al. [18] reported that even short B-DNA duplexes (less than 20 base pairs) with end-to-end stacking interaction form liquid crystal phases, including nematic and columnar phases (figure 1.5a). The B-DNA duplexes have diameter  $D \approx 2\text{nm}$  and length  $L \approx N/3\text{nm}$  where  $N$  is the number of base pairs, i.e. the elongation of those short DNA duplexes are less than 4. Both prediction by Onsager theory for long hard rod and numerical study of hard rods shows that particles of elongation  $L/D < 4.7$  do not exhibit nematic order at any packing fraction [33]. The nematic phase observed in short B-DNA duplexes therefore is likely a consequence of the self-assembly of those duplexes into rigid aggregation. Indeed, when unpaired bases were added to those duplexes to weaken stacking interaction, the liquid crystal phases disappear.



(a) Nano-length B-DNA duplexes can be idealized as hydrophilic cylinders with hydrophobic ends capable of end-to-end adhesion and stacking into units sufficiently anisotropic to orientationally and positionally order into liquid crystal phases. The nematic (N) phase is formed at lower concentration and the uniaxial columnar phase at higher concentration. Image and caption reproduced from [18].



(b) Phase behavior of DNA molecules of different lengths. Helices longer than 100 bp (full triangles) display isotropic-to-nematic transition at concentrations reasonably well described by Onsager theory once the diameter is rescaled for electrostatic repulsion (continuous black line). On the contrary, short DNA strands lack the necessary anisotropy to spontaneously align and their critical concentrations (open triangles) imply end-to-end stacking into longer aggregates. The dashed line is a guide to the eye. Long DNA and nanoDNA data are from Refs. [34] and [18], respectively. Figure and caption reproduced from Ref. [32].

# Chapter 2

## Theories

In this chapter we briefly summarize 2 important theories of isotropic-nematic phase transition and self-assembly, namely Onsager theory and Wertheim theory, which often serve as frameworks for theories that study the coupling of the 2 phenomena. Afterwards the theory proposed by De Michele et. al. [25] to study the phase transition in self-assembly systems is presented.

### 2.1 Onsager theory for isotropic-nematic phase transition in systems of long hard rods

Onsager theory [21] pioneered the study of isotropic-nematic phase transition. This was the first theory to explain the transition by considering only the competition of 2 types of entropy, namely orientational entropy and the entropy associating with the orientation-dependent excluded volume, rather than the usual energy-entropy competition. Onsager theory was also among the first ones to apply density functional theory in solving a physics problem.

Below we give a short summary of the theory, following both the original article and the review made in Ref. [35].

Onsager theory considers the dilute limit of rod systems, and the free energy consists of the ideal part and the second virial expansion:

$$\tilde{f} = \frac{\beta F}{N} = \ln(\Lambda^3 \rho) - 1 - S_{or} + B_2 \rho + \mathcal{O}(\rho^2) \quad (2.1)$$

where  $\beta$  is, as usual, the inverse of the multiplication of temperature and Boltzmann constant,  $\Lambda$  is the De Broglie wavelength,  $\rho$  is the number density of the system;  $S_{or}$  is the orientational entropy and relates to the orientational distribution of the particles  $f(\mathbf{u})$  as follows:

$$S_{or} = - \int f(\mathbf{u}) \ln[4\pi f(\mathbf{u})] d\Omega \quad (2.2)$$

$B_2$  is the second virial coefficient which accounts for the 2-particle interaction and can be computed via Mayer function:

$$B_2 = -\frac{1}{2V} \int d\mathbf{r}_1 \int d\mathbf{r}_2 \int d\Omega_1 \int d\Omega_2 f(\mathbf{u}_1) f(\mathbf{u}_2) \Phi_{1,2} \quad (2.3)$$

where  $\Phi_{1,2}$  is the Mayer function  $\Phi_{1,2} = \exp(\beta U_{1,2}) - 1$ ,  $U_{1,2}$  is the pairwise potential energy. For hard core particles,  $\Phi_{1,2} = -1$  if the 2 particles 1,2 overlap, and  $\Phi_{1,2} = 0$  otherwise. As a result,  $B_2$  relates to the excluded volume  $v_{exc}$  of the 2 particles as follows:

$$B_2 = -\frac{1}{2} \int d\Omega_1 \int d\Omega_2 f(\mathbf{u}_1) f(\mathbf{u}_2) v_{exc} \quad (2.4)$$

For long rods, where the end effects can be neglected, the excluded volume between 2 rods forming an angle  $\gamma$  with each other was estimated by Onsager as:

$$v_{exc} = 2L^2 D |\sin \gamma| \quad (2.5)$$

where  $L$  and  $D$  are the length and diameter of the rods, respectively.

In summary, the free energy (2.1) can be rewritten as:

$$\tilde{f} = \ln(\Lambda^3 \rho) - 1 + \int f(\mathbf{u}) \ln[4\pi f(\mathbf{u})] d\Omega - \rho L^2 D \int d\Omega_1 \int d\Omega_2 f(\mathbf{u}_1) f(\mathbf{u}_2) \sin \gamma \quad (2.6)$$

The orientational distribution is determined as the one that minimizes the free energy in (2.6) with the constraint  $\int d\Omega f(\mathbf{u}) = 1$ :

$$\frac{\delta}{\delta f} (\tilde{f}[f]) = \tilde{\mu} \quad (2.7)$$

where  $\tilde{\mu}$  is the Lagrange multiplier. After taking the derivation in (2.7) one arrives at:

$$\begin{aligned} \ln[4\pi f(\mathbf{u})] &= \tilde{\mu} - 2\rho L^2 D \int d\Omega f(\mathbf{u}) \sin \gamma \\ &= \tilde{\mu} - \frac{8}{\pi} \phi X_0 \int d\Omega f(\mathbf{u}) \sin \gamma \end{aligned} \quad (2.8)$$

Here one introduces the notations of packing fraction  $\phi = \rho L \frac{\pi D^2}{4}$  and elongation of a rod  $X_0 = L/D$ .

One notices that equation (2.8) always take the isotropic form of  $f(\mathbf{u})$ , namely  $f(\mathbf{u}) = 1/(4\pi)$  as one of its roots. In this case the second virial coefficient  $B_2 \equiv B_2^{iso} = \pi L^2 D / 4$  and the free energy of the system in isotropic phase is:

$$\tilde{f} = \ln(\Lambda^3 \rho) - 1 + \rho B_2^{iso} \quad (2.9)$$

which yields the excess chemical potential of the system as:

$$\beta \mu^{exc} = 2\rho B_2^{iso} \quad (2.10)$$

On the other hand substituting  $f(\mathbf{u}) = 1/(4\pi)$  into equation (2.8) one has:

$$\tilde{\mu} = 2\rho B_2^{iso} \quad (2.11)$$

i.e.,  $\tilde{\mu} = \beta \mu$ .

At high packing fraction, however, equation (2.8) can have a second root. Instead

## 2.1. Onsager theory for isotropic-nematic phase transition in systems of long hard rods

---

of resolving the equation by directly taking the derivative, Onsager chose a trial function with a variational parameter  $\alpha$  and solving the equation

$$\frac{\partial \tilde{f}}{\partial \alpha} = 0 \quad (2.12)$$

Denote  $\mathbf{z}_0$  the unit vector along the nematic director. Due to the symmetry of the hard rods, the trial function should fulfil the 2 following condition:

$$\begin{cases} f(\mathbf{u}) \equiv f(\theta) & \text{where } \theta = \cos^{-1}(\mathbf{z}_0 \cdot \mathbf{u}); \text{ and} \\ f(\theta) = f(\pi - \theta) \end{cases} \quad (2.13)$$

The trial function chosen by Onsager was:

$$f_O(\theta) = \frac{\alpha}{4\pi \sinh \alpha} \cosh(\alpha \cos \theta) \quad (2.14)$$

With this trial function, the nematic order can be computed as:

$$\begin{aligned} S(\alpha) &= \langle P_2(\cos \theta) \rangle = 2\pi \int_0^\pi P_2(\cos \theta) f_O(\theta) d\theta \\ &= 1 + \frac{3}{\alpha^2} - 3 \frac{\coth \alpha}{\alpha} \end{aligned} \quad (2.15)$$

So when  $\alpha \rightarrow 0$ ,  $f_O(\theta)$  reduces to the isotropic distribution  $\rightarrow 1/(4\pi)$  and consistently,  $S \rightarrow 0$ . In the opposite limit when  $\alpha \rightarrow \infty$   $S \rightarrow 1$ , i.e. perfect order.

The spread of  $f_O(\theta)$  in angle is measured by mean square of  $\sin(\theta'/2)$ , this time  $\theta'$  is the acute angle between the direction of a rod and the nematic director:

$$\langle (2 \sin(\theta'/2))^2 \rangle = 2 \frac{\coth \alpha}{\alpha} \approx \frac{1}{\alpha} \quad (2.16)$$

namely the standard deviation of the angle is approximately  $\alpha^{-1/2}$ .

The free energy in the nematic phase has the following form:

$$\tilde{f} = \ln(\Lambda^3 \rho) + f_1(\alpha) + 2\phi X_0 f_2(\alpha) \quad (2.17)$$

The coexistence of 2 phases is determined by imposing the condition of equality of chemical potential and pressure of the 2 phases, which relate to the free energy as follows:

$$\begin{aligned} \beta \mu &= \tilde{f} + \rho \left( \frac{\partial \tilde{f}}{\partial \rho} \right) \\ \beta P &= \rho^2 \left( \frac{\partial \tilde{f}}{\partial \rho} \right) \end{aligned} \quad (2.18)$$

For the isotropic phase one has to use expression (2.9) for the free energy, while for nematic phase one uses (2.17) and the value of  $\alpha$  fulfilling (2.12).

In the limit  $L/D \rightarrow \infty$  Onsager theory gives exact results [36]. However, given that Onsager theory neglects the virial coefficient of higher order than second, it is expected that theory only works in low packing fraction and largely elongated rods. It is suggested that the theory only produces quantitatively correct result when  $L/D > 100$ , although qualitatively reliable results can be expected for shorter rods of aspect ratio  $L/D > 20$  and higher [37, 38, 39].

## 2.2 Wertheim theory for isotropic self-assembling systems

A study which applies Wertheim perturbation theory in investigating a similar model was introduced in [22]. A very good agreement with simulation results was observed in the isotropic phase at low density ( $\rho < 0.1$ ). We briefly summarize the Wertheim theory in this section (most of this part follows Ref. [22]).

In this theory, the free energy is written as a sum of the free energy of the reference system (hard sphere in the present case)  $F_W^{hs}$ , i.e.[40]:

$$\frac{\beta F_W^{hs}}{V} = \rho [\ln \rho - 1] + \rho \frac{4\phi - 3\phi^2}{(1-\phi)^2} \quad (2.19)$$

and the contribution of the bonding part  $F_W^{bond}$ :

$$\frac{\beta F_W^{bond}}{V} = \rho (2 \ln X - X + 1) \quad (2.20)$$

where  $X$  the fraction of sites that are not bonded.

A chain of  $l$  particle consists of 2 unbonded sites, which has the probability of  $X^2$ , and  $l-1$  bonds, which occurs at the probability of  $(1-X)^{(l-1)}$ . Thus the number density  $v(l)$  of chains of length  $l$  is related to  $X$  by:

$$v(l) = \rho X^2 (1-X)^{(l-1)} \quad (2.21)$$

$X$  can be computed from the mass-action equation:

$$X = \frac{1}{1 + 2\phi X \Delta} \quad (2.22)$$

Where we have introduced  $\Delta$ :

$$\Delta = \frac{1}{v_d} \int g_{HS}(\mathbf{r}) [\exp(\beta u_0) - 1] d\mathbf{r} \quad (2.23)$$

Here  $g_{HS}(\mathbf{r})$  is the reference hard sphere pair correlation function and the integral is taken over the bonding volume.

If one considers the linear approximation of  $g_{HS}(r)$  one has [41]:

$$\Delta = \frac{V_b}{v_d} [\exp(\beta u_0) - 1] g(\phi) \quad (2.24)$$

where  $g(\phi)$  is given by

$$g(\phi) = A_0(\phi) + A_1(\phi) x_{FK} \quad (2.25)$$

The 2 functions  $A_0$  and  $A_1$  depend only in the packing fraction of the system:

$$\begin{aligned} A_0(\phi) &= \frac{1 - \phi/2}{(1 - \phi)^3} \\ A_1(\phi) &= \frac{9\phi(1 + \phi)}{2(1 - \phi)^3} \end{aligned} \quad (2.26)$$

On the other hand the factor  $x_{FK}$  depends only on the geometry of the system. For the Kern-Frenkel model one has

$$x_{FK} = 1 - \frac{3}{4} \frac{(1+\delta)^4 - 1}{(1+\delta)^3 - 1} \quad (2.27)$$

By rewriting equation (2.22) in the form of chemical equilibrium equation

$$\frac{1-X}{X^2} = 2\phi\Delta \quad (2.28)$$

one can see that within Wertheim theory, bonding can be seen as a chemical reaction between two unreacted sites forming a bonded pair where the quantity  $2\Delta$  plays the role of an equilibrium constant (in unit of inverse concentration). In view of the meaning of  $2\Delta$  one can derive the formula for the entropic contribution per bond in the low packing fraction and low temperature limit ( $e^{\beta u_0} \gg 1$ ):

$$\sigma_b = \ln(2V_b/v_d), \quad (2.29)$$

On the other hand, the theory can also be viewed as a mean field theory of chain association, with  $\ln 2\Delta$  the free energy of one bond. In this view the bonding free energy of the whole system can be written as the difference between the free energy of chains, including the transitional contribution of chains and the bonding energy to make a chain, with the ideal gas free energy of monomers:

$$\frac{\beta F_{bonding}}{V} = \sum_{l=1}^{\infty} v(l) [\ln v(l) - 1 - (l-1)\ln 2\Delta] - \rho (\ln \rho - 1) \quad (2.30)$$

Using the expression (2.21) of  $v(l)$  and definition (2.28) of  $\Delta$ , one can retrieve the expression (2.20). The advantage of the theory is that it provides a well-defined description of the bond free energy via the pair distribution function of the system. The average chain length within Wertheim theory is given as:

$$M_W \equiv \frac{\sum_{l=1}^{\infty} l v(l)}{\sum_{l=1}^{\infty} l v(l)} = \frac{1}{X} = \frac{1}{2} \left( 1 + \sqrt{1 + 8\phi\Delta} \right) \quad (2.31)$$

To lead to the final expression of  $M_W$ , we made use of the equation (2.22).

It is worth noting that the Wertheim theory uses the hard sphere system as the reference system and does not require any fitting parameter. The excluded volume interactions enter via the radial distribution function of the reference system, as shown in equation (2.24).

Figure shows simulation result obtained in Ref. [22] for hard spheres of the same size with our model, with 2 opposite spherical patches at number density  $\rho < 0.1$  (equivalent to packing fraction  $\phi \approx 0.05$ ). As can be seen in the figures, the theory produces results with excellent agreement with simulation at low density.

## 2.3 De Michele et. al. theory for isotropic-nematic phase transition in self-assembly system

We apply here the same theory that was used for studying self-assembly of bi-functional patchy superquadrics [25]. In the investigated model, the bi-functional

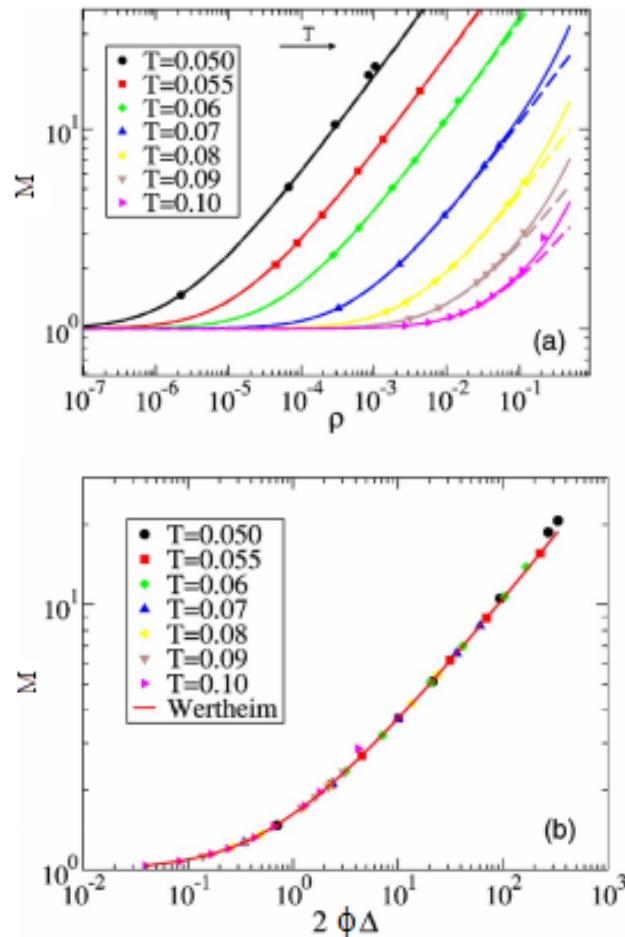


Figure 2.1: (a) Average chain length as a function of the density for all studied temperatures. Lines are the Wertheim theory predictions: dashed lines assume  $g_{HS} = 1$ , while full lines are based on the full radial dependence of  $g_{HS}$ . (b) Scaled representation of  $L$  vs  $2\phi\Delta$ . The line is the function  $M(x) = (1 + \sqrt{1 + 2x}) / 2$ . In both figures symbols are simulation data. Figures and captions from [?].

particles aggregates to form chains of various lengths. In principle, chains of any length, i.e. number of monomers,  $l$  can be generated, from  $l = 1$  to  $l = \infty$ . The system is thus composed by a polydisperse set of polymers, which continuously change their lengths, by merging with other polymers or breaking of the constituting bonds. It is natural to write the free energy of such system as a sum over all possible lengths, assuming the only interaction between the different polymers is dictated by excluded volume interactions. In this case the free energy  $F$  can be written as:

$$\begin{aligned} \frac{\beta F}{V} &= \sum_{l=1}^{\infty} v(l) \{ \ln [v_d v(l)] - 1 \} + \sum_{l=1}^{\infty} v(l) \sigma_o(l) \\ &+ \frac{\eta(\phi)}{2} \sum_{\substack{l=1 \\ l'=1}}^{\infty} v(l) v(l') v_{excl}(l, l') - (\beta u_0 + \sigma_b) \sum_{l=1}^{\infty} (l-1) v(l). \end{aligned} \quad (2.32)$$

Here  $\beta = 1/k_B T$  and  $k_B$  is the Boltzmann constant which is set to 1, so that  $T$  is measured in units of the potential well.  $v(l)$  denotes the number density of chains with length  $l$ , which obeys the normalization condition:

$$\sum_{l=1}^{\infty} l v(l) = \rho \quad (2.33)$$

where  $\rho$  is the number density of the system.  $v_d$  is volume of a monomers. Chains of different lengths are treated as different species and the first term is the free energy of a mixture of ideal gas of these species. In the second term,  $\sigma_o(l)$  describes the orientational entropy lost in the free energy of a chain of length  $l$  in the nematic phase.

The third term accounts for the contribution of steric interaction.  $v_{excl}(l, l')$  is the excluded volume of two chains of lengths  $l$  and  $l'$ . To compensate for the absence of higher virial coefficients one introduces the Parsons-Lee prefactor [42]:

$$\eta(\phi) = \frac{1}{4} \frac{4-3\phi}{(1-\phi)^2} \quad (2.34)$$

where  $\phi$  is the packing fraction of the system and relates to number density as follows:

$$\phi = \rho v_d \quad (2.35)$$

The last term accounts for the bonding free energy and includes the energetic ( $u_0$  term) and entropic ( $\sigma_b$  term) contributions.

According to Ref [24], the excluded volume has the quadratic form of the chain lengths:

$$v_{excl}(l, l') = 2A + 2BX_0^2 ll' + 2v_d K \frac{l+l'}{2} \quad (2.36)$$

$X_0$  is the elongation of the particle, and reduces to 1 if the particle is sphere. Parameters  $A$ ,  $K$  and  $B$  depend on  $f(\mathbf{u})$ , the orientational distribution function of monomers. They can be computed numerically. The detailed procedure is described in 3.4.2.

As we assume that the system is isodesmic, namely the bonding free energy lost when a particle is added to a chain is independent on the chain length, the chain length distribution takes the exponential form:

$$v(l) = \frac{\rho}{M(M-1)} e^{-l[\ln(M) - \ln(M-1)]} \quad (2.37)$$

Here,  $M$  is the average chain length and is defined by:

$$M = \frac{\sum_1^\infty l v(l)}{\sum_1^\infty v(l)} \quad (2.38)$$

Substituting equations (2.37) and (2.36) into (2.32) and calculating the sums over chain lengths, one has:

$$\begin{aligned} \frac{\beta F}{V} &= \frac{\rho}{M} \left[ \ln \left( \frac{v_d \rho}{M} \right) - 1 \right] + \rho \frac{M-1}{M} \ln(M-1) - \rho \ln M + \sum_{l=1}^{\infty} v(l) \sigma_o(l) + \\ &+ \eta(\phi) \left[ B + \frac{K}{M} + \frac{A}{M^2} \right] \rho^2 - \rho (\beta u_0 + \sigma_b) \left( 1 - \frac{1}{M} \right) \end{aligned} \quad (2.39)$$

At equilibrium, the value of  $M$  is the one that minimizes the free energy, i.e.:

$$\frac{\partial(\beta F/V)}{\partial M} = 0 \quad (2.40)$$

### 2.3.1 Isotropic phase

In the isotropic phase the orientational distribution function  $f$  is uniform, i.e.:

$$f(\mathbf{u}) = \frac{1}{4\pi} \quad (2.41)$$

The parameters of  $v_{excl}$  in this case are denoted as  $A_I$ ,  $K_I$  and  $B_I$ . Since monomers randomly orientate, the term contains  $\sigma_o(l)$  in (2.39) vanishes. The free energy is now written as:

$$\begin{aligned} \frac{\beta F_I}{V} &= \frac{\rho}{M} \left[ \ln \left( \frac{v_d \rho}{M} \right) - 1 \right] + \rho \frac{M-1}{M} \ln(M-1) - \rho \ln M + \\ &+ \eta(\phi) \left[ B_I + \frac{K_I}{M} + \frac{A_I}{M^2} \right] \rho^2 - \rho (\beta u_0 + \sigma_b) (1 - M^{-1}) \end{aligned} \quad (2.42)$$

By taking the derivative as in (2.40) and neglecting terms of order  $\mathcal{O}(1/M^2)$ , one obtains the following formula for average chain length:

$$M_I = \frac{1}{2} + \frac{1}{2} \sqrt{1 + 4\phi e^{\sigma_b} e^{K_I \phi \eta(\phi) + \beta u_0}} \quad (2.43)$$

$\sigma_b$  can be deduced if we compare the expression of  $M_I$  above with the one derived in Wertheim theory at low packing fraction and temperature limit:

$$M_I^W = \frac{1}{2} \left( 1 + \sqrt{1 + 8\phi \frac{V_b}{v_d} e^{\beta u_0}} \right) \quad (2.44)$$

At low packing fraction  $e^{K_I \phi \eta(\phi)/v_d} \approx 1$ . By comparing expressions (2.43) and (2.44) in this limit, one obtain:

$$e^{\sigma_b} = \frac{1}{2} \frac{V_b}{v_d} \quad (2.45)$$

### 2.3.2 Nematic phase

In the nematic case, the excluded volume depends on the angular distribution function of monomers  $f(\mathbf{u})$ , which is still unknown. We assume that  $f(\mathbf{u})$  has the following form, which was suggested by Onsager [21] and works quite well for spherocylinders:

$$f(\mathbf{u}) \equiv f_O(\theta) = \frac{\alpha}{4\pi \sinh \alpha} \cosh(\alpha \cos \theta) \quad (2.46)$$

where  $\theta$  is the angle between the monomer's direction and the nematic director and  $\alpha$  is the variational parameter.  $f_O$  is normalized as:

$$2\pi \int_0^\pi f_O(\theta) \sin(\theta) d\theta = 1 \quad (2.47)$$

The factor  $2\pi$  is the result of the integral over the azimuthal angle.  $2\pi f_O(\theta) \sin(\theta) d\theta$  is the probability that the direction of a monomer forms an angle in the range  $[\theta, \theta + d\theta]$  with the nematic director.

