Entropy controlled phase transitions

Sampedro Ruiz, Pablo

2020

Sampedro Ruiz, P. (2020). Entropy controlled phase transitions, Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/145375>

<https://doi.org/10.32657/10356/145375>

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

Downloaded on 26 Apr 2023 19:09:17 SGT

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

Entropy controlled phase transitions

Pablo Sampedro Ruiz

School of Chemical and Biomedical Engineering

2020

Entropy controlled phase transitions

Pablo Sampedro Ruiz

School of Chemical and Biomedical Engineering

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

2020

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research, is free of plagiarised materials, and has not been submitted for a higher degree to any other University or Institution.

July, 15, 2020

.....

Date

.....

Pablo Sampedro Ruiz

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it is free of plagiarism and of sufficient grammatical clarity to be examined. To the best of my knowledge, the research and writing are those of the candidate except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

July, 15, 2020

.....

Date

.....

Prof. Ni, Ran

Authorship Attribution Statement

This thesis contains material from 2 papers published in the following peer-reviewed journals in which I am listed as an author.

Chapter 2 is published as [Hao Hu, Pablo Sampedro Ruiz, and Ran Ni Phys. Rev. Lett. 120, 048003, \(2018\) DOI: <https://doi.org/10.1103/PhysRevLett.120.048003>](#)

The contributions of the co-authors are as follows:

- Dr. Hao Hu wrote the drafts of the manuscript. The manuscript was revised together with Prof. Ran Ni and me.
- Dr. Hao Hu programmed the code to implement the original computational method by Angioletti and Uberti[1] and its adaptation to calculate mDNACCs binding energy in the zero temperature limit. I programmed the code to simulate the 2D *sticks and balls* model used to compare with its analogous 2D mDNACCs at the zero temperature limit.
- I performed simulations to obtain the equation of state for 2D mDNACCs in the zero temperature limit and those to obtain the equation of state of the sticks and balls model. Hu Hao performed all other calculations.

Chapter 3 is published as [Sampedro Ruiz, P., Lei, Q. Ni, R. Melting and re-entrant melting of polydisperse hard disks. Commun Phys 2, 70 \(2019\). DOI:<https://doi.org/10.1038/s42005-019-0172-2>](#)

The contributions of the co-authors are as follows:

- Prof Ni suggested the materials area and edited the manuscript drafts.
- I wrote the drafts of the manuscript. The manuscript was revised together with Dr. Lei Qunli
- I developed all codes and launched all simulations, except for the event-driven molecular dynamics simulations required to calculate the diffusivity of re-entrant melting phases, which were performed by Dr. Lei Qunli.
- I performed all data analysis to build the final phase diagrams.

.... July, 15, 2020

Date

.....

Pablo Sampedro Ruiz

Acknowledgements

I wish to express my greatest gratitude to my family, all the friends that have helped me during my stay in Singapore and my advisor.

“Reality is that which, when you stop believing in it, doesn’t go away.”

—Phillip K. Dick

Abstract

A fundamental part of the design of new materials lies in the study and control of the phase transitions of solids. This control implies tweaking the free energies of the different phases of the material, in order to stabilize those we are interested in or destabilize those we are not. Although classically it has been understood that the formation of new solids is driven by enthalpic factors, entropy supposes a new and promising tool in the field. This kind of approach offers us a simpler design of assembly particles and a more reliable process. In this thesis we study the effect entropy can have in the self-assembly of colloids, by applying this approach to two experimental colloidal systems (mDNACCs and 2D hard spheres), as well as presenting a new computational method designed for the study of the assembly of spherical colloids.

Contents

Acknowledgements	ix
Abstract	xiii
List of Figures	xvii
1 Introduction	1
1.1 Preface	1
1.1.1 Thesis outline	3
1.2 Two dimensional solids	3
1.2.1 KTHNY theory	4
1.2.2 Defect energies for a 2D triangular lattice	7
1.2.3 Correlation functions	13
1.2.4 Phase diagram of the hard disk system	14
1.3 Methods	17
1.3.1 Free energy calculation in solids	17
1.3.2 Event chain Monte Carlo	20
2 Entropy stabilization of floppy crystals of mobile DNA-coated colloids	23
2.1 Introduction	23
2.2 Finite binding energy mDNACC	24
2.3 Infinite binding energy system	28
2.4 Conclusions	35
3 Melting and re-entrant melting of polydisperse hard-disks	37
3.1 Introduction	37
3.2 Model	38
3.3 Results	42
3.3.1 Phase diagram	42
3.3.2 Finite size effect on the melting in polydisperse hard-disks	44
3.3.3 Stabilization of the hexatic phase	44
3.3.4 Re-entrant melting	49