Next, we need to evaluate the term which corresponds to the orientational entropy distribution,  $\hat{\sigma}_o = \sum_{l=1}^{\infty} \sigma_o(l) v(l)$ . Up to now there is no exact analytic expression for that term. However, in the following two limits of flexible chain  $\alpha l \gg l_p$  ( $l_p$  is the persistence length) and rigid chain  $\alpha l \ll l_p$  the following formulas for flexible chains have been proposed [?, ?, 25] :

$$\begin{aligned} \sigma_o(l)^{FC} &= \ln(\alpha/4) + \frac{\alpha - 1}{4l_p} \quad \alpha l \gg l_p \\ \sigma_o^{RC}(l) &= \ln(\alpha) - 1 + \frac{\alpha - 1}{6l_p} \quad \alpha l \ll l_p \end{aligned} \quad (2.48)$$

Because the distribution of chain length is rather wide, we use a combination of both limits:

$$\hat{\sigma}_o = \sum_{l=1}^{l_0-1} \sigma_o^{RC}(l) v(l) + \sum_{l=l_0}^{\infty} \sigma_o^{FC}(l) v(l) \quad (2.49)$$

To estimate  $l_0$ , we have to rely on another approximation proposed by Odijk [?], which we cannot use directly due to its rather computational inconvenience in the minimization procedure required to evaluate the free energy in equilibrium:

$$\hat{\sigma}_o^{Od} = \sum_{l=1}^{\infty} v(l) \left\{ \ln(\alpha) + \frac{(\alpha - 1)l}{6l_p} + \frac{5}{12} \ln \left[ \cosh \left( \frac{(\alpha - 1)l}{5l_p} \right) \right] - \frac{19}{12} \ln(2) \right\} \quad (2.50)$$

In the reasonable range of  $M$  and  $\alpha$  for the models considered, we determine  $l_0$  as the value which minimize the difference between the two approximations (2.49) and (2.50) of  $\hat{\sigma}_o$ . Although  $l_0$  is slightly dependent on  $l_p$ , we noted that the value  $l_0 \approx 9$  is appropriate for the systems we studied. In summary, the free energy in the nematic phase is written as:

$$\begin{aligned} \frac{\beta F}{V} &= \frac{\rho}{M} \left[ \ln \left( \frac{v_d \rho}{M} \right) - 1 \right] + \rho \frac{M-1}{M} \ln(M-1) - \rho \ln M + \\ &+ \eta(\phi) \left[ B_N + \frac{K_N}{M} + \frac{A_N}{M^2} \right] \rho^2 - \rho (\beta u_0 + \sigma_b) \left( 1 - \frac{1}{M} \right) + \hat{\sigma}_o \end{aligned} \quad (2.51)$$

In analogy with the isotropic case, we minimize the free energy and neglect the term  $A_N/M^2$ , obtaining the expression for nematic average chain length  $M_N$ :

$$M_N = \frac{1}{2} + \frac{1}{2} \sqrt{1 + \alpha \phi e^{\sigma_b} e^{K_N(\alpha)\phi\eta(\phi) + \beta u_0}} \quad (2.52)$$

### 2.3.3 Phase coexistence

The phase boundaries of the isotropic-nematic transition  $\phi_I = v_d \rho_I$  and  $\phi_N = v_d \rho_N$  can be determined by minimizing the free energy (2.42) and (2.51) with respect to the average chain length of each phase and  $\alpha$ , along with the constraints  $P_I = P_N$  and  $\mu_I = \mu_N$ . Specifically, we have the following set of equations:

$$\frac{\partial F_I(\rho_I, M_I)}{\partial M_I} = 0 \quad (2.53)$$

$$\frac{\partial F_N(\rho_N, M_N, \alpha)}{\partial M_N} = 0 \quad (2.54)$$

$$\frac{\partial F_N(\rho_N, M_N, \alpha)}{\partial \alpha} = 0 \quad (2.55)$$

$$P_I(\rho_I, M_I) = P_N(\rho_N, M_N, \alpha) \quad (2.56)$$

$$\mu_I(\rho_I, M_I) = \mu_N(\rho_N, M_N, \alpha) \quad (2.57)$$

# Chapter 3

## Simulation techniques

In this chapter we present the procedure we applied to study the isotropic-nematic phase transition. The procedure is based on some basic Monte Carlo techniques, which we also briefly describe in this chapter.

In addition, the hit-and-miss Monte Carlo is also employed to compute some parameters required by the theory. Because some details of the simulation method relate to the specific geometry of the models, which are to be introduced in the next chapter, we will discuss them when we introduce the models, and only the general schemes are described here.

### 3.1 Basic Monte Carlo techniques

The introduction below largely follows the details provided in Refs. [43], [44] and [45].

#### 3.1.1 A brief introduction to Monte Carlo method and Metropolis sampling

Monte Carlo method stems from the idea of using random number to solve problem. As an example, let say we want to compute the integral of a particular function  $f(x)$  on  $[a, b]$ , assuming that the function is bounded in this range. One way to do it is employing numerical integration method, like the Sampson method. Alternatively we can enclose the area under the function in a rectangular (see figure 3.1), and then generate a large number of random points  $N$  in the rectangular and counts the ones that fall under  $f(x)$ , denoted  $N_h$ . The integral is given by

$$I = \frac{N_h}{N} \times \text{Area of the rectangular} \quad (3.1)$$

The method describe above is called the hit-and-miss Monte Carlo method. There is another Monte Carlo method to compute  $I$ , in which the integral is rewritten as:

$$I = (b - a) \langle f(x) \rangle \equiv (b - a) \frac{1}{N} \sum_{i=1}^N f(x_i) \quad (3.2)$$

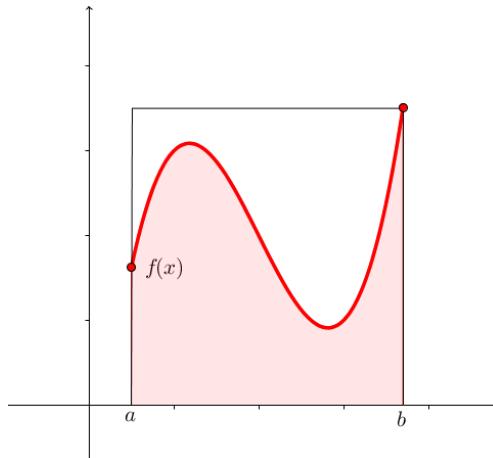


Figure 3.1: An illustration of the 'hit and miss' Monte Carlo method. The bounded function  $f(x)$  on  $[a, b]$ , over which one wishes to compute the integral, is enclosed in a rectangular. A large number of points is generated randomly in this rectangular. The ratio between the number of points falling under the function (the red shaded region) and the total number equals to that between the integral and the area of the rectangular.

In this method we compute the average value of  $f(x)$  on  $[a, b]$  by generating  $N$  random points  $x_i$  in this range and calculate the values of the function at those points.

The question is now which method is better, the numerical method or Monte Carlo. It is suggested that with the numerical method the error in  $d$  dimension decreases as  $n^{-c/d}$  where  $n$  is the number of small range  $[a, b]$  is divided into and  $c$  is some parameter independent of  $n$  and  $d$ . On the other hand the error in Monte Carlo method decreases as  $n^{-1/2}$  regardless of the dimension [45]. This means in 1D the numerical method is a better option, but in 3D and higher dimension spaces, Monte Carlo is better.

In physical problems one usually needs to compute multi-dimensional integrations such as the partition function:

$$Z = \int d\mathbf{r}^N \exp [-\beta H(\mathbf{r}^N)] \quad (3.3)$$

This integral is of dimension  $dN$ , where  $d$  is the dimension of the system. Even if we consider a system consisting of only 100 particles in a cubic box, a tiny number in simulations, and take  $m$  points along each axis, there will be  $m^{300}$  function evaluations to be made, which is impossible. So in this case Monte Carlo method not only is more precise but the only practical means.

**Metropolis sampling** Still, the error produced by uniform sampling described above is only small with smooth functions. However the functions one most often has to deal with in physics, for example the partition functions, are usually narrowly peaked. With these functions most of the random points generated uniformly will fall in regions where the value of the function is zero, while only a small portion of points falls in the regions which contribute most to the integral.

To resolve this problem, Metropolis suggested a method in which instead of just

### 3.1. Basic Monte Carlo techniques

---

generating unrelated random points uniformly, one constructs a series of biased random walk in the high dimension space. Every attempt to move from the current position in phase space will be rejected or accepted depending on the value of the function at the 2 relevant positions.

Denote as  $N(|o\rangle)$  the probability density of finding the system around state  $|o\rangle$ . We would like to construct a matrix  $\pi(|o\rangle \rightarrow |n\rangle)$  of transition from the current state  $|o\rangle$  to a new state  $|n\rangle$  so that  $N$  is proportional to the numbers of points at each state. When the system is in equilibrium, the probability that the system leaves a state has to be equal to the probability that the system arrives at that state from other states, i.e:

$$N(|o\rangle) \sum_{|n\rangle} \pi(|o\rangle \rightarrow |n\rangle) = \sum_{|n\rangle} N(|n\rangle) \pi(|n\rangle \rightarrow |o\rangle) \quad (3.4)$$

This is the balance condition. However, for simplicity's sake, in practice one usually impose a stronger condition, the detailed balance:

$$N(|o\rangle) \pi(|o\rangle \rightarrow |n\rangle) = N(|n\rangle) \pi(|n\rangle \rightarrow |o\rangle) \quad (3.5)$$

Although the detailed balance is not essential, its lack in a simulation scheme is ominous and highly likely to lead to systematic error.

Denote  $\alpha(|o\rangle \rightarrow |n\rangle)$  the matrix determining the probability that the system attempts to 'walk' from the current state  $|o\rangle$  to a new state  $|n\rangle$  and  $\text{acc}(|o\rangle \rightarrow |n\rangle)$  the probability that the trial walk is accepted. The transition matrix can then be written as:

$$\pi(|o\rangle \rightarrow |n\rangle) = \alpha(|o\rangle \rightarrow |n\rangle) \text{acc}(|o\rangle \rightarrow |n\rangle) \quad (3.6)$$

Again, for the sake of simplicity, one usually assume that  $\alpha$  is symmetric, i.e.  $\alpha(|o\rangle \rightarrow |n\rangle) = \alpha(|n\rangle \rightarrow |o\rangle)$ . With this condition and equation (3.6) one derives from the detailed balance (3.5) the equation for the acceptance probability of a trial move:

$$\frac{\text{acc}(|o\rangle \rightarrow |n\rangle)}{\text{acc}(|n\rangle \rightarrow |o\rangle)} = \frac{N(|n\rangle)}{N(|o\rangle)} \quad (3.7)$$

According to the Metropolis scheme one accepts a trial move as follow:

$$\text{acc}(|o\rangle \rightarrow |n\rangle) = \begin{cases} 1 & \text{if } N(|n\rangle) \geq N(|o\rangle) \\ \frac{N(|n\rangle)}{N(|o\rangle)} & \text{otherwise} \end{cases} \quad (3.8)$$

It should be reminded that if a trial move is rejected, the old state should be counted again when computing averages values of the system.

#### 3.1.2 Monte Carlo simulation in canonical ensemble (NVT)

In the canonical ensemble the probability density that a system is found around state  $|o\rangle$  is given by:

$$N(|o\rangle) e^{-\beta U_o} \quad (3.9)$$

The acceptance rule for a trial move from state  $|o\rangle$  to state  $|n\rangle$  is then:

$$\text{acc}(|o\rangle \rightarrow |n\rangle) = \begin{cases} 1 & \text{if } U_n \leq U_o \\ e^{-\beta(U_n - U_o)} & \text{if } U_n > U_o \end{cases} \quad (3.10)$$

where  $\beta = (k_B T)^{-1}$ ,  $T$  is the temperature of the system.

There are 2 kinds of trial moves in an NVT simulation: transitional move and rotational move of single particles, and they should be performed at equal average frequency. The procedure to perform NVT Monte Carlo simulation is as follow

1. Randomly pick a particle in the system.
2. Attempt to rotate or translate the particle. The type of trial move is randomly assigned to ensure detailed balance is fulfilled. One simple way of randomly assign the type of displacement is to generate a randome number  $r$ . If  $r < 0.5$  one goes for the translation move, otherwise rotation move is chosen.
3. Compute the energy difference of the system between the new state and the old one.
4. If  $U_n \leq U_o$ , accept the move. Otherwise compute  $e^{-\beta(U_n - U_o)}$ . Generate a random number  $r$  in the range  $[0, 1]$ .  
If  $r < e^{-\beta(U_n - U_o)}$  accept the trial move. Otherwise the move is rejected.

### 3.1.3 Monte Carlo simulation in isobaric-isothermal ensemble (NPT)

When the system is in isobaric-isothermal ensemble, apart from the displacement moves similar to that in NVT simulation, one wishes to keep its number of particles  $N$ , pressure  $P$  and temperature  $T$  fixed, while its volume (or number density) is allowed to fluctuate. Therefore in NPT simulation, apart from the translational and rotational moves of single particles as in NVT simulation, one also attempts to make a small change to the volume of the system. The scheme to perform a volume change move is as follows:

1. Assuming that the current box size of the simulation box is  $L_o$  and its volume  $V_o = L_o^3$ . One tries to change the volume of the system an amount of  $\Delta V$ , where  $\Delta V$  is a random number in the range  $[-\Delta V_{max}, \Delta V_{max}]$ . To ensure detailed balance, the value of  $\Delta V_{max}$  is fixed in the production stage (i.e. after the system has been equilibrated and now simulation is performed to gather statistics).

The scaling parameter is given by  $r = (1 + \Delta V/V)^{1/3}$ . The new box size and new positions of particles is then

$$\begin{aligned} L_n &= rL_o \\ \mathbf{r}_{i,n} &= r\mathbf{r}_{i,o} \end{aligned} \quad (3.11)$$

The direction of the particles are kept unchanged.

2. Compute the change in system energy  $\Delta U = U_n - U_o$ . The acceptance rule is given by [43]:

$$acc = \min(1, \exp\{-\beta [\Delta U + P\Delta V - N\beta^{-1} \ln(1 + \Delta V/V)]\}) \quad (3.12)$$

where  $P$  is the pressure of the system and  $N$  is the number of particles. Generate a random number  $r$  between  $[0, 1]$ . If  $r < \text{acc}$  one accepts the move, otherwise the move is rejected.

In most cases, the volume change move is a rather expensive one as the interactions between all particles have to be re-computed. The computational cost of a volume change move is comparable of that of  $N$  particle moves. It is therefore recommended that this move is performed at an average frequency of  $1/N$  of that of particle moves. Analogously to in NVT simulation, before each trial move one would decide which kind of move to be perform using a random number rather than regularly.

#### 3.1.4 Monte Carlo simulation in grand canonical ensemble (GCMC)

In this ensemble the number of particles in the system (or its number density) is allowed to fluctuate, under the condition that its chemical potential  $\mu$  is fixed (in fact,  $\mu$  is chemical potential of the system in comparison of that of ideal gas). Hence a new kind of trial move is introduced: the removal/insertion of a particle. The procedure is as follows:

1. Decide whether an insertion or removal is to be perform, again this is done using a random number. The 2 moves should be done at equal average frequency
2. If the move is a removal, one randomly picks a particle in the system to be removed; if it is an insertion, one randomly generates the position and direction of the new particle.
3. Compute the change in system energy  $\Delta U$ . The acceptance rules of removal and insertion are respectively given by [43]:

$$\begin{aligned} \text{acc}_{\text{removal}} &= \min \left\{ 1, \frac{V}{\Lambda^3(N+1)} \exp [\beta (\mu - \Delta U)] \right\} \\ \text{acc}_{\text{insertion}} &= \min \left\{ 1, \frac{\Lambda^3 N}{V} \exp [-\beta (\mu + \Delta U)] \right\} \end{aligned} \quad (3.13)$$

Here  $\Lambda$  is the De Broglie wave length, which is usually set to 1 in simulation.

The frequency of the particle exchange trial move in GCMC should not be too high. Take, for example one has a successful removal move. This will leaves a 'hole' in the system. If the frequency of exchange trial move is high, it is likely that a new particle will be attempted to add to the hole before it is eliminated by the displacement of particles around it. In this case one artificially increases the time the system visit the same state. In this study we set the average frequency of the exchange trial move approximate to the number of particles in the system, therefore on average each particle is displaced around once before another insertion/removal trial is made.

Again, the frequency of exchange trial move should be kept unchanged during the course of simulation production stage in order not to violate the detailed balance.

### 3.1.5 Periodic boundary condition and reduced units in simulation

#### Periodic boundary condition (PBC)

In simulation one can only deal with some thousand or dozens of thousand of particles, while macroscopic systems usually contains much larger number of particles than that (around the order of Avogadro number). With that small number of particles, the interfacial effect becomes significant. Hence to simulate bulk behaviour PBC is often employed. The idea is consider the system as a cell of an infinite lattice, i.e. for each particle at position  $\mathbf{r}$  in the simulation box, there will be an infinite numbers of its images located at:

$$\mathbf{r}_{l,m,n} = \mathbf{r} + \mathbf{i}lL_x + \mathbf{j}mL_y + \mathbf{k}nL_z \quad (3.14)$$

where  $L_x, L_y, L_z$  is the sizes of the simulation box,  $\mathbf{i}, \mathbf{j}, \mathbf{k}$  are the unit vectors along the box edges and  $l, m, n$  is a set of arbitrary 3 integers.

PBC along all 3 dimensions are applied in all of our simulations.

#### Reduced unit

It is more convenient to use reduced units in simulations, namely one chooses the units of length, energy and mass as the basis and units of other quantities are expressed through these units. In addition, many constants like Boltzmann constant, De Broglie wave length, etc. are set as 1. The reason of convenience is twofold. First, many physics constants are of many orders of magnitude different from unit, which presents a difficulty for numerical computation with computers and can heavily affect its precision. Second, many states with different values of quantities in real units actually correspond to just one state in simulation of reduced units. Therefore by employing reduced units one can describe all those equivalent states with just one state in simulation.

From now on to make clear the reduced units are used, we will use the superscription \* whenever we report simulation results.

## 3.2 Scheme to compute phase coexistence lines

### 3.2.1 Determine coexistence packing fraction using successive umbrella sampling (SUS)

One of the main obstacles in studying first order phase transition and locating coexistence point is that to transit from one phase to another, the system has to overcome a large free energy barrier, which relates to the free energy cost to create the interface separating the 2 phases. Straightforward techniques such as NPT and GCMC usually suffer from metastability, where the system remains in the metastable phase even when its pressure or density is far beyond the transition point.

Ref. [46] proposed another method, successive umbrella sampling, to compute the free energy, with which the coexistence point can be determined. When the number of particles can be used as the order parameter, the free energy of the system can be computed via the equilibrium probability distribution  $P(N)$ ,  $F(N) =$

$-1/\beta P(N) + \text{constant}$ . Similarly to  $-F$ , near transition  $P(N)$  also exhibits 2 maxima and a valley in between. The probability difference between the maxima and the valley is usually of many orders of magnitude. Due to the low probability at the valley, if one samples the whole range of number of particles that contains the 2 peaks in one simulation, it is very likely that the system will be trapped under one of the peaks in most of, or even all, simulation time. The result therefore suffers heavily from poor statistics in the region of the valley and the other peak.

The idea of SUS is to divide the range of interest into many small windows of size  $w$ . A GCMC simulation is then performed on each window to measure how frequently the system visit a particular state in that window. The simulations on different windows are independent of each other. Once the data for each window are obtained, the continuity condition is imposed on the window edges to achieve the unnormalized probability distribution in the whole range:

$$\frac{P(N)}{P(0)} = \frac{H_{1r}}{H_{1l}} \cdot \frac{H_{2r}}{H_{2l}} \cdots \frac{H_k(N)}{H_{kl}} \quad (3.15)$$

where  $H_{kl}$ ,  $H_{kr}$  and  $H_k(N)$  is the histogram of the left and right boundary of the  $k$ th windows, and a state  $N$  in that window, respectively.

At coexistence, the equal weight rule applies, i.e. the areas under 2 peaks have to be equal [47, 48, 49]:

$$\int_0^{\langle N \rangle} P(N) dN = \int_{\langle N \rangle}^{\infty} P(N) dN \quad (3.16)$$

where the average value of number of particles  $\langle N \rangle$  is defined by:

$$\langle N \rangle = \frac{\int_0^{\infty} NP(N) dN}{\int_0^{\infty} P(N) dN} \quad (3.17)$$

When one starts the GCMC simulation one does not know the chemical potential at coexistence, and the resulting  $P_{sim}(N)$  is very likely to not fulfil the condition of equal areas mentioned above. However one can reweight it [48] to obtain the distribution at coexistence  $P_{coex}$  thanks to the following relations:

$$\begin{aligned} P_{sim}(N) &= \frac{1}{Z_{GC}} Z_N e^{\beta \mu_{sim} N} \\ P_{coex}(N) &= \frac{1}{Z_{GC}} Z_N e^{\beta \mu_{coex} N} \end{aligned} \quad (3.18)$$

and therefore

$$\frac{P_{coex}(N)}{P_{sim}(N)} = c' \left[ e^{\beta(\mu_{coex} - \mu_{sim})} \right]^N \quad (3.19)$$

Here  $Z_N$  is the canonical partition function and  $Z_{GC}$  is the grand partition function;  $c'$  is a constant independent of  $N$ . Please note that in equation (3.18) the distributions are normalized but normalization is not necessary to apply the reweighing

formula (3.19), as the normalization constants will all go into  $c'$ .

Once one has obtained the coexistence probability distribution, the numbers of particles at coexistence in the 2 phases are defined as:

$$N_1 = 2 \frac{\int_0^{\langle N \rangle} NP(N)dN}{\int_0^{\infty} P(N)dN}$$

$$N_2 = 2 \frac{\int_0^{\langle N \rangle} NP(N)dN}{\int_0^{\infty} P(N)dN} \quad (3.20)$$

In practice, one replaces the  $\infty$  limit in equations (3.20) by the largest number of particles simulated  $N_{max}$ .  $N_{max}$  is chosen so that the probability distribution at this value is smaller than that at the minimum. The steps of SUS are sketched in figure 3.2.

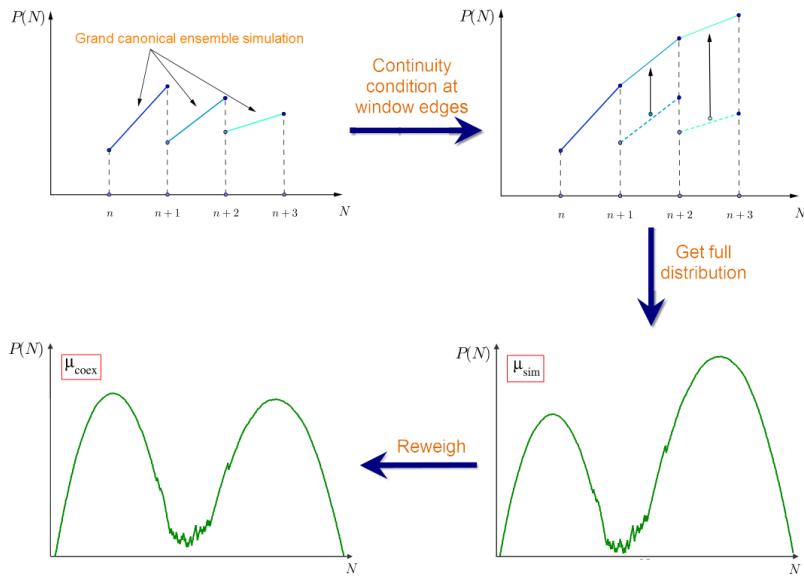


Figure 3.2: A sketch of steps to compute a coexistence point using SUS.

By investigating many small windows instead of a large one, the system is forced to spend a sufficient amount of time in every state and hence one can gather a good statistics for the whole range of  $N$ . Moreover, the waiting time for simulation is also significantly improved if the simulations on different windows are done in parallel. Alternatively one can sample one window after another (as the authors of Ref [46] did) and use the distribution of previous windows to extrapolate the weight function for the next window to even further flatten the distribution of that window.

There are some comments in order for the implementation of SUS.

First, care has to be taken for an attempt to make the system leave its current windows. Apart from rejecting the attempt, one also has to increment the histogram at the corresponding window edge by 1 to fulfil detailed balance.

Second, one has to consider how large an window should be. As discussed earlier, since  $P(N)$  is quite steep for most part, a larger window results in poorer statistics for the same simulation time. Therefore one tempts to limit the window size to 1 whenever possible. However in some cases, for example when the free energy landscape is complex and involves other variables, that practice can lead to difficulty in sampling, as the author discussed in his application of the method in Ref. [50].

Third, one also has to consider whether adjacent windows should overlap by only 1 state or more. Increasing the overlap between adjacent windows can improve the accuracy of  $P(N)$ , but it is rather expensive, because one has to simulate the overlapped states twice. However, in Ref. [46] the authors suggested that the precision should not be significantly lowered if the windows overlap only by 1 state.

**Implementation** In this study we limited both the overlap between adjacent windows and the size of each window to 1, i.e. a particular window  $k$ th includes 2 states  $k$  and  $k + 1$ . In all of our calculations of  $P(N)$  we had to investigate a range of around 400 – 500 particles, therefore it is infeasible to simulate them consecutively to benefit from histogram reweigh. Thus we sampled the windows simultaneously. Furthermore, we figured that it is not always necessary to start sampling from empty state, i.e.  $N = 0$ . By performing a short NVT simulation on a number of systems of different packing fraction, one can have a rough estimate on where phase transition occurs, and afterwards sample only systems around that area. In this case the LHS of equation (3.19) will be  $P(N)/P(N_{min})$ , where  $N_{min}$  is the smallest number of particles one studies. Since all of the systems we sampled with SUS contained no less than around 2000 particles, this practice spared us of the unnecessary effort to sample 2000 windows below the interested range.

All of our initial configurations to start SUS are in nematic phase, for 2 reasons. The first is to reduce metastability, as it is easier for a nematic system to make the transition to isotropic phase if it is the stable one than the other way round. The second reason is to predetermine the director rather than letting random moves in the simulation course decide it. We will explain why this is necessary in section 3.2.3.

After initiating the simulations we first let the systems equilibrate, during this period the histograms was not counted. Depending on specific models, number density of the systems and temperature, equilibrating takes around  $4 \times 10^6$  –  $10 \times 10^6$  MC steps. Once the systems had equilibrated, we stopped the simulations and restarted from the last configuration obtained in the equilibrating stage. We now had entered production stage. In this stage the histograms were updated every time an attempt to insert/delete a particle was made. The statistic were collected over around  $4 \times 10^6$  –  $10 \times 10^6$  MC steps, depending on the systems. As a check whether the statistic had been good enough, we first looked into the evolution of the histograms to see if they were close to convergence, i.e. exhibiting a plateau, or not. However this check alone is not enough. Since the full  $P(N)$  in (3.19) is obtained by 'stitching' hundreds of windows together, even a minor change in the histogram of each window can result in a significant change in  $P(N)$ . Therefore we computed full  $P(N)$  and the values of average numbers of particles at coexistence after every approximately  $1.5$  –  $2 \times 10^6$  MC steps. If the updated values we got did not significantly different from the previous ones (for the average numbers of particles, the criterion is precision to 1 decimal place), we stopped the simulations.

### 3.2.2 Kofke integration

Kofke integration method [51] allows us to obtain the coexistence lines over a range of temperature (or pressure) once one has an initial point of coexistence. When one moves along the coexistence lines the temperature, pressure and chemical potential of the systems remain in equal with each other and their variations obey the following equations:

$$\begin{aligned} d\mu &= -s_1dT + v_1dP \\ d\mu &= -s_2dT + v_2dP \end{aligned} \quad (3.21)$$

where  $s$  and  $v$  are molar entropy and volume, while the subscripts refer to the 2 phases of interest. Take the subtraction of 2 equations in equation set (3.21) and recall that the Gibb potentials of the 2 systems are also equal along the coexistence lines, one obtains:

$$\frac{P}{T} = -\frac{1}{T} \frac{h_2 - h_1}{v_2 - v_1} \quad (3.22)$$

which is nothing but Clausius-Clapeyron equation. Here  $h$  is the molar enthalpy. As equation (3.22) is a differential equation, it can be solved numerically by calculating the integral. For this purpose, it is preferable to rewrite equation (3.22) in the following form so that the slope of the RHS becomes more gentle:

$$\frac{d \ln P}{d\beta} = -\frac{h_2 - h_1}{\beta P(v_2 - v_1)} \quad (3.23)$$

where  $\beta = 1/(k_B T)$  with  $k_B$  is the Boltzmann constant.

For the sake of generality, one rewrites the equation (3.23) as follows:

$$\frac{dy}{dx} = f(x, y) \quad (3.24)$$

where the integral can be taken either over  $x$  or  $y$ . The difference between integrating a differential equation such as the Clausius-Clapeyron equation and integrating a “normal” mathematical function is that the right hand side of (3.24) is only implicitly dependent on the variable. Therefore for each integration step, one needs to perform 2 separate simulations to compute the relevant quantities of both phases and their differences. In the specific case of the Clausius-Clapeyron equation, for example, two NPT simulations at the same pressure and temperature are required to compute the enthalpy and volume per particle. Here lies the advantage of Kofke technique: the 2 simulations can be performed independently and do not have to rely on particle exchange, a huge obstacle in systems of high volume fraction or consisting of large molecules.

Because the RHS of equation (3.24) depends on both the dependent and independent variables, when the independent variable, let say  $x$ , makes an increment  $h$  from  $x_i$  in step  $i$  to  $x_{i+1}$  in step  $i+1$ , the value of  $f$  is still unknown. Therefore the computation of  $y_{i+1}$  has to be done in two steps, the first step to get a rough estimation of how  $f$  varies from  $x_i$  to  $x_{i+1}$  and the second step to refine the result, sometimes relying also on the values obtained in more than 1 previous steps. Such a process points to predictor-corrector techniques. A number of predictor-corrector formulas have been proposed, and the ones suggested by Kofke are listed in table 3.1.

### 3.2. Scheme to compute phase coexistence lines

---

Table 3.1: List of predictor-corrector formulas proposed in Ref. [51]. The denotations P and C stand for predictor steps and corrector steps, respectively.  $n$  is the number of previous steps required to be able to use the formula for current step. The subscripts refer to the step number. The denotations  $x$ ,  $y$ ,  $h$  and  $f$  are the same as in the text.

Name	Type	$n$	Formula
Trapezoid	P	1	$y_{i+1} = y_i + hf_i$
	C	1	$y_{i+1} = y_i + \frac{h}{2}(f_{i+1} + f_i)$
Midpoint	P	1	$y_{i+1} = y_{i-1} + 2hf_i$
	C	1	$y_{i+1} = y_{i-1} + \frac{h}{3}(f_{i+1} + 4f_i + f_{i-1})$
Adams	P	4	$y_{i+1} = y_i + \frac{h}{24}(55f_i - 59f_{i-1} + 37f_{i-2} - 9f_{i-3})$
	C	3	$y_{i+1} = y_i + \frac{h}{24}(9f_{i+1} + 19f_i - 5f_{i-1} + f_{i-2})$

Predictor-corrector algorithm of 3rd and higher order improve accuracy but can lead to instability at large integration steps. Another difficulty arises from employing those algorithms is that integration steps are fixed in the course of integration. Ref. [52] proposed another predictor-corrector scheme that balances between stability and accuracy, and at the same time provides the flexibility in adjusting integration step. In this scheme the predictor step is of little relevance and a simple Euler formula is used. The formula for the corrector step is as follows:

$$y_{i+1} = Ay_{i-1} + By_i + D \sum_{k=-1}^1 C_k f_{i+k} \quad (3.25)$$

Again, the subscripts refer to the integration step. There are 2 alternative sets of the parameters available to use, as listed in table 3.2.

The greatest concern usually raised about Kofke integration is that it is hard to control how errors propagate during the course of integration and how far the integrated data deviate from the actual coexistence lines. Kofke suggested that when it is possible to choose the direction of integration, one can rely on this formula to decide which way leads to smaller error:

$$\Delta P = \Delta P|_i \frac{\Delta v|_i}{\Delta v} \quad (3.26)$$

where  $i$  refers to the initial point; the symbol  $\Delta$  indicates the difference of the variables follow it between 2 phases.

The application of Kofke integration does not limit to the temperature-pressure plane. In fact, equation (3.21) is a special case of the following general equation [53]:

$$\rho_0 d\phi_0 = \rho_1 d\phi_1 + \rho_2 d\phi_2 + \rho_3 d\phi_3 + \dots \quad (3.27)$$

Table 3.2: List of predictor-corrector formulas proposed by Ref. [52]. The subscript  $i$  refers to the step number. The denotations  $x$ ,  $y$  and  $f$  are the same as in the text. In both variants  $h_j = x_{j+1} - x_j$  and  $r = h_{i-1}/h_i$ .

Variant 1	Variant 2
$A = 0$	$A = \frac{1}{r^2(3+2r)}$
$B = 1$	$B = 1 - \frac{1}{r^2(3+2r)}$
$C_{-1} = -\frac{1}{r}$	$C_{-1} = 0$
$C_0 = \frac{1}{r} + 4 + 3r$	$C_0 = \frac{1}{r} + 2 + r$
$C_1 = 2 + 3r$	$C_1 = 1 + r$
$D = \frac{h_i}{6(1+r)}$	$D = \frac{h_i}{3+2r}$

$\phi_i$  are “field” variables and  $\rho_i$  are their conjugate “densities”. Field variables have to be equal among coexistence phases;  $\phi_0$  is the field variable that does not explicitly involve in the integration process. Examples of  $\phi - \rho$  couples are pressure and molar volume, chemical potential and species fraction in a mixture. When one considers the coexistence only between 2 phases, the general form of Cladius - Clapeyron equation is:

$$\frac{d\phi_2}{d\phi_1} = -\frac{\Delta(\rho_1/\rho_0)}{\Delta(\rho_2/\rho_0)} \quad (3.28)$$

This generalized equation opens up to numerous applications. Less “conventional” applications includes studying how isotropic-nematic phase transition depends on the flexibility of polymers [54], studying the effect of polydispersity on fluid-solid transition of hard spheres [55]. Bolhuis and Frenkel also applied this method to trace the coexistence lines of monodisperse hard spherocylinder systems when their aspect ratio varies [33]. Similarly, Philip J. Camp et. al. studied in Ref. [56] the isotropic-nematic transition of uniaxial hard ellipsoids when the elongation varies. Another important class of problems can be addressed with the method is studying the effect of the variational of a potential on coexistence, starting from a reference system, as did by Agrawal and Kofke to study fluid-solid transition in soft spheres with hard sphere as reference system [57] or Abascal and Vega to develop a new model for water [58]. In this case the pairwise potential is written as  $u(\lambda)$  where  $u(\lambda = 0) = u_{ref}$ , the potential of the reference system and  $u(\lambda = 1) = u_d$ , that of the destination system. The conjugate variable of  $\lambda$  is then

$$x_g = \frac{\partial g}{\partial \lambda} = \left\langle \frac{\partial u}{\partial \lambda} \right\rangle \quad (3.29)$$

## Implementation

**Pressure of SUS points** To start Kofke from a point obtained by SUS first one needs to estimate the pressure of the coexistence systems obtained from SUS. If ones starts SUS from an empty states, this can be done easily using equations (3.18) for a normalized distribution:

$$P(0) = \frac{1}{Z_{GC}} \quad (3.30)$$

with  $Z_{GC} = e^{\beta PV}$ . Hence the pressure at coexistence can be computed as  $P = -(1/\beta V) \ln P(0)$ .

However since we did not start SUS from empty state, we turned to a conventional method of using NPT simulations. We performed a batch of NPT simulations at guessed values of pressure and check which pressure yielded the average packing fraction coinciding the values from SUS.

**Predictor-corrector implementation** The predictor-corrector implementation is as follows:

1. Update  $\beta_i$  from the previous step  $i$  to  $\beta_{i+1} = \beta_i + h$ .
2. Compute the value of predictor pressure  $P_p$  from the previous step, using one of the equations in table 3.1. We will report in details about the choices when we discuss each system in chapter 4.
3. Perform NPT at  $P_p$ , computing the average value of molar enthalpy and volume for both phases and the RHS of (3.23).
4. Compute the corrector pressure  $P_c$  using the formula corresponding to the one we used to compute the predictor pressure in table 3.1, or a formula in table 3.2.
5. Perform NPT at  $P_c$  and compute the RHS of (3.23) at this pressure.
6. If the difference between the 2 pressures fulfils the convergence condition we set, i.e

$$\left| \frac{P_c - P_p}{P_p} \right| \leq 5 \times 10^{-3} \quad (3.31)$$

$P_c$  is taken as the pressure at  $\beta_{i+1}$ . Otherwise repeat from step 4, this time using the values of  $P_c$  and the RHS one obtains at this pressure as the predictor values.

At this point almost all our systems yielded a new value of  $P_c^{(1)}$  fulfilling (3.31).  $P_c^{(1)}$  is then accepted as the pressure at  $\beta_{i+1}$ . In the only occasion it did not, the value of  $P_c^{(1)}$  also went further from  $P_p$  than  $P_c$ , signifying that the step size was too large. In this case we reduced the step size and restarted from step 1.

Depending on the temperatures, the RHS of (3.23) is computed over  $4 - 8 \times 10^6$  MC steps, excluding equilibrating.

To check the accuracy of Kofke, we integrated equation (3.23) along both direction and on each direction we computed at least one other coexistence point using SUS.

### 3.2.3 Reducing finite size effect

When performed SUS simulations we tried to keep a stable interface between isotropic and nematic phase in coexistence systems. Due to the employment of PBC, there were in fact 2 parallel interfaces in these systems. When such interfaces exist, the excess free energy in the system is dominated by the interface free energy, which is given by:

$$\Delta F = 2\gamma S \quad (3.32)$$

where  $S$  is the area of the interface, the factor 2 accounts for the presence of 2 interfaces. If one can keep the interaction between the 2 interfaces negligible, the resulting  $P(N)$  will exhibit 2 distinctively separated peaks with a flat valley in between, which clearly defines the areas under each peak and hence the position of phase coexistence. Indeed, if the 2 interfaces do not interact, there will be no free energy cost when one slightly changes their distance by changing the number of particles of a phase between them. On the other hand, without the presence of these stable interfaces, the interface free energy in different systems where coexistence occurs may vary considerably depending on the spontaneous formation of bubbles of a phase inside the other. As a result, the valley  $P(N)$  will have a ragged shape instead of a plateau.

As suggested in Ref. [59], to keep the interaction between the interfaces small we made the simulation boxes elongated, with the box edge normal to the interfaces large in relevance to the areas of the interfaces. To minimize the well depth, the interface should be parallel to the nematic director. Specifically, in our SUS simulations we used boxes whose sizes  $L_x \times L_y \times L_z$  satisfy:

$$L_x \gg L_y \quad \text{and} \quad L_x \gg L_z \quad (3.33)$$

where  $\mathbf{z}$  is parallel to the nematic director. All initial configurations consists of particles aligning perfectly with the nematic director. As the simulations go, in the systems where coexistence occurs interfaces will form perpendicularly to  $x$ . As an illustration, in figure 4.6a we present a snapshot of a hard cylinder system obtained from SUS simulation.

Another concern is the box size parallel to the nematic director. When the system is in nematic phase, the chains can get quite long and the box should be large to not suppress the growth of long chains. In Kofke simulations  $L_z$  was chosen to approximate length of chains consisting of around 40 particles. As for SUS, since  $L_z$  is constrained by condition (3.33),  $L_z$  was usually around the length of 15-mers chains, which was still larger than the average chain length in all of the systems we studied. We assessed the affect of finite size effect by comparing the results of SUS and Kofke integration, as the two simulations were implemented independent of each other except for the initial points of Kofke were obtained from SUS.

## 3.3 Some other details of the simulations

### 3.3.1 Linked list

Most of the simulation effort is spent on calculating the interaction between particles. In each attempt to move a particle in MC simulation, one has to estimate the

interaction of the particle with other  $N - 1$  particles in the systems, i.e. the computation effort scales as  $N$ . However, as the interaction investigated in this study are all short range, each particle can only interact with some few particles close to it and it is not necessary to compute its interaction with the rest. Thus we apply the trick represented in [44] in order to eliminate those unnecessary effort.

The idea is simple: one divides the simulation box into small cells, so that each particle can interact with only other particles in the same cell and 26 neighbour cells. The computation cost for the construction scales as  $N$ . For NVT and GMC simulations, the cells only have to be constructed once at the beginning. The list of neighbour cells of each cell remains during the simulation. After the first construction one also has a list of particles in each cells. Afterwards each time a particle is displaced only its cell has to be updated. In NPT simulation, the cells have to be reconstructed each time the volume is changed, which is performed around once out of every  $N$  attempts to move a single particle. Therefore in all cases the total interaction computation cost is approximately  $N^{-1}$  times less than when the trick is not implemented.

### 3.3.2 Creating an initial configuration at high packing fraction

In our simulations sometimes we need to build up a initial configuration densely packed without any positional order. If one just generates particles randomly the packing fraction cannot reach a very high value. For example with hard sphere this method can hardly go beyond around 0.37. Other options include performing NPT or GCMC simulations, but they are generally slow.

We therefore employed another scheme, adapted from strategies proposed in Refs. [60, 61] to create configurations of densely packed polydispersed hard spheres. Assuming one needs to create a box of size  $L_x \times L_y \times L_z$  containing  $N$  spheres of diameter  $\sigma$ . At first one fills the box of  $N$  particles not of the desired diameter but of diameter  $\sigma^* < \sigma$ . Next one performs a simulation with the following trial moves:

1. Positional displacement: similarly to that of an NVT simulation. However at each attempt one only needs to check that the displaced particle do not overlap with other particles and the usual acceptance rule relating to energy does not apply.
2. Particle enlargement: pick a random particle and expand it to the largest possible diameter not beyond  $\sigma$  that does not involve any overlap.

With this scheme, a configuration of hard sphere with packing fraction of 0.51 can be achieved in a matter of minutes.

## 3.4 Computation of parameters required by theory

### 3.4.1 Bonding volume

In a cubic box of size  $b$  one generates  $N_{cf} = 2 \times 10^8$  random configurations of 2 particles and counts the number of times they form a bond,  $N_b$ . The bonding

volume is then computed as:

$$V_b = \frac{N_b}{N_p^2 N_{cf}} \times b^3 \quad (3.34)$$

Here  $N_p$  is the number of patches of a particle. The size of the box should be large enough for both particles, regardless of their directions, to fit in when they merely touch (both hard core and patches considered).

### 3.4.2 Excluded volume and parameters

To estimate the parameters  $A_I$ ,  $K_I$ ,  $B_I$  and  $A_N$ ,  $K_N$ ,  $B_N$  we performed separate MC simulations to calculate the excluded volume as a function of chain length and of nematic parameter  $\alpha$  in nematic case. For simplicity we take  $l = l'$ , and the formula for the excluded volume of two chain of length  $l$  becomes:

$$v_{excl}(l) = 2(A + v_d Kl + 2BX_0^2 l^2) \quad (3.35)$$

We generate 2 chains, each of which consists of  $l$  monomers in a cubic box of size  $s$  where  $s$  is larger than two times of the maximum length of a chain, including the space among particles. The directions of the monomers conform to (2.41) for the isotropic case or (2.46) for the nematic case. That procedure was repeated for  $N_{trial} \approx 5 \times 10^8$  times and we counted the number of times that the two chains overlapped, denoted by  $N_{overlap}$ . The excluded volume is then given by:

$$v_{excl} = \frac{N_{overlap}}{N_{trial}} s^3 \quad (3.36)$$

#### Isotropic case

$A_I$ ,  $K_I$  and  $B_I$  can be estimated by a quadratic fit of the excluded volume as a function of  $l$ . For all of the models we studied in this report, the values of  $A_I$  obtained from the fitting, which is indicated by the fits of  $v_{excl}(l)$  at  $l = 0$ , are rather small. Hence it is reasonable to neglect this term when deducing the expression of the average chain length, as is done in the theory.

#### Nematic case

The method discussed in this part was proposed in Ref. [62]. One assumes that  $A_N(\alpha) \approx 0$ , as in the case of isotropic chains and  $K_N(\alpha) = K_N^{HC}(\alpha)$ , where  $K_N^{HC}(\alpha)$  is the parameter corresponding to the end-midsection contribution of the excluded volume of 2 hard cylinders, which has been computed in Ref. [62]:

$$K_N^{HC}(\alpha) \approx \frac{4}{\pi} \sum_{i=0}^4 \frac{d_i}{\alpha^i} \quad (3.37)$$

with  $d_0 = 3.0846$ ,  $d_1 = -4.0872$ ,  $d_2 = 9.0137$ ,  $d_3 = -9.009$  and  $d_4 = 3.3461$ .

In analogy to  $K_N$ ,  $B_N$  can be written as a series of  $1/\alpha^{1/2}$ :

$$B_N(\alpha) = \frac{\pi}{4} D^3 \left( \eta_1 + \frac{\eta_2}{\alpha^{1/2}} + \frac{\eta_3}{\alpha} \right) \quad (3.38)$$

### 3.4. Computation of parameters required by theory

---

By fitting the numerical values of  $v_{excl}$  as a function of 2 variables  $l$  and  $\alpha$  as in equation (3.35), one obtained the values of  $\eta_k$  where  $k = 1 - 3$ , and hence the analytical expression of  $v_{excl}$ .

#### 3.4.3 Persistence length

The persistence length measures the stiffness of an isolated polymer. It can be extracted by computing the orientational correlation function  $C_o$ :

$$C_o(i) = \langle U_{m+i} \cdot U_m \rangle \quad (3.39)$$

where  $i$  or  $i + m$  are indexes of monomers within a chain;  $U_i$  is the unit vector representing the direction of the  $i$ th monomer. The brackets indicate an ensemble average, i.e. taken over all configurations and values of  $m$ .  $C_o(i)$  decays as  $\exp(-i/l_p)$ . Thus to compute persistence length, one needs to generate a large number of chains of random configurations (i.e. directions and positions of particles within a chain) and take average of the quantity  $U_{m+i} \cdot U_m$ . After computing the orientational correlation function with average taken over those chains, one fits it exponentially to get  $l_p$ .