3.3.5	Non monotonic packing fraction behaviour for the ideal gas phase	53
3.4	Conclusions	57
4	Effect of particle size distribution in polydisperse hard-disks	59
4.1	Introduction	59
4.2	Lognormal polydisperse hard-disks	60
4.3	Triangular polydisperse hard-disks	64
4.4	Conclusions	67
5	Adaptive bias Boltzmann generators	69
5.1	Boltzmann generators	70
5.2	Method	73
5.3	Results	75
5.4	Conclusions	79
6	Conclusions	83
6.1	Summary	83
6.2	Final remarks	86
6.2.1	Polydisperse hard-disks	86
6.2.2	Adaptive bias Boltzmann generators	87
A	Additional figure for Chapter 3	89

List of Figures

1.1	Dislocation in a triangular lattice with an added crystal semi-row. The seven neighbour particle is represented in red while the five neighbour in green, the black arrows show the path around the dislocation and the red arrow shows the Burgers vector associated to the dislocation.	6
1.2	Displacement field induced by a single dislocation on the 2D triangular lattice centred at the origin and Burguess vector $(a_0, 0)$	10
1.3	Displacement field induced by a two dislocations on the 2D triangular lattice centred at the origin, Burguess vectors $\vec{b}_1 = -\vec{b}_2 = (a_0, 0)$ and $\vec{r}_{12} = (a_0, 0)$	10
2.1	Phase diagrams for mobile DNA-coated colloids in two and three dimensional space, in the area/packing fraction η - binding strength $\beta\Delta G_0$ plane. Black circles mark the position of the triple points obtained by extrapolation. The dashed line is an estimate of the transition from the hexagonal phase to the disordered hexagonal phase. Snapshots are shown for the ordered crystal phases. Gray areas represent the coexistence regions.	26
2.2	Interaction potential $\beta U/N$ and vibrational entropy S_{vib}/kN versus the binding strength $\beta\Delta G_0$, at the square - hexagonal phase coexistence in 2D, and CsCl - CuAu phase coexistence in 3D. . . .	29
2.3	Simulation results for mobile DNA-coated colloids at the infinite binding strength limit $\beta\Delta G_0 \rightarrow -\infty$. (a,b) Probability density distribution of area/packing fraction η of the system versus pressure P in 2D (a) and 3D (b). (c) Interaction potential $\beta U/N$ and enthalpy $\beta H/N$ versus the pressure P . Solid curves are obtained by data fitting.	30
2.4	Probability distribution function for packing fraction for the 2D infinite binding mDNACCs system, showing the average of the distribution (red line) and the center of two fit Gaussian distributions that correspond to the square (green line) and the hexagonal phase (blue line). A clear coexistence region can be distinguished for pressures between $1.0kT/\sigma^2$ and $1.5kT/\sigma^2$	33