For each model in this study the number of random configurations generated were  $N_c \approx 10^6 - 10^7$  and chain lengths were set at 50.

# Chapter 4

## Simulation results and comparison to theory

### 4.1 Hard cylinders with 2 spherical patches at bases

#### 4.1.1 Model

Our model of sticky hard cylinder is based on the short DNA duplexes. We investigated hard cylinders of diameter  $D$  and length  $L = 2D$ . To account for the end-to-end stacking, each cylinder is decorated with 2 attractive spherical patches of diameter  $\delta = 0.125D$  and whose centers lie on the axis of the cylinder, at the distance of  $h = 0.075D$  away from the nearest base. The patches on different cylinders interact via a square-well potential:

$$V_{SW}(x) = \begin{cases} -u_0 & \text{if } r \leq \delta \\ 0 & \text{if } r > \delta \end{cases} \quad (4.1)$$

where  $x$  is the distance between the centers of 2 patches. An illustration of the model is provided in figure 4.1.

The chosen geometry of the patches ensure no branching can occur. The estimation

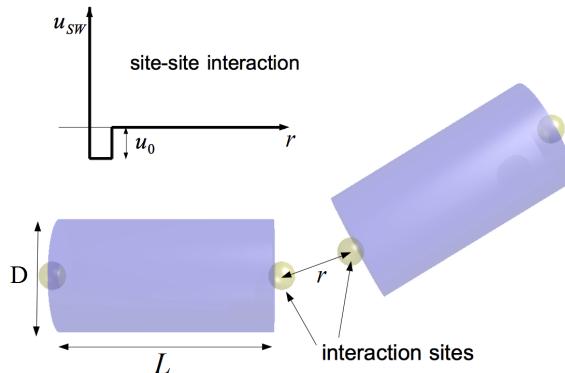


Figure 4.1: An illustration of the model being investigated. Particles are hard cylinders of diameter  $D$  and length  $L = 2D$  with two interaction sites on their axes, represented by the yellow spheres. 2 particles are bonded when 2 spheres of them overlap.

of the persistence length of the model is around 11 monomers,  $\approx 22D$ , comparable

with the persistence length of B-DNA measured in experiments,  $\approx 50\text{nm}$  [32]. The volume of a particle of this model is given by

$$v_d \equiv v_{hc} = \frac{\pi LD^2}{4} \quad (4.2)$$

### 4.1.2 Checking overlap between 2 hard cylinders

We based on the outline provided in Ref. [63] to build an algorithm to check if 2 hard cylinders overlap. Denote  $\mathbf{C}_i$ ,  $i = 1, 2$  the centers of mass of the cylinders,  $\mathbf{O}_{i,a}$ ,  $a = 1, 2$  the centers of their bases, and  $\mathbf{u}_i$  the unit vector indicating their directions. The outline of the steps is as follows:

1. Check if the 2 cylinders are parallel, i.e.  $\mathbf{u}_1 = \pm \mathbf{u}_2$ . In case they do, they overlap if and only if the 2 following condition fulfil:  $|\mathbf{u}_1 \cdot \mathbf{C}_1 \mathbf{C}_2| < L$  and  $|\mathbf{u}_1 \cdot \mathbf{C}_1 \mathbf{C}_2| < D$ .
2. Spherocylinders overlap.

We encapsulated each cylinder in a spherocylinder of the same diameter and length, constructed by adding to its bases 2 hemispheres (figure 4.2). Checking overlap between the 2 spherocylinders is a cheap calculation and if they do not overlap then the 2 cylinders too do not.

To do this first one computes the distance between the lines 2 containing

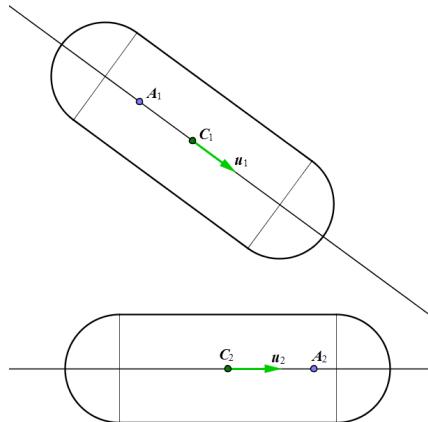


Figure 4.2: 2 cylinders are encapsulated in 2 spherocylinders of the same length and diameters.  $\mathbf{C}_i$  are centers of the cylinders;  $\mathbf{u}_i$  are the unit vector indicating their directions.  $\mathbf{A}_i$  are arbitrary points on the axes of the 2 cylinders.

their axes. Any couple of points  $A_i$ ,  $i = 1, 2$  on these 2 lines can be written as  $\mathbf{A}_i = \mathbf{C}_i + t_i \mathbf{u}_i$ . Minimizing the distance between them with respect to  $t_1$  and  $t_2$  yields the distance between the lines. The results are:

$$t_1^* = \frac{X - KY}{1 - K^2} \quad (4.3)$$

$$t_2^* = -\frac{Y - KX}{1 - K^2} \quad (4.4)$$

$$(\mathbf{A}_1 \mathbf{A}_2)_{min}^2 = \mathbf{C}_1 \mathbf{C}_2^2 + 2t_2^* Z - 2t_1^* Y - 2t_1^* t_2^* K + (t_1^*)^2 + (t_2^*)^2 \quad (4.5)$$

where one defines:

$$\begin{aligned} K &= \mathbf{u}_1 \cdot \mathbf{u}_2 \\ X &= \mathbf{u}_1 \cdot \mathbf{C}_1 \mathbf{C} \\ Y &= \mathbf{u}_2 \cdot \mathbf{C}_1 \mathbf{C} \end{aligned} \quad (4.6)$$

If  $d \geq D$  the two spherocylinders and hence the 2 cylinders do not overlap. Otherwise one needs to check if the 2 points  $\mathbf{A}_{i,j}^*$  corresponding to the minimized distance both lie inside the 2 cylinder rims, in which case the 2 cylinder rims overlap. This happens if the 2 following conditions satisfy:

$$|t_1^*| < L/2 \quad \text{and} \quad |t_2^*| < L/2 \quad (4.7)$$

If not, either the 2 rims do not overlap, or the overlap also involves a disk-disk or disk-rim overlap, in either case one needs to continue to next step.

### 3. Disk-disk overlap

If the 2 disks overlap, obviously their intersection must lie on the intersection line between the 2 plane containing them. Therefore one begins by checking if the intersection line is a secant to both the 2 disks: if it does not intersect at least 1 of the disks at 2 points, one can conclude that the 2 disk do not overlap. Otherwise, denote the intersection points  $\mathbf{B}_{i,a}$  where  $i$  refers to the disks and  $a$  refers to any of the 2 points on the same disk. Their possible relations in space are illustrated in figure 4.3.

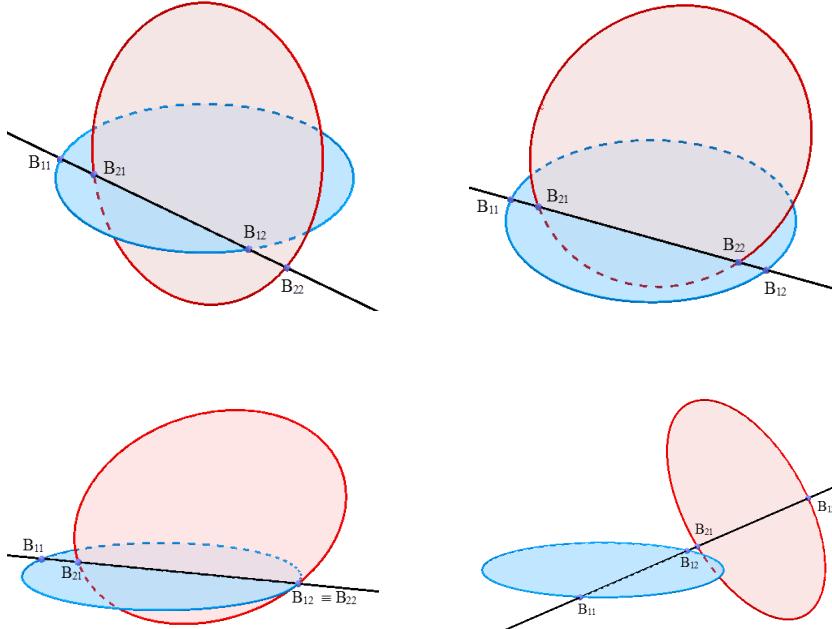


Figure 4.3: Possible relative positions of the intersection points of the circumferences of the disks and the intersection line of the 2 planes containing the 2 disks

The 2 disks overlap if and only if for at least one point  $\mathbf{B}_{i,a}$  one has:

$$\mathbf{B}_{i,a} \mathbf{B}_{j,1} \cdot \mathbf{B}_{i,a} \mathbf{B}_{j,2} < 0 \quad (4.8)$$

The 2 different subscripts  $i$  and  $j$  implies 2 different disks.

#### 4. Disk-rim overlap.

This step is the most expensive one. We used the algorithm suggested in Ref. [64] here, with some alteration.

The idea of the algorithm is using an iteration scheme to determine the shortest distance between a point on a disk (on cylinder 1) and a point on the axis of the second cylinder. Denote  $\mathbf{A}_2$  a point on the axis of the second cylinder and  $\mathbf{T}_1$  the point on the circumference of the disk that is nearest to  $\mathbf{A}_2$ .  $\mathbf{T}_1$  can be determined by minimizing  $(\mathbf{T}_1 - \mathbf{A}_2)^2$  with the following constraints:

$$(\mathbf{T}_1 - \mathbf{O}_1)^2 = \frac{D^2}{4} \quad \text{and} \quad (\mathbf{T}_1 - \mathbf{O}_1) \cdot \mathbf{u}_1 = 0 \quad (4.9)$$

which yields 2 solutions:

$$\mathbf{T}_1^\pm = \mathbf{O}_1 \pm \frac{D \mathbf{A}_2 - \mathbf{O}_1 - ((\mathbf{A}_2 - \mathbf{O}_1) \cdot \mathbf{u}_1) \cdot \mathbf{u}_1}{\sqrt{((\mathbf{A}_2 - \mathbf{O}_1) \times \mathbf{u}_1)^2}} \quad (4.10)$$

Denote  $\mathbf{T}_\perp^\pm$  and  $\mathbf{T}_{\parallel}^\pm$  the absolute value of the perpendicular and parallel projections of  $(\mathbf{T}_1^\pm - \mathbf{C}_2)$  on  $\mathbf{u}_2$ . The next points  $\mathbf{T}_1$  and  $\mathbf{A}_2$  in the iteration are given by:

$$\begin{aligned} \mathbf{T}_1^{\text{new}} &= \begin{cases} \mathbf{T}_\perp^- & \text{if } \mathbf{T}_\perp^- < \mathbf{T}_\perp^+ \\ \mathbf{T}_\perp^+ & \text{if } \mathbf{T}_\perp^+ < \mathbf{T}_\perp^- \end{cases} \\ \mathbf{A}_2^{\text{new}} &= \mathbf{C}_2 + \mathbf{T}_{\parallel}^{\text{new}} \cdot \mathbf{u}_2 \end{aligned}$$

Once the iteration convergence condition has been met, the disk and rim can be determined to overlap if and only if the following conditions are met:

$$\mathbf{T}_\perp < D/2 \quad \text{and} \quad \mathbf{T}_{\parallel} < L/2 \quad (4.11)$$

In Ref. [64] the authors tested the algorithm with the iteration convergence condition set at 5 decimal places. As for the choice of initial point, they chose  $\mathbf{A}_2 = \mathbf{C}_2$ . On the other hand we set the convergence condition at  $(\mathbf{T}_1^{\text{new}} - \mathbf{T}_1)^2 \leq 1e-14$  and at first chose a random point on the cylinder axis as initial point. While implementing the algorithm we noticed that sometimes the convergence point of the scheme depended on the initial point. So we chose initial point differently: we considered 32 points equally placed on the disk circumference and among them looked for the point that was nearest to the axis and the point on the axis nearest to that point was taken as the first point.

### 4.1.3 Computed excluded volume, persistence length and bonding volume

#### Generation of a particle bonded with an available particle

To compute the excluded volume and persistence length one needs to generate chains of particular length, by adding one particle after another. A straightforward

scheme would be repeatedly generate a new particle of random direction and position in the proximity of the destination particle until they are bonded. However, as the bonding volume of the models of interest are very small (values will be given later), this method would be not very effective. A more effective method is putting the new particle directly in a bonded configuration with the destination particle. The description of the method is provided below.

1. Generate the patch of the new particle which is to be bonded with the destination particle.

Denote the center of the destination patch  $\mathbf{A}_d$ . The center of the patch of new particle to be bonded with the destination patch, denoted by  $\mathbf{A}_n$ , then has to be within the sphere of radius  $\delta$  and centering at  $\mathbf{A}_d$ .  $\mathbf{A}_n$  is given by:

$$\mathbf{A}_n = \mathbf{A}_d + \left( r^{1/3} \delta \right) \mathbf{v} \quad (4.12)$$

Here  $\mathbf{v}$  is a random unit vector and  $r$  is a random number in the range  $[0, 1]$ .

2. Generate another unit vector  $\mathbf{u}_n$ , which is to be the direction of the new particle.

In the isotropic case,  $\mathbf{u}_n$  is simply a random unit vector. In the nematic case one wants the distribution of the angle  $\theta_1$  formed by  $\mathbf{u}_n$  and the nematic director to conform to a specific distribution function  $f(\theta)$  (the usual dependence on the azimuthal angle does not present thanks to the cylindrical symmetry of the nematic phase). It can be done with the following steps:

- (a) Generate a random unit vector  $\mathbf{v}$ . The angle it forms with the director  $\mathbf{z}$  is determined as  $\theta_1 = \arccos(\mathbf{v} \cdot \mathbf{z})$ .
- (b) The probability that the direction of a particle of the chain form an angle  $(\theta)$  with the nematic director is

$$p(\theta)d\theta = f(\theta)\sin(\theta)d\theta \quad (4.13)$$

which means that the probability density is  $p(\theta) = f(\theta)\sin(\theta)$ .

Generate a random number  $r$  in the range  $[0, 1]$ . If  $r < p(\theta_1)$  one accepts the generated vector. Otherwise it is rejected and one returns to step (2a). If the nematic angle distribution  $f$  can be written as a function of only  $\cos \theta$ , as in the case of the Onsager distribution we chose in our study:

$$f_O(\cos \theta) = \frac{\alpha}{4\pi \sinh \alpha} \cosh(\alpha \cos \theta) \quad (4.14)$$

then the probability distribution in equation (4.13) can be rewritten as

$$p'(\cos \theta)d(\cos \theta) = f(\cos \theta)\sin(\theta)d\theta = -f(\cos \theta)d(\cos \theta) \quad (4.15)$$

In this case one only needs to generate random number  $x$  in the range  $[-1, 1]$  and check if  $x < f(\mathbf{v} \cdot \mathbf{z})$ . This will be slightly faster as one does not have to compute the trigonometric functions and their reverses, which are more expensive.

3. Once one has determined the position of one patch and the direction of the new particle, given the geometry of the particle, one can easily determine its center.

## Results

To compute the excluded volume in the isotropic case we generated  $10^8$  configurations of 2 chains of length  $1 - 8$ , whereas in nematic case  $5 \times 10^8$  configurations were generated for 4 different values of  $l$  with  $\alpha$  ranges from 10 to 30 with step 5. The simulation results and fittings are presented in figure 4.4.

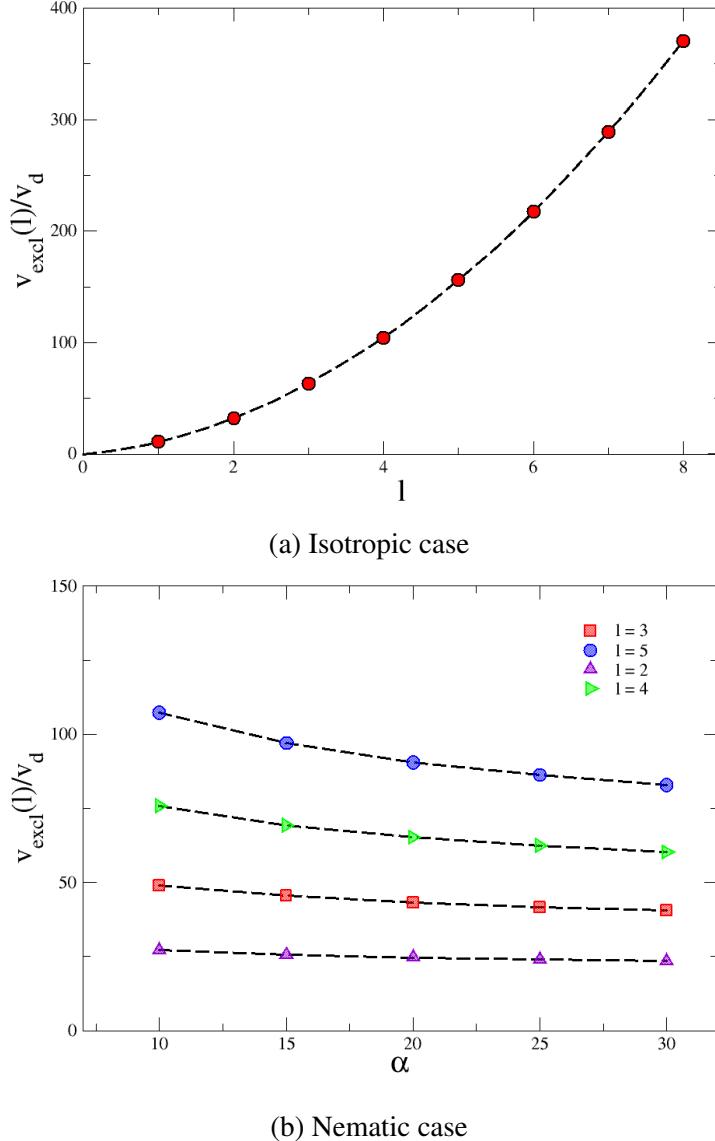


Figure 4.4: Numerical estimation and fitting of  $v_{excl}$  of sticky hard cylinders as a function  $l$  (isotropic case) and  $\alpha$  (nematic case). Symbols are simulation results and dashed lines are fits according to equation (3.35), using expressions (3.37) for  $K_N$  and (3.38) for  $B_N$ .

The results of the orientational correlation function is presented in figure 4.5. To obtain this function around  $2 \times 10^6$  chains of length 50 were generated. By fitting this to an exponential function we got  $l_p = 11.62$  (monomers). As for bonding

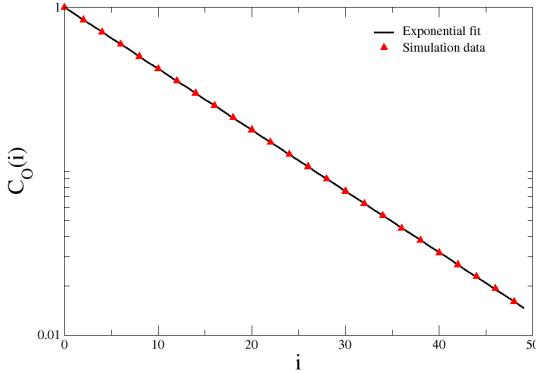


Figure 4.5: Correlation function as defined in (3.39). Symbols are simulation results and line is exponential fit, from which the persistence length can be extracted.

volume, after generating  $2 \times 10^8$  configurations we determined  $V_b/v_{hc} = 1.706 \times 10^{-3}$ .

#### 4.1.4 Simulation results

We computed the isotropic-nematic phase coexistence lines exploiting SUS and Kofke techniques. SUS simulations were performed at 4 temperatures  $T^* = 0.120, 0.135, 0.149$  and  $0.158$ . The points at  $T^* = 0.149$  was used as starting points for Kofke integration, while other points served as check points. The integration was taken over  $\beta$  with the step  $h = 0.2$ , using formula listed in table 3.1.

As discussed in section 3.2.3, all the boxes used in SUS are elongated along  $x$  axis while the nematic director is along  $z$  axis. The number of particles in the systems ranges from  $1750 - 2850$ , depending on the temperature. For each temperature the range of particles for which  $P(N)$  is computed is around 400. The details are reported in table 4.1.

As an illustration for the points discussed in section 3.2.3, we provide in figure

Table 4.1: Details of box sizes and number of particles in SUS simulations. The motivation of using elongated boxes are explained in section 3.2.3

$T^*$	$L_x/D$	$L_y/D$	$L_z/D$	$N$
0.120	52.5	9.5	29.5	2400 - 2850
0.135	47.0	8.0	26.0	1750 - 2150
0.149	47.0	8.0	26.0	1930 - 2440
0.158	47.0	8.0	26.0	2000 - 2500

4.6 a snapshot of a system simulated using SUS where the two phases coexist, as

#### 4.1. Hard cylinders with 2 spherical patches at bases

---

well as two of the reweighted  $P(N)$ s obtained by SUS. In the snapshot one can see clearly two parallel interfaces, a consequence of periodic boundary condition. Thanks to the elongated shape of the boxes the interfaces between the two phases are stable: they establish parallel to the nematic director  $\mathbf{z}$ , which minimizing the free energy barrier, and the long box size along  $\mathbf{x}$  axis increases the distance between them, therefore reducing their interaction. As a result, two peaks are well defined and separated by a flat region, as can be seen in 4.6b. We also made the box along  $\mathbf{y}$  axis shorter to reduce the number of particles needed to handle. As for

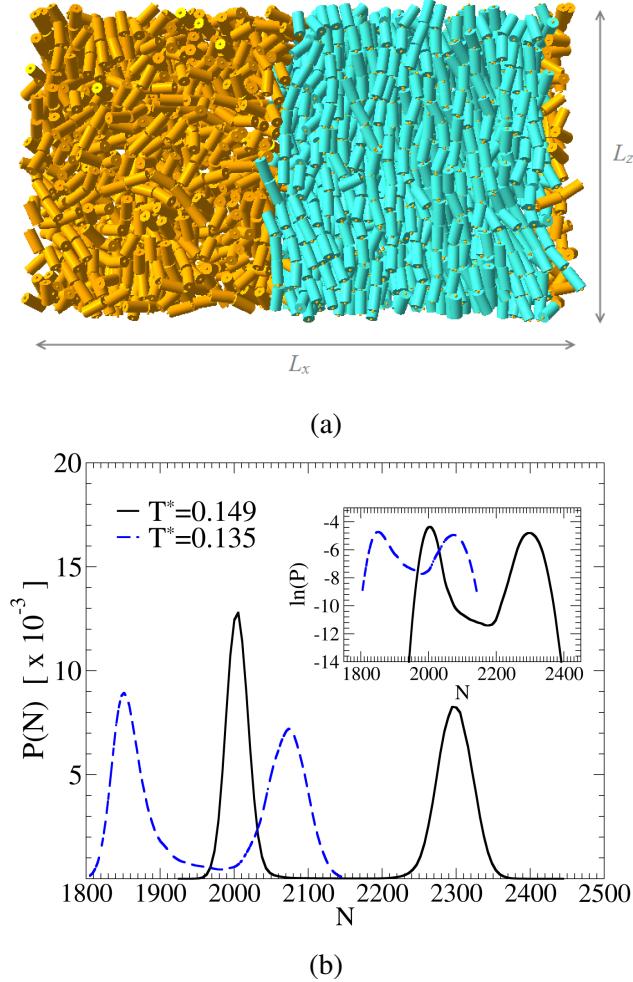


Figure 4.6: (a) A snapshot of a system where coexistence occurs, obtained in SUS simulations. To guide the eye, the two phases are roughly distinguished by different colours: orange for isotropic phase and cyan for nematic phase. There are 2 interfaces between the phases in the system, a consequence of periodic boundary condition. The snapshot is taken at  $T^* = 0.149$  and the system contains  $N = 2150$  particles. (b) Some of the reweighted probability distributions of number of particles obtained in SUS simulations. A clear flat region separating the two peaks is the result of stable interfaces, making possible by elongated shape of the boxes.