2.5	(a) Two consecutive “floppy” moves of the sticky-sphere model from an initial square lattice. In simulations only particles in a single column or row are allowed to move relatively to other particles at a time. Here we show a column move in the middle (top) followed by a row move in the third row (bottom). (b) Probability density distribution of area fraction η of the system versus pressure P for the sticky-sphere model in 2D, for which distances between neighboring A and B particles are fixed at $d_{AB} = 1.04\sigma$	35
3.1	Phase behaviour of polydisperse hard-disks. (a): Equation of state (EOS) for polydisperse hard-disk systems (PHDS) with various polydispersity parameter $\nu/\sigma_0 = 0.005$ to 0.08 in the representation of $(P - P^*)\sigma_0^2/k_B T$ vs $(\rho^{-1} - \rho_{hex}^{-1})\sigma_0^{-2}$, where P^* and ρ_{hex} are the pressure and density of the hexatic phase at the fluid-hexatic transition, respectively, and the solid lines are fits of the EOS using 5th order polynomials. (b): $ \langle\Psi_6\rangle $ as functions of $(\rho^{-1} - \rho_{hex}^{-1})\sigma_0^{-2}$ for systems with $\nu/\sigma_0 = 0.07$ and 0.08 . P and $\rho = N/V$ are the pressure of the system and the density of particles in the system, respectively. (c): EOSs for PHDS of various numbers of particles $N = 64^2 \sim 512^2$ at $\nu/\sigma_0 = 0.07$. The solid symbols are the obtained fluid-hexatic transition points.	39
3.2	Phase behaviour of polydisperse hard-disks. (a): Phase diagram of the PHDS in the representation of $\rho\sigma_0^2$ and $s/\langle\sigma\rangle$, in which the dashed lines are the interpolated phase boundaries for re-entrant melting of solid and hexatic phases. The phase boundaries are obtained from $NVT - \Delta\mu$ MC simulations for PHDS with $\nu/\sigma_0 = 0.005$ to 0.0835 . Inset: the enlarged view of the region of phase diagram at $0 \leq s/\langle\sigma\rangle \leq 0.02$. The dotted lines connect the re-entrant transitions at $\nu/\sigma_0 = 0.08, 0.0805, 0.081, 0.082, 0.083$, and 0.0835 from left to right. and (b): same phase diagram but including the equations of state along the re-entrant transitions for $\nu/\sigma_0 = 0.08, 0.0805, 0.081, 0.082, 0.083$, and 0.0835 . Figure b) is mostly referenced in Chapter 4 of this work.	40
3.3	Stabilization of hexatic phase. (a): The fractions of topological defects and dislocations in the hexatic phase as functions of $(\rho - \rho_{hex})\sigma_0^2$ for systems of various polydispersity parameter ν . Here ρ and ρ_{hex} are the density of the system and the lowest density of stable hexatic phase, respectively. The vertical dashed lines indicate the hexatic-solid transition points. (b): Low pressure phase diagram of polydisperse hard-disks in the representation of η vs $s/\langle\sigma\rangle$, where η is the packing fraction of the system with σ_i the diameter of particle i . The state points obtained from simulations at each ν/σ_0 from 0.005 to 0.08 are shown as the symbols, which are color coded with $\langle\sigma\rangle/\sigma_0$. The error bars are smaller than the symbols.	41

3.4	Re-entrant melting transitions. (a,b): The pair correlation function $g(x, 0) - 1$ along the major axis of the system (a) and the six fold orientation correlation function $g_6(r)$ (b) for polydisperse hard-disk systems (PHDS) with $\nu/\sigma_0 = 0.08$ at various densities. (c): $ \langle\Psi_6\rangle $ and $s/\langle\sigma\rangle$ as functions of density $\rho\sigma_0^2$ for PHDS with $\nu/\sigma_0 = 0.082, 0.083$, and 0.0835 . (d): Contour plot of the probability density distribution of particle size σ/σ_0 for systems of different densities at $\nu/\sigma_0 = 0.0835$. (e): Mean square displacement $\langle\Delta r^2\rangle/\sigma_0^2$ in PHDS of $\nu/\sigma_0 = 0.0835$ at densities $\rho\sigma_0^2 = 1.592, 2.292$, and 3.565 as marked in the inset, where $\tau = \sigma_0\sqrt{m/k_BT}$ is the time unit of molecular dynamics simulations with m the mass of particles. Inset: the diffusion coefficient D as a function of density in PHDS with $\nu/\sigma_0 = 0.0835$.	43
3.5	The equation of state for polydisperse hard-disk systems (PHDS) of various system sizes with $\nu/\sigma_0 = 0.005, 0.01, 0.02$ and 0.05 . The solid lines are fits of the EOS using 5th order polynomials, and solid symbols are the obtained coexisting densities from the Maxwell construction.	45
3.6	The fraction of particles in topological defects clusters with nonzero disclination charges in the hexatic phase as functions of $(\rho - \rho_{hex})\sigma_0^2$ for systems of various ν . The representation as a function of ρ rather than η was chosen to simplify the figure. This is however not important and the main purpose of Fig. The vertical dashed lines indicate the hexatic-crystal transition points.	48
3.7	Probability density distribution of particle size σ in polydisperse hard-disk systems (PHDS) with various ν and ρ , in which the solid lines are fitted Gaussian distribution.	48
3.8	Six-fold bond orientation order parameter of the system $\langle\Psi_6\rangle$ as functions of Monte Carlo (MC) sweeps in simulations from perfect crystal and random fluid at $\rho\sigma_0^2 = 1.528, 2.292$ and 3.565 with $\nu/\sigma_0 = 0.0835$. From these, one can see that the correlation time is around 10^7 MC sweeps, which is smaller than our simulation of around 10^8 MC sweeps, and the our simulations have been well equilibrated.	50
3.9	Equation of state for polydisperse hard-disk system with $\nu/\sigma_0 = 0.082, 0.083$ and 0.0835 .	50
3.10	$s/\langle\sigma\rangle$ (a) as well as η and $D\tau/\sigma_0^2$ (b) as functions of density $\rho\sigma_0^2$ for a system with $\nu/\sigma_0 = 0.0835$.	52
3.11	Packing fraction η as a function of density $\rho\sigma_0^2$ for Gaussian-like polydisperse hard-disks with $\nu/\sigma_0 = 0.1$.	52
3.12	a) Average particle size ($\langle\sigma\rangle/\sigma_0$) and b) absolute polydispersity (s/σ_0) VS density ($\rho\sigma_0^2$) for a system of Gaussian PDHD with $\nu/\sigma_0 = 0.1$ (black) and for an ideal gas configuration of the Gaussian polydisperse hard-disk system under equal pressure (red).	54