Kofke, we used cubic boxes for isotropic systems and elongated boxes for nematic ones. Naturally the longest edge is the one along the nematic director  $\mathbf{z}$ . When implementing Kofke we noticed that the result was not very sensitive to the precision of enthalpy and volume per particle of the systems in previous steps. Therefore at

each temperature, after estimating the packing fractions of 2 phases by NPT simulation, we switched to NVT simulations to collect statistics of average chain lengths of the systems, as NVT simulation is faster. For both NPT and NVT simulations, the box sizes of isotropic systems are  $L_x = L_y = L_z \approx 8D$  and those of nematic ones are  $L_x = L_y \approx 10D$  and  $L_z \approx 40D$ .

The phase boundary computed by simulation and theory are shown in figure 4.7. As

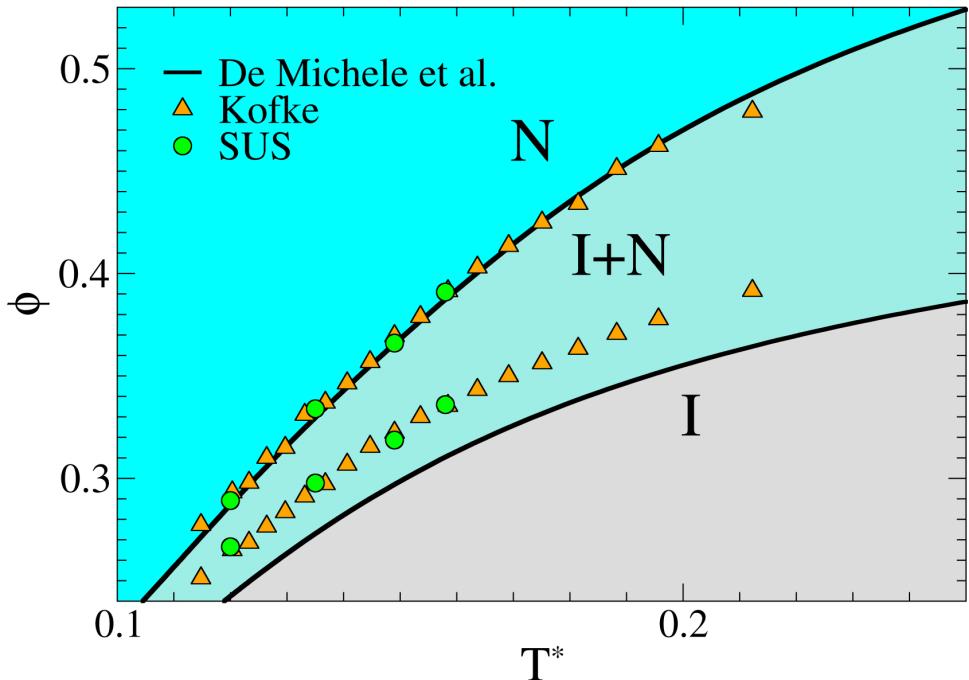


Figure 4.7: Numerical and theoretical computation of phase boundary of hard cylinders with two attractive sites at their bases. Solid lines are prediction according to the theory proposed in Ref. [25]. Circles are estimations by SUS, while triangles are results of Kofke integration, which starts from SUS points at  $T^* = 0.149$ .

can be seen in the figures, the 2 methods yield highly consistent results. Considering the largely different box sizes using in the 2 methods, particularly the size along the nematic director, this indicates that when it comes to packing fraction, the finite size effect in SUS simulations is rather small. The numerical result is well captured by the theory, although the isotropic coexistence packing fraction is slightly underestimated. As expected, the packing fraction at boundary of both phases increase with temperature, as the formation of bonds is less favourable. According to both the simulation data and theory prediction, the coexistence region narrows monotonically as temperature lowers.

In systems without self-assembly that display isotropic-nematic phase transition, particles have a fixed aspect ratio that determines the packing fraction at phase boundary. On the contrary, in self-assembling systems such as the ones we studied, particles form chains of different number of monomers, whose average value  $M$  vary according to the packing fraction at phase boundary. The quantity  $X_0M$  can be used to approximate the aspect ratio of an average chain in a self-assembling system as an equivalent to the that of a particle in non self-assembling systems. In figure 4.8a we investigate the relationship between the quantity  $X_0M$  with the pack-

#### 4.1. Hard cylinders with 2 spherical patches at bases

---

ing fraction at phase boundary in our system. Interestingly, it differs quantitatively from that of non self-assembling systems. As opposed to the monotonic dependence in the latter (the longer the particles, the lower packing fraction at which the system nematicize), what we observe here is a re-entrance: as the packing fraction at the nematic boundary increases, the average aspect ratio exhibits a minimum, beyond which it increases with the packing fraction. Although the theory overestimates the

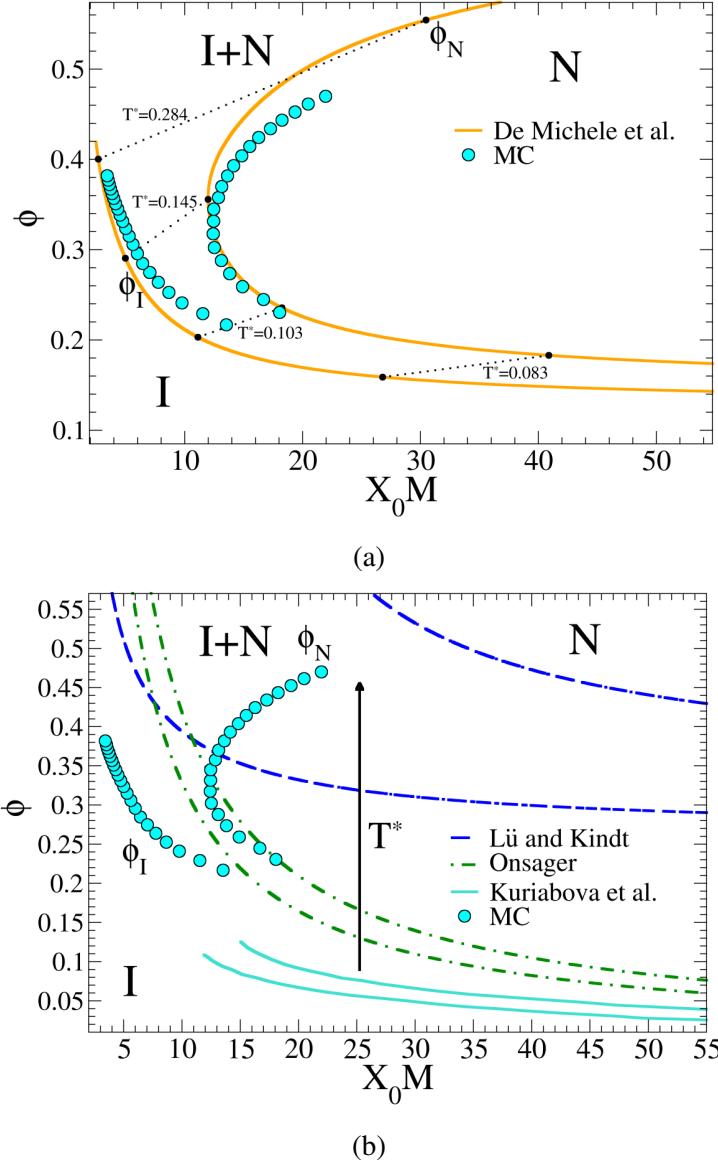


Figure 4.8: Re-entrance of the average aspect ratio  $X_0M$  as coexistent packing fraction increases. In both figures, symbols are numerical estimations while lines are theoretical predictions. (a) Comparison between simulation results and the theory proposed in Ref [25], which is represented by solid lines. Some couples of points on the isotropic and nematic phase boundary lines at the same temperature are jointed by dotted lines with the value of the corresponding temperature reported next to them. (b) Comparison between simulation results and other theories: 1-Onsager (dot-dashed lines), 2-Lü and Kindt (dashed lines), and 3-Kuriabova et al. (solid lines).

average chain length at high packing fraction, it captures the re-entrant behaviour. To understand the cause of the re-entrance one can consider the expressions (2.52) and (2.43) of  $M_N$  and  $M_I$ . At low temperature limit the nematic packing fraction at phase boundary is low and  $M_N \propto \sqrt{e^{\beta u_0}}$  goes to infinity. On the other hand at rather high temperature and hence nematic packing fraction at coexistence  $M_N \approx \sqrt{\alpha}M_I$ . While  $M_I$  approaches 1 at high temperature,  $\alpha$  quickly increases as the system becomes more ordered due to high packing fraction. That explains why  $M_N$  exhibits a minimum in intermediate range.

Next, in figure 4.8b we compared the numerical results with other available theories:

- 1 - Onsager's theory, which studies isotropic-nematic phase transitions of hard cylinders without self-assembly.
- 2 - Lü and Kindt's theory, which studies systems of hard spheres self-assembling to stiff chains (with persistence length of order hundreds to thousands of particles). The parameters are chosen to fit the present hard cylinder model.
- 3 - Kuriabova et al.'s theory, which studies self-assembling hard cylinders.

According to our computation, none of the theories predict the non-monotonic dependence of average aspect ratio on packing fraction at coexistence. Moreover, Lü and Kindt's theory largely overestimates the values of the two quantities.

Next we investigate the chain length distribution of the systems at isotropic and nematic coexistence. According to suggestion by Ref. [19] for spheres, the distribution is bi-exponential because short chains remain isotropic even when the system as a whole is nematic. On the contrary, the theory proposed in chapter 2 assumes a single exponential one. Therefore it is important to check the validity of the assumption. We plot the logarithm of chain length distribution computed by simulation and theory at different temperature in figure 4.9. As can be seen in these plots, in both isotropic and nematic systems, the distributions are well approximated by a single exponential function. Also, in the inset of figure 4.9b we plot the nematic order of subsystems containing chains of the same number of monomers. Only chains containing less than 3 monomers are isotropic and deviate from the fits.

As the nematic packing fraction at coexistence at high temperature is rather high, we computed the 3-dimensional pair distribution function  $g(\mathbf{r})$  to check if there is any transitional order presents in the systems. The expression of  $g(\mathbf{r})$  is given in (1.6).

The contour plots of  $g(\mathbf{r})$  on 2 planes  $y = 0$  (i.e. parallel to the nematic director) and  $z = 0$  (perpendicular to the nematic director) at a random chosen points on the higher half of nematic peaks are represented in figure 4.10, showing that no transitional order exists in those systems.

To conclude, we performed thorough simulations to determine the isotropic-nematic coexistence lines of self-assembling cylinders. The two techniques exploited in our study produce highly precise results comparing to each other, showing their robustness as means of locating phase boundary, overcoming the obstacle of metastability one usually have with other methods like NPT or GCMC.

The numerical data, for the first time, confirm the re-entrant behaviour of average chain length on increasing coexistence nematic packing fraction, predicted by the theory proposed in Ref. [25]. The theoretical prediction of phase boundary is also in good agreement with numerical results. This provide a benchmark in theoret-

#### 4.1. Hard cylinders with 2 spherical patches at bases

---

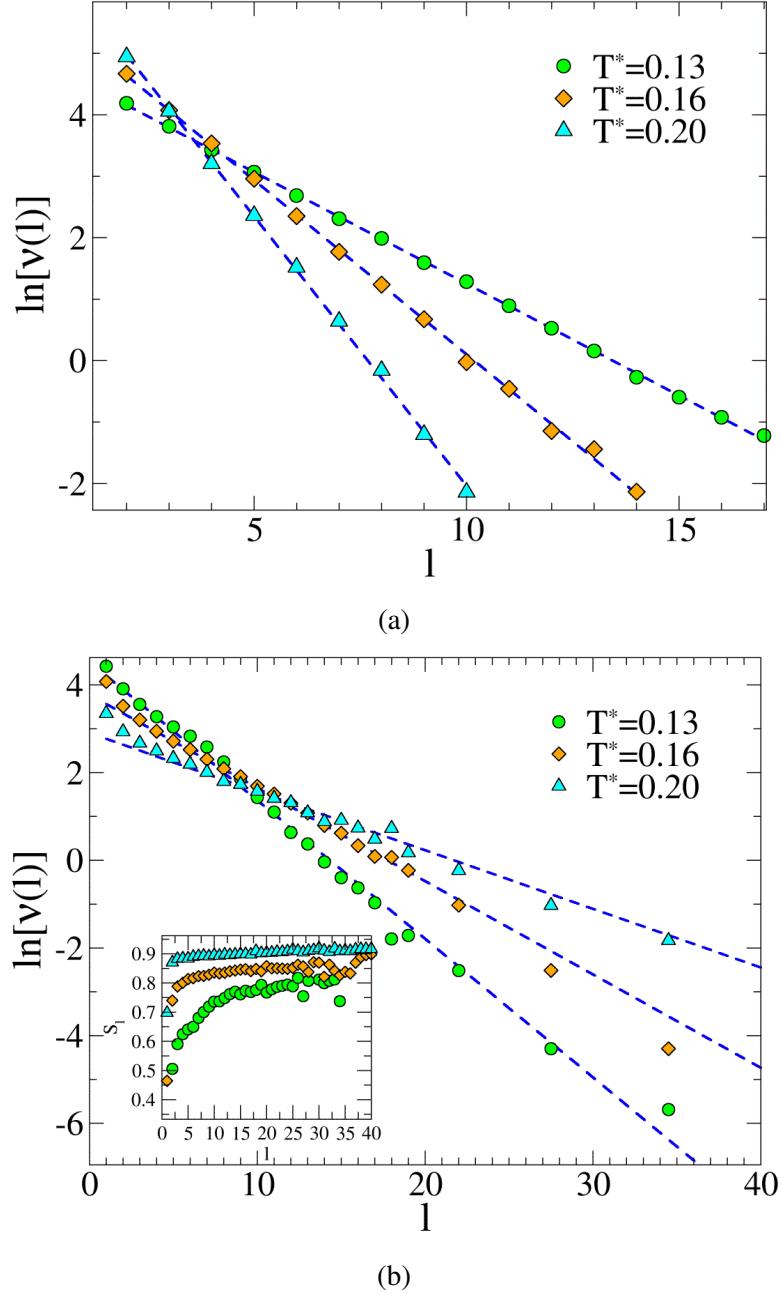


Figure 4.9: Chain length distribution at phase boundary at different temperatures of (a) isotropic systems and (b) nematic systems. The symbols are simulation data while dashed lines are exponential fits. The inset of (b) shows the nematic order as a function of chain length in nematic systems.

ical study of isotropic-nematic phase transition coupling with self-assembly. Our simple model can be a reference for a wide class of biological systems with aggregation. The values of average chain length can also be estimated in experiments via measurements of the elastic constant.

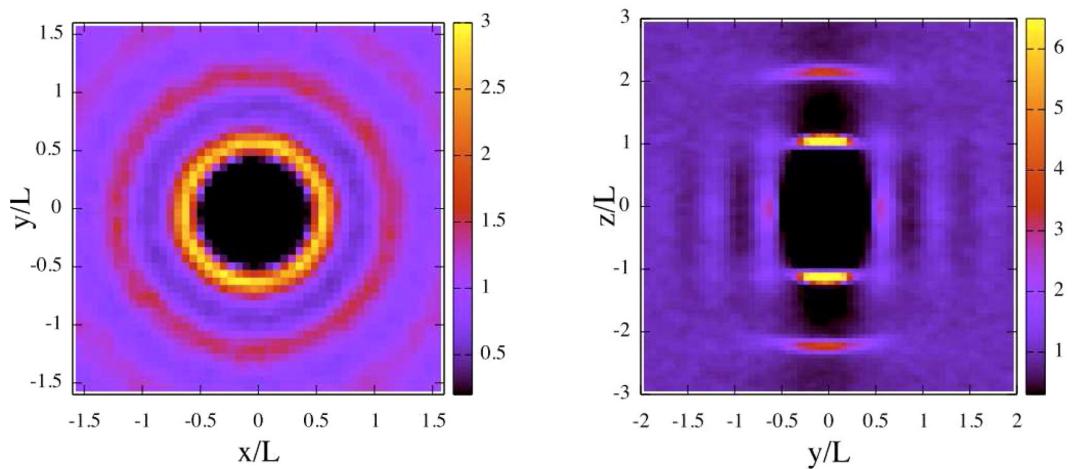


Figure 4.10: An example of 3-dimensional pair distribution function  $g(\mathbf{r})$  in the simulated systems. The data is of a system at temperature  $T^* = 0.20$  and  $\phi = 0.47$ . (a)  $g(\mathbf{r})$  on plane  $z = 0$  (perpendicular to the nematic director) and (b) on plane  $x = 0$  (perpendicular to the nematic director).

## 4.2 Bent cylinder

Experiments have shown that for short DNA duplexes ( $\leq 12$  base pairs), the isotropic-nematic phase boundary is sequence dependent. This suggests that more details of the structure of the duplexes should be taken into account in studying their phase transition. We hypothesize that the duplexes can be considered as bent core particles whose bending angle is a parameter of the process. The idea was inspired by recent findings on banana-shaped molecules. Followed by the work of Niori et al. [65] which shows that achiral bent core molecules display unusual ferromagnetic properties, much attention has been given to them from all theoretical, numerical and experimental points of view. Those systems are also the first to exhibit biaxial nematic phase [26, 27].

### 4.2.1 Model

We modelled the DNA duplexes as a bent hard cylinder, which can be either symmetric and asymmetric (to be referred to as SYBC and ASBC from now on, respectively). In both cases, the bent cylinder consists of two unaligned cylinders of the same diameter  $D$ , joining each other at one point on their axes. The bending angle, denoted  $\theta_b$ , is defined as the complementary angle of the one formed by the 2 axes. The contour length of the bent cylinder, measuring along the axes of the two component cylinders, is  $L_c$ . The contour length is the sum of two parts  $L_1$  and  $L_2$ , as shown in figure 4.11. It should be noted that the lengths of the 2 component cylinders were chosen so that there is no void between their middle bases, and therefore are slightly larger than  $L_1$  and  $L_2$ . In the case of SYBC,  $L_1 = L_2$ , while  $L_1 \neq L_2$  in ASBC case (see figure 4.11 ).

Two spherical patches near the bases of the duplexes are added to mimic stacking

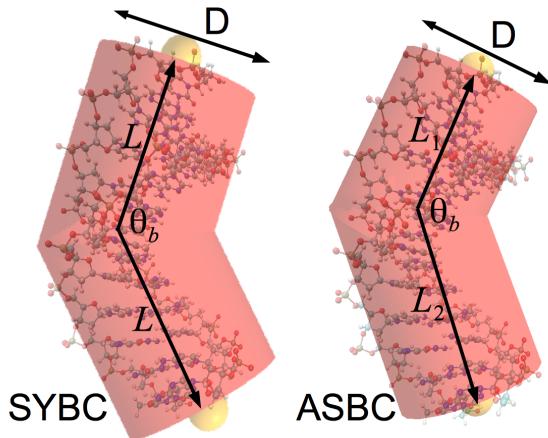


Figure 4.11: An illustration of symmetric bent cylinder (SYBC) and asymmetric bent cylinder (ASBC). Particles are hard cylinders of diameter  $D$  and length  $L$  with two interaction sites on their axes, represented by the yellow spheres.

interaction. The centers of them lie on the axes of the component cylinders at a distance  $h = 0.075D$  from the nearest bases. The diameter of the patches is  $\delta = 0.25D$ . As in the case of hard cylinders described above, these values are chosen to prohibit branching on one hand, and yield large bonding volume and reasonable persistence

length on the other hand.

### Determination of bending angle by all-atom simulations

To determine the geometry of the duplexes, we performed atomistic molecular dynamics to determine the geometry of the bent core, using both TIP4/2005 and TIP3P models of water. The sequences that we investigated are listed in table 4.2, along with their nematic concentration at coexistence obtained experimentally. The equi-

Sequence	Label	$c_N$ (mg/ml)
AATGAATTCTT	AAT	500
CGCGCCGGCGCG	allCG2	570
AATAAAATTATT	allLAT	600
AACGAATTCTGTT	AAC	620
CCGGCGCGCCGG	allCG1	670
CGCGAATTCTGCG	DD	730
ACCGAATTCTGGT	ACC	850

Table 4.2: List of the DNA sequences studied by all-atom simulations, their labels and the nematic concentration at coexistence ( $c_N$ ) from Ref. [66].

librium conformation of some of the sequences yielded by all-atom simulation is shown in figure 4.12. Note that they slightly bend at different level. In figures 4.13

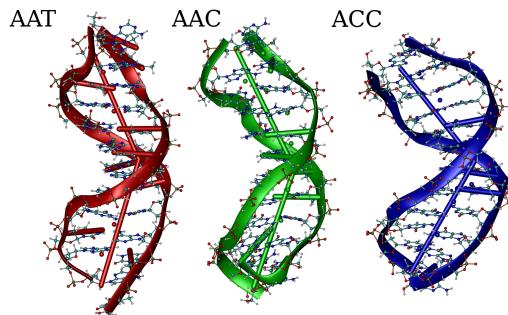


Figure 4.12: Some equilibrium conformations of 3 different sequences obtained in all-atom simulation. Note the slightly different bending angles of them.

we put together the experimental data of nematic concentration at phase coexistence of the sequences and their bending angle  $\theta_b$  which we got from all-atom simulation. The plots display an obvious correlation between the two quantities.

The detailed geometry of the sequences computed by all-atom simulation are listed in table 4.3. The data suggest that the SYBC model is a good approximation of the sequence because even if ASBC model is considered, the ratio  $L_1/L_2$  is always rather close to 1. Based on the results of all-atom simulation we opted for the model of symmetry bent cylinder of  $L = 3.6\text{nm}$  and  $D = 1.8\text{nm}$ , i.e.  $L = 2D$ .

## 4.2. Bent cylinder

---

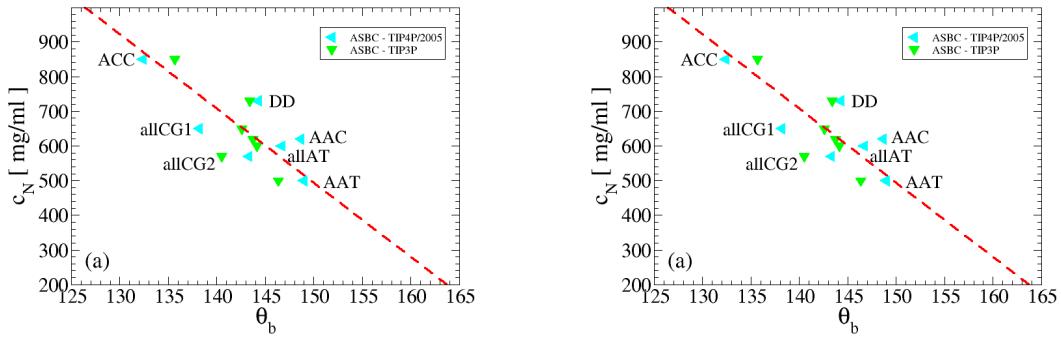


Figure 4.13: Critical nematic concentration  $c_N$ , obtained from Ref. [?], versus the bending angles  $\theta_b$  of the sequences obtained from all-atom simulations for (a) ASBC model and (b) SYBC. The red lines are to guide the eye. The label of each sequence is put next to its corresponding symbol.

TIP4P/2005		SYBC		ASBC		
Sequence	$\theta_b$	$\theta_b$	$L_1$ (nm)	$L_2$ (nm)	$L_1/L_2$	
AT	159.8	149.0	1.88	2.13	0.88	
allCG2	144.3	143.3	1.89	2.12	0.89	
allAT	151.0	146.7	1.89	2.12	0.89	
AAC	157.3	148.7	1.88	2.13	0.88	
allCG1	138.6	138.2	1.89	2.13	0.89	
DD	145.3	144.3	1.88	2.12	0.89	
ACC	130.9	132.4	1.90	2.10	0.90	

TIP3P		SYBC		ASBC		
Sequence	$\theta_b$	$\theta_b$	$L_1$ (nm)	$L_2$ (nm)	$L_1/L_2$	
AAT	152.8	146.3	1.89	2.12	0.89	
allCG2	141.3	140.4	1.89	2.12	0.89	
allAT	147.5	144.1	1.88	2.12	0.89	
AAC	147.0	143.7	1.88	2.12	0.89	
allCG1	146.0	142.6	1.89	2.13	0.89	
DD	143.3	143.4	1.88	2.12	0.89	
ACC	134.4	135.7	1.89	2.11	0.90	

Table 4.3: Parameters of the SYBC and ASBC models evaluated from the all-atom simulations for two different water-models (TIP4P/2005 and TIP3P) for the studied DNA sequences.

### 4.2.2 Simulation results of isotropic-nematic phase transition

We performed NPT simulations to locate the phase boundaries of bent cylinders of different values of  $\theta_b$  at temperature  $T^* = 0.12$ . To reduce metastability we started all the simulations from initial nematic configurations. The equations of state of three values of  $\theta_b$  was displayed in figure 4.14. The discontinuity of those lines indicates phase transition in the systems. The fact is confirmed by the jump in nematic order, which is plotted on the right panel of the same figure.

With the data obtained from NPT simulations, we represent the dependence of

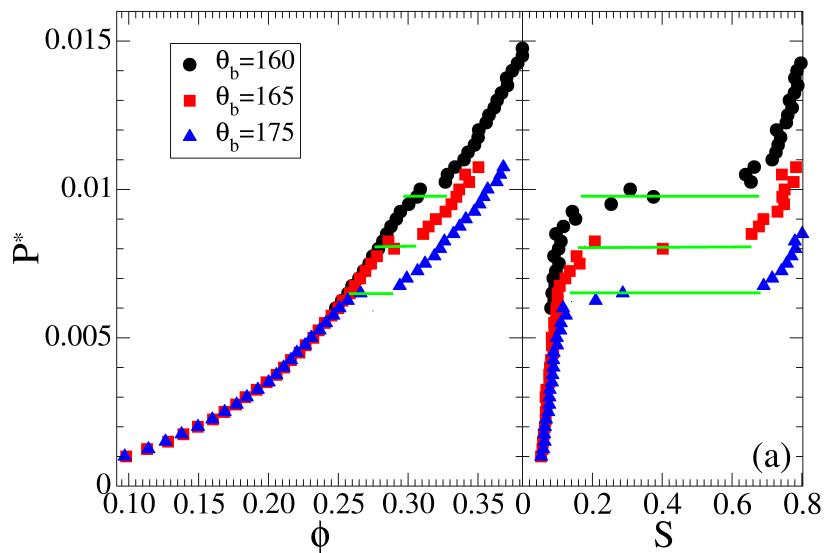


Figure 4.14: The equation of states of bent cylinders with different bending angles  $\theta_b$ , obtained by NPT simulations. The jump in packing fraction  $\phi$  indicates the transition from isotropic phase to nematic phase. It is confirmed by the sudden increase of the nematic order  $S$ .

packing fraction at phase boundary on bending angle in figure 4.15 against theory prediction. The simulation results agree reasonably well with theory, capturing both the rapid increase of the phase boundary packing fraction as well as the narrowing of the gap when bending angle decreases. This implies the potential of extending the theory for systems with smaller bending angles, for which phase coexistence occurs at packing fraction so high that can be hard to access by simulation.

Further more, to evaluate the effect of metastability on the NPT simulation results, we re-estimate the phase boundary for bending angle  $\theta_b = 160$  using SUS. The result is also shown in figure 4.15 by green circle symbols. As expected, the coexistence range of packing fraction computed by SUS is bracketed by that obtained by NPT, but the discrepancy is less than 10%.

We speculated that the steep increase of phase coexistence packing fraction stemmed from the higher flexibility of chains formed by more bending monomers. Two snapshots of systems of different bending angles at the same temperature and pressure provided in figure 4.16 give the visual impression that more bending particles, i.e. with smaller  $\theta_b$ , self-assemble to more flexible chains.

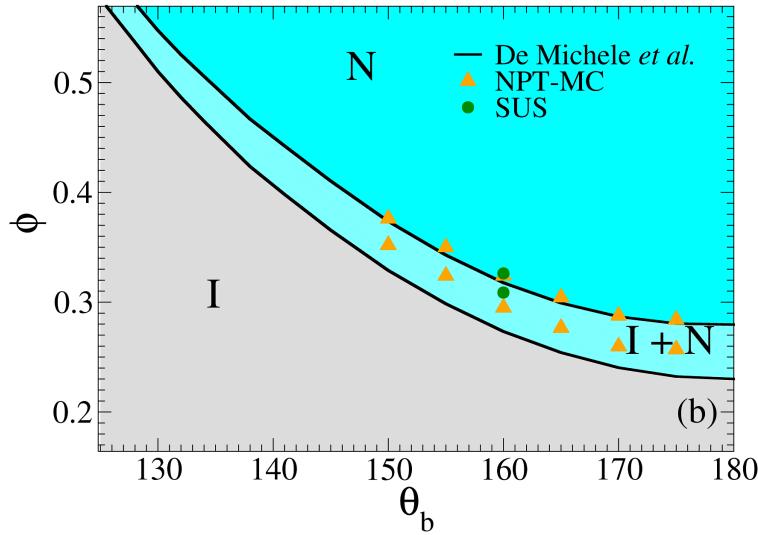


Figure 4.15: The dependence of isotropic-nematic phase boundary on bending angle, simulation (symbols) versus theory (solid lines).

To provide a proof for the hypothesis, we investigate the dependence of persis-

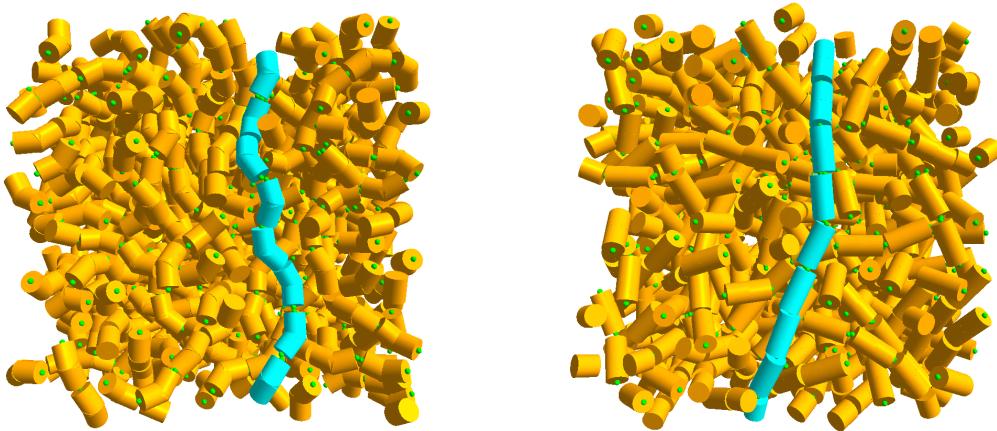


Figure 4.16: Snapshots of two systems of hard bent cylinders with different bending angles  $\theta_b$  at the same temperature and pressure, obtained from NPT simulation. The particles of higher level of bending self-assemble to chains which are significantly more flexible.

tence length  $l_p$  on  $\theta_b$ , and the theoretical phase diagram with all parameters kept unchanged except for the persistence lengths, which are replaced by the value corresponding to  $\theta_b = 180$ . The results are shown on figures 4.17.  $l_p$  shows a remarkable similar sharp increase as  $\theta_b$  decreases to the coexistence packing fraction. In addition, as the value of persistence length is replaced by that of  $\theta_b = 180$ , the packing fraction only slightly changes as  $\theta_b$  varies. The results confirm that the persistence length is the main contribution to the strongly dependence of packing fraction on  $\theta_b$ .

We also compared the theory prediction of nematic concentration at coexistence  $c_N$  as a function of bending angles with experimental data. To do it first we needed to decide on the geometrical parameters of the monomers to use in the theory. We

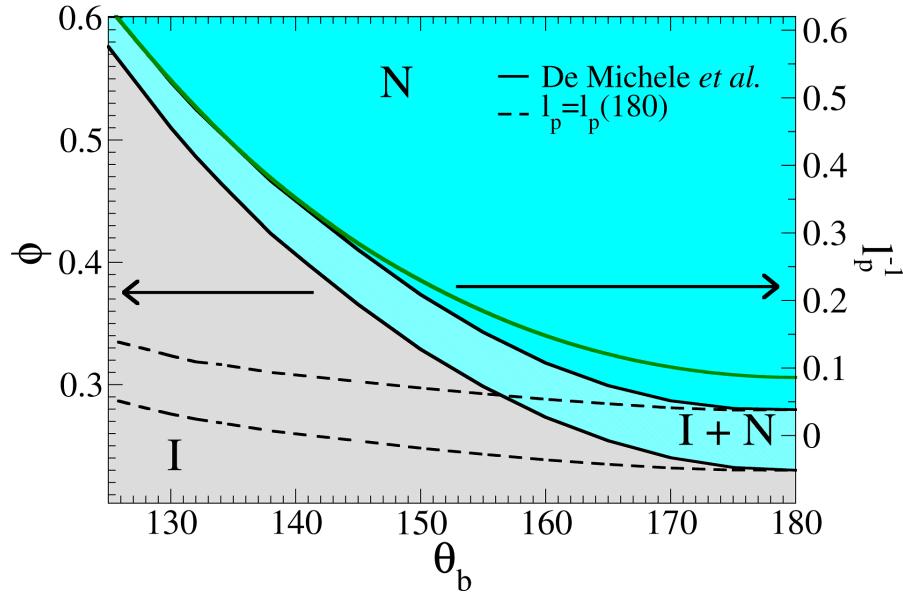


Figure 4.17: The dependence of isotropic-nematic phase boundary on bending angle, simulation (symbols) versus theory (solid lines).

set the diameter and contour length of monomer at  $1.8\text{nm}$  and  $3.6\text{nm}$ , respectively, compatible with the geometry of DNA dodecamers. As for the bending angles, the estimation by all atom simulation is taken as there are no experimental data available.

Another parameter required by the theory is the stacking free energy  $G_{ST}$ , which was obtained from experimental measures. Since there is no consensus on its value, we considered 2 limited cases  $G_{ST} = -0.9$  and  $G_{ST} = -2.5\text{kcal/mol}$ , compatible with the estimations made in Ref. [62].

The numerical and experimental evaluation of  $c_N$  versus  $\theta_b$  is reported in figure 4.18a. The experimental data with error bar falls with the grey theoretical band, again reflecting the possible range of  $G_{ST}$ . Interestingly, the grey band narrows as the bending angle decreases, indicating that for highly bent particles,  $c_N$  becomes less sensitive to  $G_{ST}$ . The explicit dependence of  $c_N$  on  $G_{ST}$  is shown in figure 4.18b. Because  $G_{ST}$  varies with temperature, this indicates that for highly bent particles  $c_N$  only weakly depends on temperature. A further experimental measurement of the dependence of  $c_N$  on temperature in these systems can provide a validation of the theory.

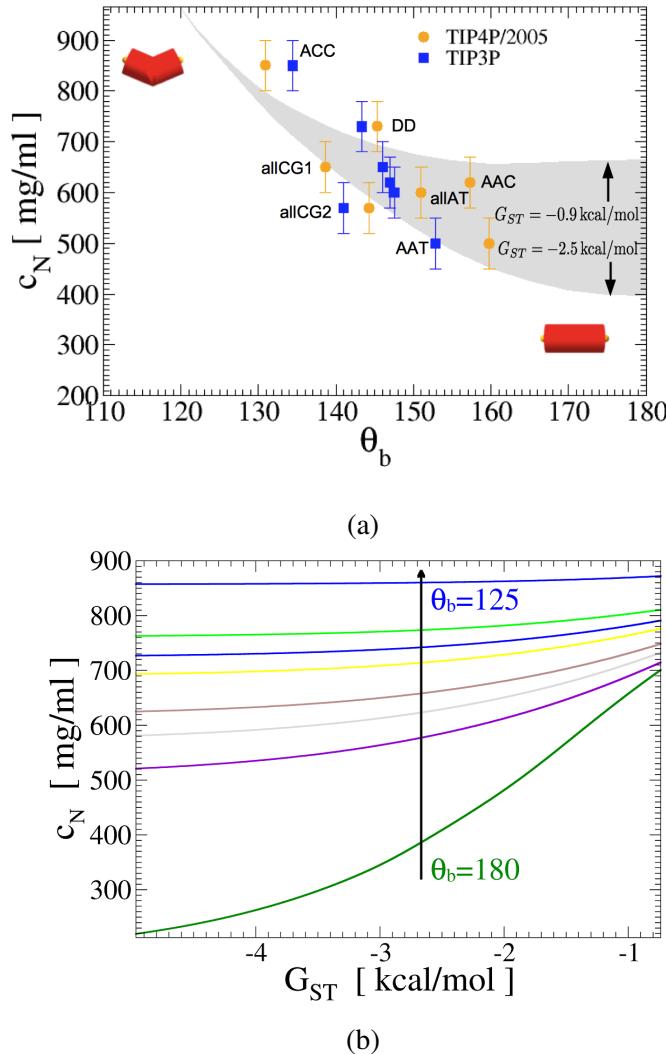


Figure 4.18: (a) The critical nematic concentration  $c_N$  obtained experimentally versus bending angle. The value of bending angles are results of all-atom simulation. (b) The dependence of  $c_N$  on  $G_{ST}$ .

## 4.3 Spheres

Hard spheres with sticky spots are interesting in the theoretical point of view to build the theory framework for self-assembling systems since its reference systems, hard spheres, have been thoroughly studied with many analytical descriptions of their properties available.

### 4.3.1 Model

In this report we study hard spheres of diameter  $D$  with 2 opposite attractive sites. The sites on two particles 1 and 2 interact via Kern-Frenkel potential [67]:

$$V_{KF}(\mathbf{r}) = \begin{cases} -u_0 & \text{if } \begin{cases} D < r < D + \delta \\ \cos(\theta_1) > \cos(\theta_{max}) \\ -\cos(\theta_2) > \cos(\theta_{max}) \end{cases} \text{ and} \\ 0 & \text{otherwise} \end{cases} \quad (4.16)$$

where  $\theta_1$  and  $\theta_2$  are the angle between the directions of the site on particle 1 and 2, respectively, and vector  $\mathbf{r}$  which roots at the center of particle 1 and points to the center of particle 2;  $\delta$  and  $\theta_{max}$  are parameters of the potential. Please refer to figure 4.19 for an illustration of the model.

One can enforce the rule of no more than 1 bond per patch by imposing the following condition for the open angle  $\theta_{max}$  for a particular value of  $\delta$ :

$$\sin(\theta_{max}) \leq [2(1 + \delta/D)]^{-1} \quad (4.17)$$

The Kern-Frenkel potential offers the flexibility to manipulate the bonding volume and persistence length independently, to some extent, with each other. To study the isotropic-nematic transition we set  $\delta = 0.546D$  and  $\cos(\theta_{max}) = 0.9462$ , which results in maximum bonding volume. In addition, we also studied the average chain length in the isotropic phase for different values of the parameters.

Please refer to figure 4.19 for an illustration of the model. For this model the

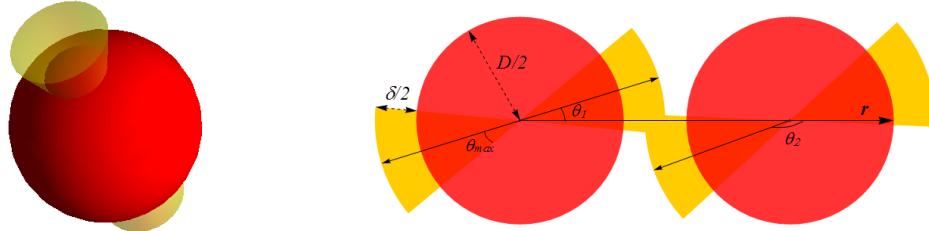


Figure 4.19: The model of hard sphere with Kern-Frenkel patches. (a) A illustration of a particle and its Kern-Frenkel patches. (b) In 2D, an example of 2 bonded particles with explanations of the parameters of the model; note that the two bonding patches (yellow spherical cones) overlap and vector  $\mathbf{r}$  joining the centers of 2 particles penetrates the caps of both patches.

volume of a monomer is given by:

$$v_d \equiv v_s = \frac{1}{6}\pi D^3 \quad (4.18)$$

### 4.3.2 Computed excluded volume, persistence length and bonding volume

As discussed earlier, when generating chains to compute excluded volume and persistence length, it is more effective to put a new particle directly in bonding region of the destination particle. The scheme to do so for Kern-Frenkel model is described below.

#### Generation of a particle bonded with an available particle

Denote the unit vector along the direction of the destination particle  $\mathbf{u}_d$ ,  $\mathbf{u}_d$  points to the destination patch, and the center of the hard sphere  $\mathbf{C}_d$ . The generation of a new particle bonded to the destination particle involves 3 steps:

1. Generate the center of the new particle,  $\mathbf{C}_n$ .

First,  $\mathbf{C}_d\mathbf{C}_n$  has to be outside of the spherical core of the destination particle. Second, it has to be within the spherical cone whose axis is  $\mathbf{u}_d$ , open angle is  $\theta_{max}$  and radius is  $(D + \delta)$ . The problem is actually similar to that in step 2 of the scheme in the case of spherical patch, the only difference is that now we have to take into account also the limit in solid angle.

Denote as  $\mathbf{a}$  the unit vector parallel to  $\mathbf{C}_d\mathbf{C}_n$ , and  $\mathbf{v}_1, \mathbf{v}_2$  2 random unit vectors which are both perpendicular to  $\mathbf{u}_d$  and to each other.  $\mathbf{a}$  is given as:

$$\mathbf{a} = \sin t (\sin \phi_1 \mathbf{v}_1 + \cos \phi_1 \mathbf{v}_2) + [r(1 - \cos \theta_{max}) + \cos \theta_{max}] \mathbf{u}_d \quad (4.19)$$

where  $\phi_1$  is a random number in the range  $[0, 2\pi]$ ,  $r$  is random number in the range  $[0, 1]$  and  $T^* = \arccos(r(1 - \cos \theta_{max}) + \cos \theta_{max})$ . After generating  $\mathbf{a}$  one still needs the norm of  $\mathbf{C}_d\mathbf{C}_n$ , denoted  $cc$ . Recall that  $cc$  has to be no smaller than the diameter of the hard spheres. With this in mind one keep generating a random number  $r$  and until its value satisfies  $r^{1/3}(D + \delta) \geq D$ .  $\mathbf{C}_n$  is then given by:

$$\mathbf{C}_n = \mathbf{C}_d + [r^{1/3}(D + \delta)] \mathbf{a} \quad (4.20)$$

2. Generate the direction of the patch of the new particle which is to be bonded to the destination patch, denoted  $\mathbf{u}_n$ .

Again, this vector has to be within the spherical cone whose open angle is  $\theta_{max}$  and radius is  $(D + \delta)$ , but whose axis is now  $\mathbf{C}_n\mathbf{C}_d$ . Therefore to generate  $\mathbf{u}_n$  we use the same scheme used to generate the direction of  $\mathbf{C}_d\mathbf{C}_n$ .

If we compute the excluded volume in the nematic case we also require that  $\mathbf{u}_n$  conforms to the angle distribution, which is done similarly to step 2 of the scheme in the case of spherical patch.

When generating chains to compute the persistence length (to be discussed in the next part), the direction of the new particle is always set to be the unit vector anti-parallel to  $\mathbf{u}_n$ . This choice ensures the consistency with how the persistence length is defined.

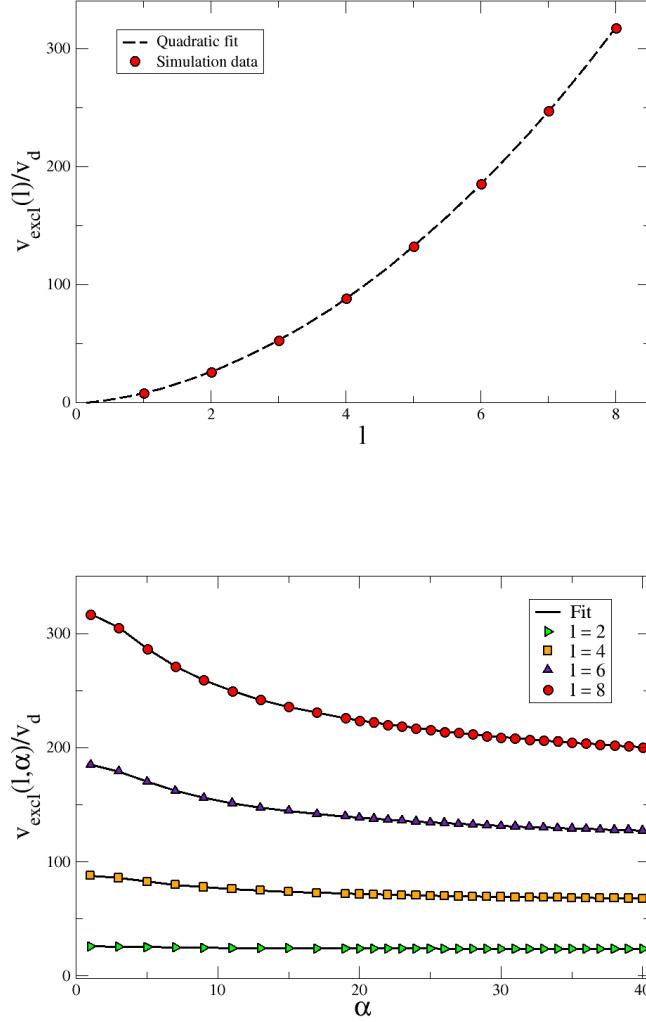


Figure 4.20: Numerical estimation and fitting of  $v_{excl}$  of sticky hard spheres as a function  $l$ . Symbols are simulation results and dashed lines are fits according to equation (3.35), using expressions (3.37) for  $K_N$  and (3.38) for  $B_N$ .

## Results

To compute the excluded volume in the isotropic case we generated  $2 \times 10^8$  configurations of 2 chains of length 2 – 9, whereas in nematic case  $1.5 \times 10^8$  configurations were generated for 4 different values of  $l$  with  $\alpha$  ranges from 1 to 40 with step 2 for  $\alpha < 20$  and step 1 for  $\alpha \geq 20$ . The simulation results and fittings are shown in figure 4.20.