3.13	a) Packing fraction (η) VS density ($\rho\sigma_0^2$) for a system of Gaussian polydisperse hard-disks with $\nu/\sigma_0 = 0.1$ (black) and for an ideal gas configuration of the Gaussian polydisperse hard-disk system under equal pressure (red).	56
4.1	(a) EOS for lognormal polydisperse hard-disk systems with various polydispersity parameter $\nu/\sigma_0 = 0.05$ to 0.15 in the representation of $(P - P^*)\sigma_0^2/k_B T$ vs $(\rho^{-1} - \rho_{hex}^{-1})\sigma_0^{-2}$, where P^* and ρ_{hex} are the pressure and density of the hexatic phase at the fluid-hexatic transition, respectively, and the solid lines are fits of the EOS using 5th order polynomials. (b) $ \langle\Psi_6\rangle $ as functions of $(\rho^{-1} - \rho_{hex}^{-1})\sigma_0^{-2}$ for systems with $\nu/\sigma_0 = 0.10, 0.13$ and 0.15	61
4.2	(a) Phase diagram of the polydisperse hard-disk systems of lognormal distribution in the representation of $\rho\sigma_0^2$ and $s/\langle\sigma\rangle$. The phase boundaries are obtained from $NVT - \Delta\mu$ simulations for systems with $\nu/\sigma_0 = 0.005$ to 0.2 . Inset: the enlarged view of the region of phase diagram at $0.06 \leq s/\langle\sigma\rangle \leq 0.08$. The crosses are the simulation results along the the re-entrant transitions at $\nu/\sigma_0 = 0.13, 0.15, 0.17$, and 0.2 from bottom to top. (b) Low pressure phase diagram of polydisperse hard-disks with lognormal distribution in the representation of η vs $s/\langle\sigma\rangle$, where η is the packing fraction of the system. The state points obtained from simulations at each ν/σ_0 from 0.08 to 0.20 are shown as the symbols. The error bars are smaller than the symbols	62
4.3	Chemical potential difference $\Delta\mu(\sigma)/k_B T$ for Gaussian-like (Eq. 4.1) and lognormal polydisperse (Eq. 4.2) hard-disks with $\nu/\sigma_0 = 0.1$	63
4.4	EOS for triangle polydisperse hard-disk systems with various polydispersity parameter $\nu/\sigma_0 = 0.25$ to 0.3 in the representation of $(P - P^*)\sigma_0^2/k_B T$ vs $(\rho^{-1} - \rho_{hex}^{-1})\sigma_0^{-2}$, where P^* and ρ_{hex} are the pressure and density of the hexatic phase at the fluid-hexatic transition, respectively, and the solid lines are fits of the EOS using 5th order polynomials.	64
4.5	Phase diagram for the systems of triangular polydisperse hard-disks in the representations of $\rho\sigma_0^2$ vs $s/\langle\sigma\rangle$ (a) and η vs $s/\langle\sigma\rangle$ for ν/σ_0 from 0.02 to 0.3 . The inset in (a) is the enlarged view of the low polydispersity region.	65
4.6	Polydispersity $s/\langle\sigma\rangle$ as a function of packing fraction η for highly polydisperse hard-disks with triangle distributions at $0.08 \leq \nu/\sigma \leq 0.3$	65
5.1	Potential Energy Surface for the 2D four well system used to test the adaptive bias Boltzmann generators algorithm	76
5.2	Energy sampling of the four-well model using a single Boltzmann Generator and energy training. Non colored regions are not sampled by the method.	77