As for the persistence length, we generated around  $2 \times 10^6$  chains of length 50 for each set of values of  $\delta$  and  $\theta_{max}$ . The results are plotted in figure 4.21. The data fit well to quadratic function. From the fitting function, we extrapolated the persistence length at  $\delta \rightarrow 0$  of  $l_p = 7.3$ . From now on, whenever we study systems

### 4.3. Spheres

---

of different patch sizes,  $\theta_{max}$  always takes the largest value that satisfies condition (4.17).

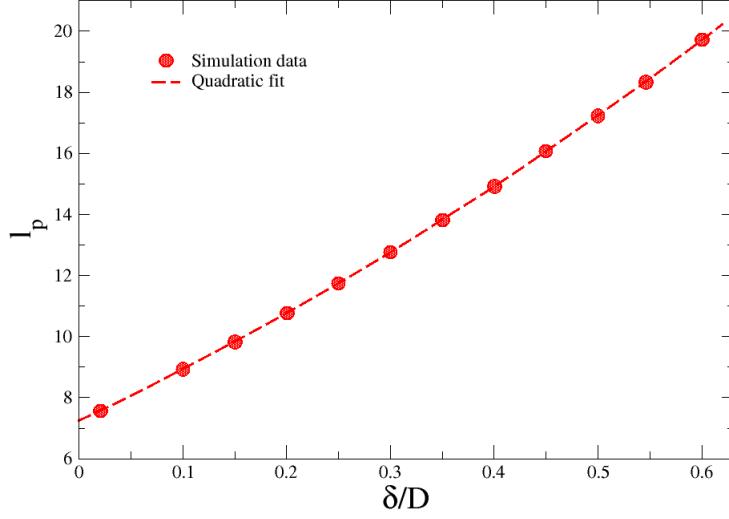


Figure 4.21: Persistence length corresponding to different sizes of the patches. Symbols are simulation data whereas dashed line is quadratic fit.

The bonding volume of sphere with Kern-Frenkel patches can be computed analytically [41]:

$$V_b/v_s = 2 \left[ (1 + \delta/D)^3 - 1 \right] [1 - \cos(\theta_{max})]^2 \quad (4.21)$$

The dependence of bonding volume on patch range is plotted in figure 4.22. It reaches the maximum at  $\delta/D \approx 0.546$ .

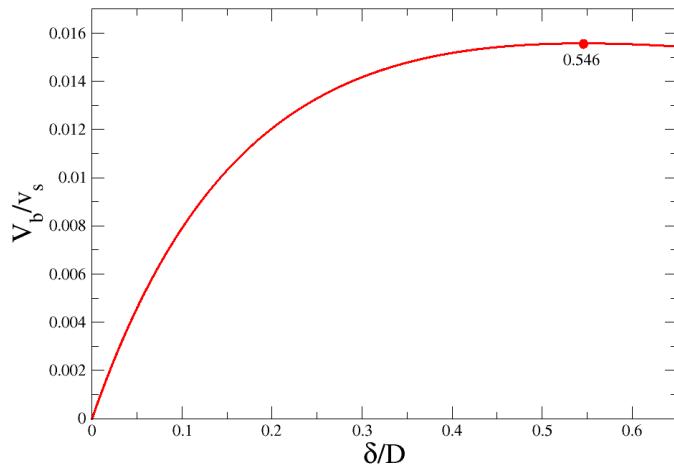


Figure 4.22: Bonding volume as a function of patch range.

### 4.3.3 Simulation results

#### Properties of isotropic phase

We were interested in how Wertheim theory works with higher packing fraction reported in [22]. In figure 4.23 the results of average chain length as a function of packing fraction in isotropic phase for different patch sizes at different temperatures  $T^* = 0.12, 0.13, 0.14$  and  $0.15$  are reported. To obtain the results, we use NVT simulations. The boxes are cubic and their sizes are around  $21D$ . The predictions of Wertheim theory and theory by De Michele et. al. are also reported for comparison.

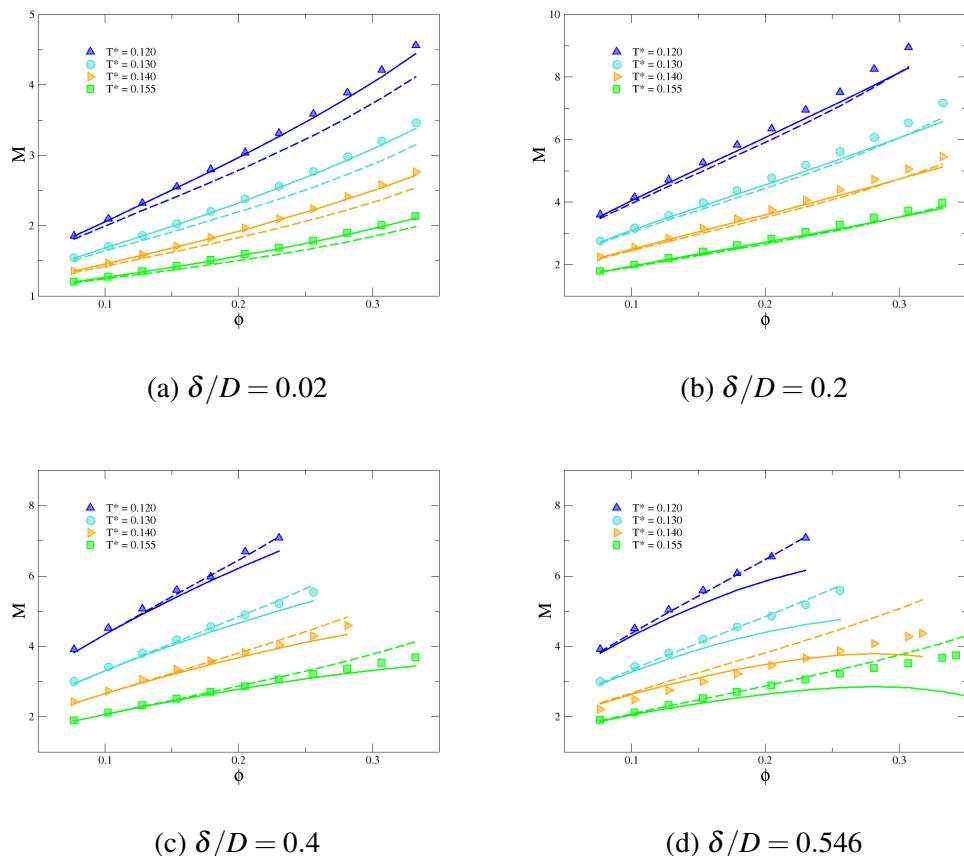


Figure 4.23: The average chain length in isotropic phase for different patch sizes at different temperatures. As discussed earlier, as the patch ranges shorten, their open angles widen accordingly to keep the bonding volume at the maximum volume without permitting multi-bonding. In all figures, symbols are simulation data, solid lines are predictions by Wertheim theory and dashed lines are predictions by De Michele et. al. theory.

As can be seen in the figures, the agreement between simulation data and theory predictions depend not only on packing fraction of the systems but also on temperature and the patch size. At the smallest patch range, Wertheim theory still works very well up to a quite high packing fraction of around 0.4 for all temperatures. The range over which the Wertheim theory is in good agreement with simulation data narrows as temperature decreases or the patch range increases. For the largest patch range studied, at high temperatures, namely  $T^* = 0.145$  and  $T^* = 0.155$ , the theory

becomes unreliable, yielding an artificial turning point of the average chain length on the increase of packing fraction.

In contrast to Wertheim theory, the theory by De Michele et al. works better as the patch range increases. Surprisingly, while at very short patch range the theory works better at higher temperatures, at larger patch ranges, its agreement with simulation seems to improve as the temperature lowers.

### Isotropic - nematic phase transition

As with hard cylinders, we computed the isotropic-nematic phase coexistence lines exploiting SUS and Kofke techniques. SUS simulations were performed at 3 temperatures  $T^* = 0.120, 0.130$  and  $0.155$ . The points at  $T^* = 0.130$  was used as initial points for Kofke integration, while other points served as check points for the integration from  $T^* = 0.130$ . The integration was taken over  $\beta$  with the step  $h_- = 0.4$  for decreasing  $\beta$  and  $h_+ = 0.2$  for increasing  $\beta$ . The reason  $h_+$  is chosen smaller than  $h_-$  is because at the first integration step up on  $\beta$  with step 0.4, we figured that the predictor-corrector seemed not convergent. The fact that the scheme converges with decreasing  $\beta$  but not with increasing  $\beta$  with the same step is consistent with the discussion in section 3.2.2 about the accumulation of error is proportional to  $\Delta v^{-1}$  and according to our theory prediction as well as the results we obtained for hard cylinder,  $\Delta v$  increases when  $\beta$  decreases. At the same time, we also use SUS point at  $T^* = 0.155$  as a starting integration point to obtain 2 more points at lower  $\beta$ , with integration step 0.2. In all integration, we use trapezoid and midpoint formula for the first and second integration step, and variant 1 listed in table 3.2 for the 3 steps and beyond.

As discussed in section 3.2.3, all the boxes used in SUS are elongated along  $x$  axis while nematic director was along  $z$  axis in the cases of  $T^* = 0.120$  and  $0.130$ ; as for  $T^* = 0.155$ , the nematic director is parallel to the diagonal vector of the  $y-z$  plane, to give more room for longer chains. In all cases, the box size was  $40 \times 16 \times 16$ . The numbers of particles in the systems range from  $4500 - 7380$ , depending on the temperature. For each temperature the range of particles for which  $P(N)$  was computed was around 500. The details are reported in table 4.4.

The phase boundary computed by simulation and theory are shown in figure 4.24.

Table 4.4: Details of number of particles in SUS simulations. The motivation of using elongated boxes are explained in section 3.2.3

$T^*$	$N$
0.120	4500 - 5020
0.130	5100 - 5630
0.155	6830 - 7380

Again, the 2 methods yields very agreeable results although the step of  $\beta$  is somewhat large. At low packing fraction the theory captures the numerical result well. However as the coexistent packing fraction increases, the discrepancy becomes significant. A possible explanation is that Parsons-Lee factor does not work well with

spheres [68].

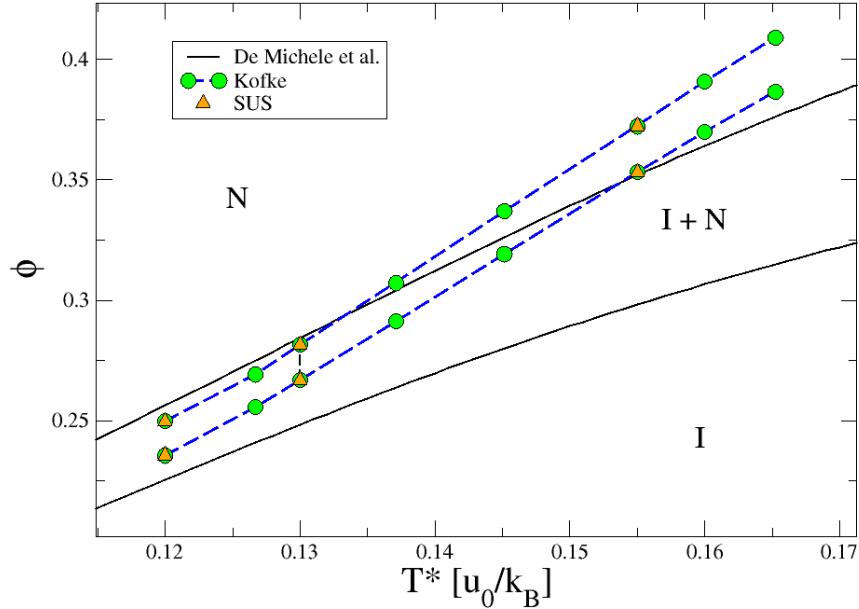


Figure 4.24: Numerical and theoretical computation of phase boundary of sticky hard spheres. Solid lines are prediction by the theory proposed in Ref. [25]. Circle symbols are estimations by SUS, while triangles are results of Kofke integration, which starts from SUS points at  $T^* = 0.130$ .

In figure 4.25 we present the dependence of coexistence pressure on temperature. The values of pressure were obtained from Kofke integration and from NPT simulation for SUS points. In the range of temperature we studied, the coexistence pressure fits well to a cubic function. Again, the results yielded by Kofke and SUS are in quite good agreement with each other.

An interesting feature of the systems at phase boundary is the distribution of chain length. Kindt et al. suggested in Ref. [19], which is then generally accepted, that in self-assembling systems of attractive hard spheres in nematic phase the chain length distribution are bi-exponential. According to Ref. [19], in nematic systems the short chains remain isotropic and the entropy loss when a particle is added to a short chain quickly increases as the chain grows longer. However Ref. [19] only investigated systems with rather stiff chains whose persistence lengths are of order of 100 or 1000 particles. Our study on the other hand provides a complement with flexible chains whose persistence lengths are around 18 particles.

The numerical results of chain length distribution in some of the isotropic we studied are presented in figures 4.26 and 4.28. In figure 4.26 we also plotted the theoretical distribution (2.37) with the value of  $\rho$  and  $M$  taken from simulation results. Apart from long chains whose statistics is poor due to their scarceness, the theoretical distribution fits extremely well to the numerical data, suggesting that the distribution proposed by (2.37) describes highly accurately the chain lengths in

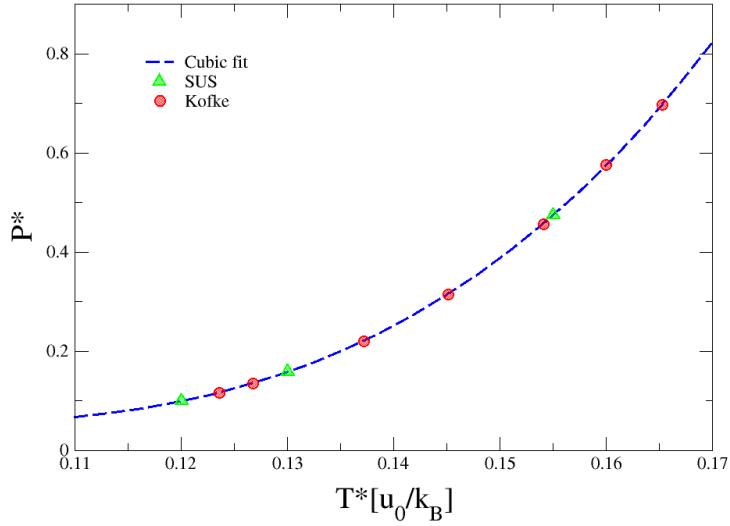


Figure 4.25: Coexistence pressure as a function of temperature. Symbols are simulation data, obtained from Kofke integration or from NPT simulation for SUS points. Dashed line is a cubic fit of the points obtained from Kofke.

isotropic systems.

As for nematic systems, to study the behaviour of their short chains, we computed the nematic order  $S_l$  of subsystems of chains of the same length  $l$ , shown in figure 4.27. For all temperatures we studied, only very short chains, around 3-mers or less, are isotropic. As expected, at higher temperature and hence higher nematic packing fraction at phase boundary, the length limit under which the chains remains isotropic decreases. At the highest temperature  $T^* = 0.155$  for example, even dimers can be considered to be nematic.

The distributions of nematic systems at coexistence are reported in figure 4.28. They slightly bended around short chain limits. The positions of the bends coincide with the region where the corresponding  $S(l)$  steeply increases, moving towards shorter chains as the number density becomes higher, which is consistent with the reason provided by Ref. [19] for the bi-exponential behaviour. However they still fit nicely to a single exponential function (when doing the fitting we excluded the very long chains, namely chains consisting of more than 100 monomers, for their poor statistics). Again, only 3-mers or shorter chains clearly deviate from the single exponential fit.

As in the case of hard cylinders, the re-entrant behaviour is also observed by both simulations and theory, as is seen in figure 4.29. The reduced temperature where the re-entrance occurs is around  $T^* = 0.130$  while the theory predicts 0.144.

It is interesting to check the appropriateness of the Onsager trial distribution 2.14. Although the function is known to work very well with non-assembling hard spherocylinders, to our knowledge, it has not been validated for self-assembling systems. We extracted the trial function from simulation results in two ways:

1. Fit the simulation data to the function. The results are shown in figure 4.30a
2. Compute the nematic order from simulations and numerically solve equation

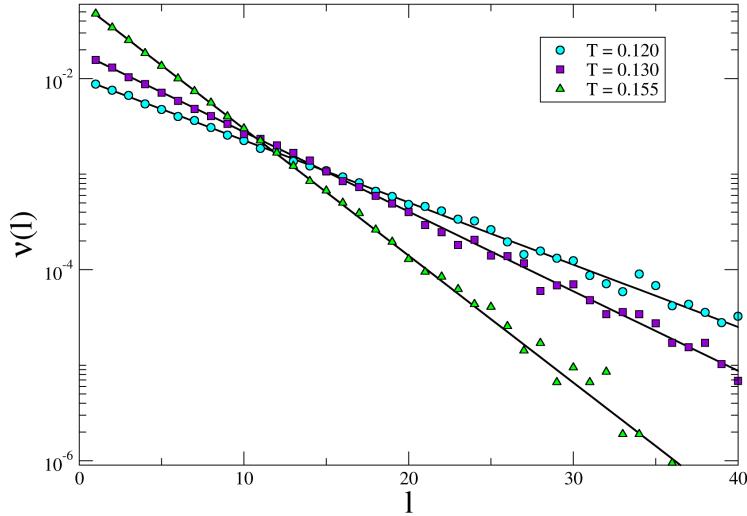


Figure 4.26: Chain number density of some isotropic systems at coexistence. Symbols are simulation data. Lines are theoretical exponential distribution according to equation (2.37), using the values of number density  $\rho$  and average chain length  $M$  obtained from simulation.

(2.16) to get  $\alpha$ , from which the function can be determined. The results are shown in figure 4.30b

As can be seen in the figures, at higher temperatures (and hence packing fractions) the results obtained from 2 methods are consistent and confirm that the trial function capture very well the angle distributions of the systems. However at lower packing fractions, the agreement deteriorates. Although the simulation data fit very well to the function form, they fit at the wrong values of  $\alpha$ . When computing the function from the values of  $\alpha$  obtained from nematic order, the results significantly deviate from simulation data. For comparison, the values of  $\alpha$  obtained by the 2 methods are plotted in figure 4.31. A possible explanation for the function working better at

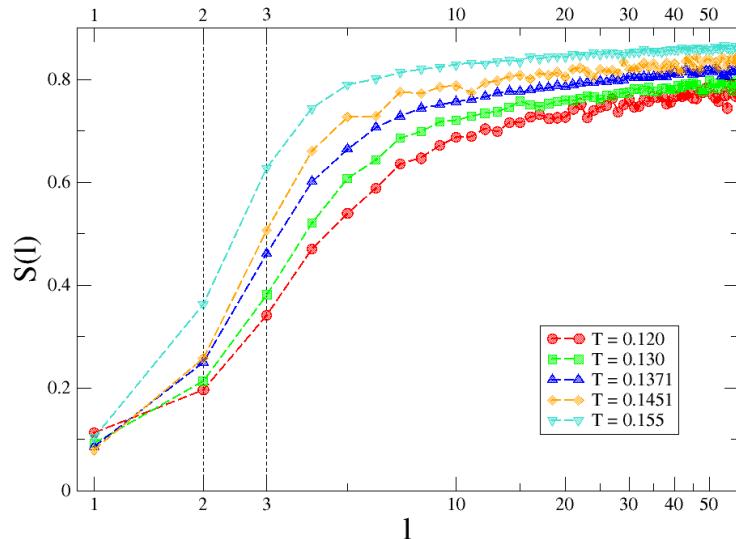


Figure 4.27: The dependence of nematic order on chain length in systems at different temperatures. Symbols are simulation data. Two dashed lines indicate the points of dimers and trimers.

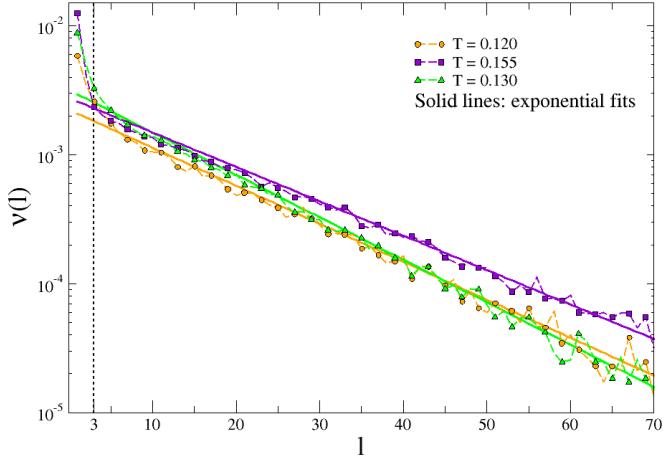


Figure 4.28: Chain number density of some isotropic systems at coexistence. Symbols are simulation data. Lines are single exponential fits. Long chains of more than 100 monomers are excluded in the fitting as their statistic are rather poor due to scarceness.

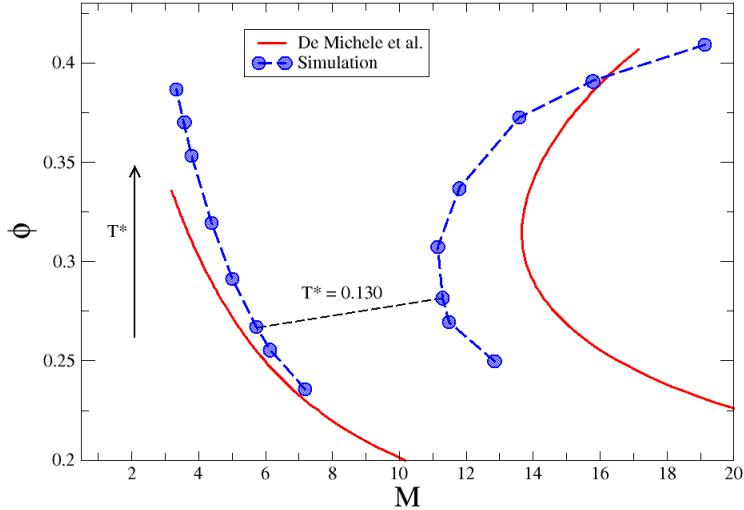


Figure 4.29: The re-entrance behaviour of sticky hard spheres. Red circles are numerical results. Solid lines are estimates by theory [25].

higher packing fraction is that at higher packing fractions the chains are less flexible and their shapes approximate straight hard rods more closely.

Figure 4.31 also confirms that  $\alpha$  increases steeply at high packing fraction, consistent with our argument that it is the large values of  $\alpha$  at high packing fraction that leads to re-entrance behaviour.

Finally, we check that at high packing fraction the systems are still in nematic phase, by computing the pair distribution function defined in equation (1.6). The result for a configuration at  $T^* = 0.1652$ , the highest temperature we studied, is shown in figure 4.32. As can be seen in the figure, there is no transitional order occurs in the system in both planes parallel and perpendicular to nematic director.

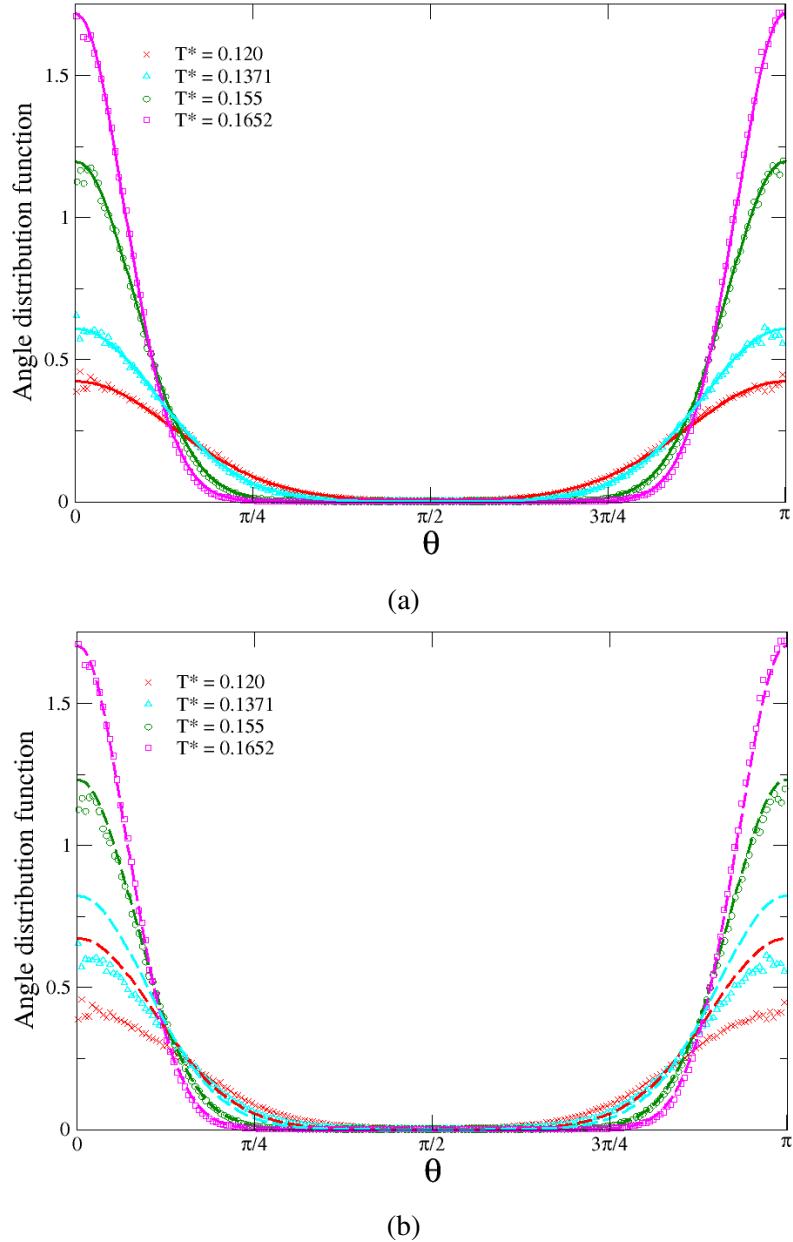


Figure 4.30: Angle distribution in simulation systems (symbols) compared to Onsager trial function (lines). (a) Solid lines are fitting results of simulation data to the function; (b) the parameter  $\alpha$  of the function is determined from the value of nematic order of the systems, making use of relation (2.16). Dashed lines represent the according function. In both figures, the lines are plotted in the same color of their corresponding simulation data.

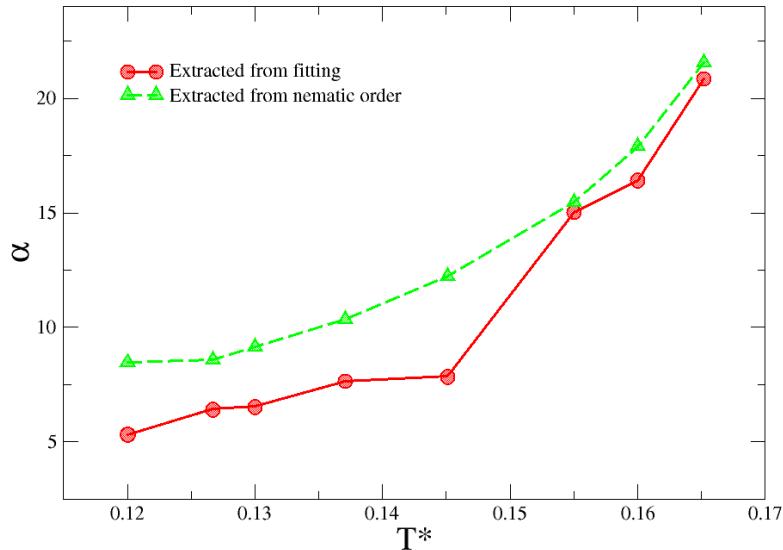


Figure 4.31: Orientational distribution in nematic systems. Symbols are simulation data, while lines are fit to Onsager's distribution, which is assumed in the theory. Inset: the values of  $\alpha$ , which are obtained by the fitting, for systems below and beyond the nematic re-entrant number density.

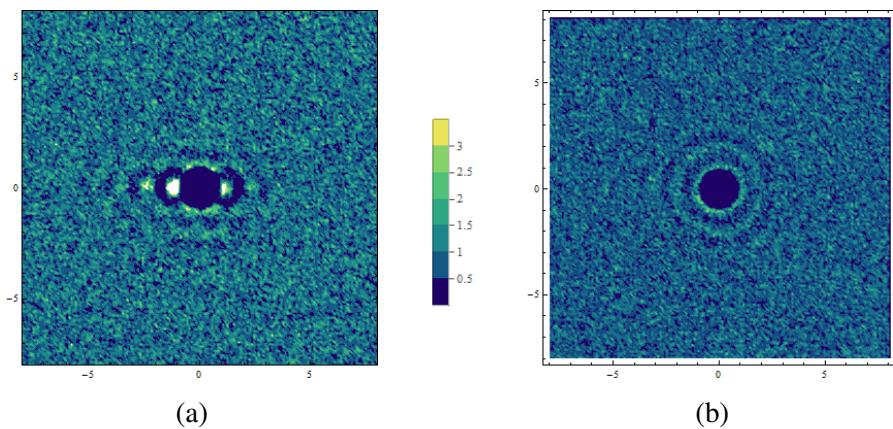


Figure 4.32: Pair distribution function of a configuration at  $T^* = 0.1652$  (a) in a plane parallel to the nematic director and (b) in a plane normal to the nematic director.

# Conclusion

We have presented in this thesis a procedure to accurately locate isotropic-nematic phase boundary in self-assembling systems over a range of temperature, based on a combination of simulation techniques proposed by Ref. [46] (SUS), Ref. [51] (Kofke integration) and Ref. [59]. By investigating simultaneously many small windows instead of a large windows to compute the probability distribution of number of particles, not only SUS reduces the computation time but also improve precision as the distribution in each windows is rather flat. The tricks suggested by Vink helps reduce the finite size effect in SUS and improve its precision without greatly increasing the number of particles. On the other hand Kofke integration allows one to calculate coexistent lines from the first coexistent points obtained from SUS. For both sticky hard cylinders and hard spheres we performed extra SUS points (3 for hard cylinders and 2 for spheres) as checkpoints for Kofke integration and found out that the results the 2 methods produced are in excellent agreement with each other. This suggests that the techniques are robust as a tool to study phase transition in self-assembling systems numerically. This is important in building a theory framework, as phase transition in self-assembling systems are complicated and theories need to start from simple models whose parameters can be estimated accurately. These parameters are often not available in experiments.

We also compared numerical results with the theory developed by De Michele et. al. [25]. The theory attempts to take into account many different aspects of self-assembling systems that are sometimes omitted by other theories: flexibility, polydispersity and modelling differently the entropy loss in nematic phase of short and long chains. The theory is also parameter free, in the sense that all the required parameters can be computed by simulations. For the systems of sticky hard cylinders and bent cylinders, the theory quantitatively agrees with simulation data when it comes to phase coexistent packing fraction. For spheres, the agreements deteriorate as the temperature, and hence coexistent packing fraction, increases.

Remarkably, the simulation results confirm a prediction made by the theory about the re-entrance behaviour of self-assembling systems, namely as the nematic coexistent packing fraction increases, the average aspect ratio of the chains exhibits a minimum rather than monotonically decreases as in non self-assembling systems. The re-entrance behaviour is the results of coupling between nematic order and self-assembly. The confirmation of the theory's prediction by simulation results indicates the promise of the theory as a framework to study isotropic-nematic phase transition in self-assembling systems.

In addition, we applied the theory to study the dependence of isotropic-nematic phase transition of sticky bent cylinders on their bending angles. Recent experiments have shown that the nematic critical concentration of DNA dodecamers is

---

sequence dependent. We showed that such phenomenon can be explained if the DNA dodecamers are modelled as symmetric bent cylinders. Using theoretical and numerical approach, we showed that bent cylinders with higher bending level, i.e. smaller bending angle, self-assemble into chains with smaller persistence length, which results in higher nematic packing fraction at coexistence. The theory therefore can be considered the first to study the isotropic-nematic phase transition in bent core systems.

Finally, we studied in details the structure of sticky hard sphere systems at coexistence. We pointed out that for systems of high flexibility, namely when the persistence length of isolated chain is of order 20 monomers, the chain length distribution even in nematic phase can be well described by a single exponential function. This is different from self-assembling systems of rigid chains studied by Lü and Kindt, where the double exponentiality is more pronounced and have to be taken into account. We also verified that the trial distribution proposed by Onsager works quite well at high packing fraction in system of hard spheres. This result is not trivial considering that Onsager proposed it for non self-assembling systems of highly anisotropic particles. In addition, we computed explicitly the parameter of the orientational distribution function proposed by Onsager [21] and adopted in the theory of De Michele et. al. [25] and confirmed that the re-entrance of the average aspect ratio is associated with a steep increase in such parameter.

Although this study focuses on isotropic-nematic transition, the same simulation procedure can be exploited to study other phase transitions. In fact, a possible extension of this study can be other liquid crystal phases in self-assembling systems, such as smectic and columnar phases. Another extension include improving the theory proposed by De Michel et. al. so that it works better with spheres and yields more accurate predictions for average chain length at high packing fraction.

# Bibliography

- [1] B. Alberts, A. Johnson, J. L. J, and *et al.*, *Molecular Biology of the Cell*. New York: Garland Science, 4th ed., 2002. Available from: <http://www.ncbi.nlm.nih.gov/books/NBK26871/>.
- [2] G. D. Rose, P. J. Fleming, J. R. Banavar, and A. Maritan *Proceedings of the National Academy of Sciences*, vol. 103, no. 45, pp. 16623–16633, 2006.
- [3] C. B. Anfinsen *Science*, vol. 181, no. 4096, pp. 223–230, 1973.
- [4] H. Fraenkel-Conrat and R. C. Williams *Biochemistry*, vol. 41, 1955.
- [5] G. M. Whitesides and M. Boncheva *Proceedings of the National Academy of Sciences*, vol. 99, no. 8, pp. 4769–4774, 2002.
- [6] Q. Chen, S. C. Bae, and S. Granick, “Directed self-assembly of a colloidal kagome lattice,” *Nature*, vol. 469, pp. 381–384, 2011.
- [7] S. C. Glotzer and M. J. Solomon, “Anisotropy of building blocks and their assembly into complex structures,” *Nat Mater*, vol. 6, no. 12, pp. 557–562, 2007.
- [8] F. Sciortino, “Primitive models of patchy colloidal particles. a review,” *Collection of Czechoslovak Chemical Communications*, vol. 75, pp. 349–358, Unspecified 2010.
- [9] F. C. Bawden, N. W. Pirie, J. D. Bernal, and I. Fankuchen *Nature*, vol. 138, pp. 1051–1052, 1936.
- [10] S. Fraden, *Phase Transitions in Colloidal Suspensions of Virus Particles*, vol. 460 of *NATO ASI Series*. Springer Netherlands, 1995.
- [11] Z. Dogic and S. Fraden, “Ordered phases of filamentous viruses,” *Current Opinion in Colloid & Interface Science*, vol. 11, no. 1, pp. 47–55, 2006.
- [12] R. Mezzenga, J.-M. Jung, and J. Adamcik *Langmuir*, vol. 26, no. 13, pp. 10401–10405, 2010.
- [13] J.-M. Jung and R. Mezzenga *Langmuir*, vol. 26, no. 1, pp. 504–514, 2010.
- [14] A. Aggeli, M. Bell, L. M. Carrick, C. W. G. Fishwick, R. Harding, P. J. Mawer, S. E. Radford, A. E. Strong, and N. Boden *Journal of the American Chemical Society*, vol. 125, no. 32, pp. 9619–9628, 2003.

## Bibliography

---

- [15] D. M. Kuntz and L. M. Walker, “Nematic phases observed in amphiphilic polyelectrolyte-surfactant aggregate solutions,” *Soft Matter*, vol. 4, pp. 286–293, 2008.
- [16] F. Livolant, A. M. Levelut, J. Doucet, and J. Benoit, “The highly concentrated liquid-crystalline phase of dna is columnar hexagonal,” *Nature*, vol. 339, pp. 724–726, 1989.
- [17] F. Tombolato and A. Ferrarini, “From the double-stranded helix to the chiral nematic phase of b-dna: A molecular model,” *The Journal of Chemical Physics*, vol. 122, no. 5, 2005.
- [18] M. Nakata, G. Zanchetta, B. D. Chapman, C. D. Jones, J. O. Cross, R. Pindak, T. Bellini, and N. A. Clark *Science*, vol. 318, no. 5854, pp. 1276–1279, 2007.
- [19] X. Lü and J. T. Kindt, “Monte carlo simulation of the self-assembly and phase behavior of semiflexible equilibrium polymers,” *J. Chem. Phys.*, vol. 120, pp. 10328–10338, 2004.
- [20] T. Kuriabova, M. D. Betterton, and M. A. Glaser *J. Mater. Chem.*, vol. 20, pp. 10366–10383, 2010.
- [21] L. Onsager *Ann.NY Acad.Sci*, vol. 51, p. 627, 1949.
- [22] F. Sciortino, E. Bianchi, J. F. Douglas, and P. Tartaglia *J. Chem. Phys.*, vol. 126, p. 194903, 2007.
- [23] T. Odijk *Macromolecules*, vol. 19, no. 9, 1986.
- [24] P. V. D. Shoot and M. E. Cates *Europhys. Lett.*, vol. 25, p. 515, 1995.
- [25] C. De Michele, T. Bellini, and F. Sciortino, “Self-assembly of bifunctional patchy particles with anisotropic shape into polymers chains: Theory, simulations, and experiments,” *Macromolecules*, vol. 45, no. 2, pp. 1090–1106, 2012.
- [26] B. R. Acharya, A. Primak, and S. Kumar, “Biaxial nematic phase in bent-core thermotropic mesogens,” *Phys. Rev. Lett.*, vol. 92, p. 145506, Apr 2004.
- [27] L. A. Madsen, T. J. Dingemans, M. Nakata, and E. T. Samulski, “Thermotropic biaxial nematic liquid crystals,” *Phys. Rev. Lett.*, vol. 92, p. 145505, Apr 2004.
- [28] M. J. Freiser, “Ordered states of a nematic liquid,” *Phys. Rev. Lett.*, vol. 24, pp. 1041–1043, May 1970.
- [29] G. R. Luckhurst, “Liquid crystals A missing phase found at last?,” *Nature*, vol. 430, pp. 413–414, 2004.
- [30] P. G. de Gennes and J. Prost, eds., *The Physics of Liquid Crystals*. Oxford: Clarendon Press, second edition ed., 1993.

- [31] “Phase transitions in liquid crystals.” [http://soft-matter.seas.harvard.edu/index.php/Phase\\_transitions\\_in\\_liquid\\_crystals](http://soft-matter.seas.harvard.edu/index.php/Phase_transitions_in_liquid_crystals).
- [32] T. Bellini, R. Cerbino, and G. Zanchetta, *DNA-Based Soft Phases*, vol. 318 of *Topics in Current Chemistry*. Springer Berlin Heidelberg, 2012.
- [33] P. Bolhuis and D. Frenkel, “Tracing the phase boundaries of hard spherocylinders,” *The Journal of Chemical Physics*, vol. 106, no. 2, 1997.
- [34] K. Merchant and R. Rill, “{DNA} length and concentration dependencies of anisotropic phase transitions of {DNA} solutions,” *Biophysical Journal*, vol. 73, no. 6, pp. 3154 – 3163, 1997.
- [35] J.-L. Barrat and J.-P. Hansen, *Basic Concepts for Simple and Complex Liquids*. Cambridge University Press, 2003.
- [36] G. J. Vroege and H. N. W. Lekkerkerker, “Phase transitions in lyotropic colloidal and polymer liquid crystals,” *Reports on Progress in Physics*, vol. 55, no. 8, p. 1241, 1992.
- [37] H. N. W. Lekkerkerker, P. Coulon, R. Van Der Haegen, and R. Deblieck, “On the isotropic liquid crystal phase separation in a solution of rodlike particles of different lengths,” *The Journal of Chemical Physics*, vol. 80, no. 7, 1984.
- [38] D. Frenkel, “Onsager’s spherocylinders revisited,” *The Journal of Physical Chemistry*, vol. 91, no. 19, pp. 4912–4916, 1987.
- [39] D. Frenkel, “Onsager spherocylinders revisited [erratum to document cited in ca107(14):124951h],” *The Journal of Physical Chemistry*, vol. 92, no. 18, pp. 5314–5314, 1988.
- [40] W. G. Chapman, G. Jackson, and K. E. Phase *Mol. Phys.*, vol. 65, p. 1, 1988.
- [41] J. Russo, J. M. Tavares, P. I. C. Teixeira, M. M. T. da Gama, and F. Sciortino, “Re-entrant phase behaviour of network fluids: A patchy particle model with temperature-dependent valence,” *The Journal of Chemical Physics*, vol. 135, no. 3, 2011.
- [42] D. C. Williamson and G. Jackson *J. Chem. Phys.*, vol. 108, p. 10294, 1998.
- [43] D. Frenkel and B. Smit, eds., *Understanding Molecular Simulation*. San Diego: Academic Press, second edition ed., 2002.
- [44] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*. New York, NY, USA: Clarendon Press, 1989.
- [45] H. Gould, J. Tobochnik, and C. Wolfgang, *An Introduction to Computer Simulation Methods: Applications to Physical Systems*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., third edition ed., 2005.

## Bibliography

---

- [46] P. Virnau and M. Müller, “Calculation of free energy through successive umbrella sampling,” *The Journal of Chemical Physics*, vol. 120, no. 23, pp. 10925–10930, 2004.
- [47] K. Binder and D. P. Landau, “Finite-size scaling at first-order phase transitions,” *Phys. Rev. B*, vol. 30, pp. 1477–1485, Aug 1984.
- [48] A. M. Ferrenberg and R. H. Swendsen, “New monte carlo technique for studying phase transitions,” *Phys. Rev. Lett.*, vol. 61, pp. 2635–2638, Dec 1988.
- [49] K. Binder, B. J. Block, P. Virnau, and A. Trster, “Beyond the van der waals loop: What can be learned from simulating lennard-jones fluids inside the region of phase coexistence,” *American Journal of Physics*, vol. 80, no. 12, 2012.
- [50] L. G. MacDowell, P. Virnau, M. Müller, and K. Binder, “The evaporation/condensation transition of liquid droplets,” *The Journal of Chemical Physics*, vol. 120, no. 11, pp. 5293–5308, 2004.
- [51] D. A. Kofke and E. D. Glandt, “Nearly monodisperse fluids. i. monte carlo simulations of lennardjones particles in a semigrand ensemble,” *The Journal of Chemical Physics*, vol. 87, no. 8, pp. 4881–4890, 1987.
- [52] F. A. Escobedo and J. J. de Pablo, “Pseudo-ensemble simulations and gibbs-duhem integrations for polymers,” *The Journal of Chemical Physics*, vol. 106, no. 7, pp. 2911–2923, 1997.
- [53] D. Kofke, *Semigrand canonical Monte Carlo simulation: Integration along coexistence lines*, vol. 108. John Wiley & Sons, Inc., 2007.
- [54] M. Dijkstra and D. Frenkel, “Simulation study of the isotropic-to-nematic transitions of semiflexible polymers,” *Phys. Rev. E*, vol. 51, pp. 5891–5898, Jun 1995.
- [55] P. G. Bolhuis and D. A. Kofke, “Monte carlo study of freezing of polydisperse hard spheres,” *Phys. Rev. E*, vol. 54, pp. 634–643, Jul 1996.
- [56] P. J. Camp, C. P. Mason, M. P. Allen, A. A. Khare, and D. A. Kofke, “The isotropicnematic phase transition in uniaxial hard ellipsoid fluids: Coexistence data and the approach to the onsager limit,” *The Journal of Chemical Physics*, vol. 105, no. 7, 1996.
- [57] R. Agrawal and D. A. Kofke, “Solid-fluid coexistence for inverse-power potentials,” *Phys. Rev. Lett.*, vol. 74, pp. 122–125, Jan 1995.
- [58] J. L. F. Abascal, E. Sanz, R. Garca Fernndez, and C. Vega, “A potential model for the study of ices and amorphous water: Tip4p/ice,” *The Journal of Chemical Physics*, vol. 122, no. 23, pp. –, 2005.
- [59] R. L. C. Vink and T. Schilling, “Interfacial tension of the isotropic-nematic interface in suspensions of soft spherocylinders,” *Phys. Rev. E*, vol. 71, p. 051716, May 2005.

- [60] B. Lubachevsky and F. Stillinger, “Geometric properties of random disk packings,” *Journal of Statistical Physics*, vol. 60, no. 5-6, pp. 561–583, 1990.
- [61] A. R. Kansal, S. Torquato, and F. H. Stillinger, “Computer generation of dense polydisperse sphere packings,” *The Journal of Chemical Physics*, vol. 117, no. 18, 2002.
- [62] C. De Michele, L. Rovigatti, T. Bellini, and F. Sciortino, “Self-assembly of short dna duplexes: from a coarse-grained model to experiments through a theoretical link,” *Soft Matter*, vol. 8, pp. 8388–8398, 2012.
- [63] R. Blaak, D. Frenkel, and B. M. Mulder, “Do cylinders exhibit a cubatic phase?,” *The Journal of Chemical Physics*, vol. 110, no. 23, 1999.
- [64] N. Ibarra-Avalos, A. Gil-Villegas, and A. Martinez Richa, “Excluded volume of hard cylinders of variable aspect ratio,” *Molecular Simulation*, vol. 33, no. 6, pp. 505–515, 2007.
- [65] T. Niori, T. Sekine, J. Watanabe, T. Furukawa, and H. Takezoe, “Distinct ferroelectric smectic liquid crystals consisting of banana shaped achiral molecules,” *J. Mater. Chem.*, vol. 6, pp. 1231–1233, 1996.
- [66] G. Zanchetta, F. Giavazzi, M. Nakata, M. Buscaglia, R. Cerbino, N. A. Clark, and T. Bellini, “Right-handed double-helix ultrashort DNA yields chiral nematic phases with both right- and left-handed director twist,” vol. 107, no. 41, pp. 17497–17502, 2010.
- [67] N. Kern and D. Frenkel, “Fluidfluid coexistence in colloidal systems with short-ranged strongly directional attraction,” *The Journal of Chemical Physics*, vol. 118, no. 21, 2003.
- [68] D. C. Williamson and G. Jackson, “Liquid crystalline phase behavior in systems of hard-sphere chains,” *The Journal of Chemical Physics*, vol. 108, no. 24, 1998.

# List of publications

1. T. K. Nguyen, F. Sciortino and C. De Michele, “*Self-assembly-driven nematicization*”, Langmuir, 2014, 30 (16), pp 4814–4819
2. T. K. Nguyen, A. Battisti, D. Ancora, F. Sciortino and C. De Michele, “*Self-assembly of mesogenic bent-core DNA nanoduplexes*”, Soft Matter, 2014, DOI: 10.1039/C4SM01571A