5.3	Successive biased potentials $u_k(\mathbf{x})$ used to sample $p_{k+1}(\mathbf{x})$. It can be seen that as the algorithm fills the different energy wells the potential becomes flat	78
5.4	Successive repulsive potentials $\hat{u}_k(\mathbf{x})$ obtained	79
5.5	Results for the final ML+KL optimization of a Boltzmann generator using the information collected during the sampling process: a) Non normalized ($\log(p_X(\mathbf{x}))$) free energy surface, b) Normalized ($u(\mathbf{x}) - \log(p_X(\mathbf{x}))$) free energy surface, c) and d) a collection of equilibrium configurations in Cartesian space taken from MC simulations and their respective representation in the latent space	80
A.1	Packing fraction η as functions of density $\rho\sigma_0^2$ for PHDS with $\nu/\sigma_0 = 0.082$ (black), 0.083 (red), and 0.0835 (green). Errors are smaller than the symbols.	89

Chapter 1

Introduction

1.1 Preface

It has been said that technology is currently determined by the state of Material Science. Material properties such as the strength of steel or the highest working temperatures of superconductors are hard limits that no amount of engineering can surpass. Therefore, the need of our society for improved materials and techniques to synthesize them never stops.

A fundamental part of the design of new materials lies in the study and control of the phase transitions of solid matter. This allows us to determine in which conditions a certain crystal will form and how we should modify the different control parameters (temperature, pressure, interactions between composing particles) in order to stabilize it. The subject of this thesis is precisely the study of phase transitions and to explore how we can control them.

As we know, at constant pressure and temperature the most stable conformation of a material will be that with the lowest free energy. In turn, this free energy has both an enthalpic and an entropic contribution. Classically, the most common approach to stabilize new crystals has been to use enthalpy, as indeed most of the crystallizations we observe in our day to day lives (the freezing of water, the precipitation of a non soluble salt, etc) are driven by enthalpy. Entropy, we believe, tends to favour more disordered forms of matter, such as liquids and gases.

However, this view was challenged when it was proven that hard spheres, an athermal system for which all possible configurations present identical internal energy, can crystallize. This very notion was controversial for a time, as it was argued that no true phase transition to the solid could occur in this system, as the entropy change to form an ordered solid from a disordered fluid would always be negative. However, this reasoning did not take into account that the solid uses space much more efficiently, which leads to a higher vibrational freedom for the solid phase when compared to the jammed fluid, at high enough packing fractions. Therefore, the system presents a first order transition to the solid [2] for a 0.545 packing fraction.

This surprising finding shows us that entropy can indeed play an important part in the formation of solids, and offers us a new instrument to control phase transitions to add to our tool-shed. As the title of this thesis implies, our focus was to study and develop new strategies and tools to control phase transitions using entropy. More specifically, this study was centred in the field of colloidal chemistry science and was done using computational techniques.

One field that can significantly benefit from this kind of entropic stabilization of crystals is the development of colloids that assemble to form open lattices. Open lattices, in opposition to compact ones, are crystal structures with packing fractions smaller than the optimal ones and that are generally characterized for presenting less constraints in their structure than the total amount of degrees of freedom, allowing therefore deformation under no energy cost. Several of them are of interest due to presenting unusual properties, such as negative thermal expansion [3] and negative Poisson's ratio[4–6], as well as applications for photonic crystals [7, 8].

When assembling open structures using colloids, several approaches have focused on stabilizing them enthalpically, adding additional potential energy terms between particles that stabilize the open lattice[9–13]. However, most of these designs are difficult to realize with current experimental techniques for colloidal particles.

On the other hand, some recent experimental implementations have managed to stabilize open lattices over compact ones due to their higher mobility and therefore rotational or vibrational entropy[14–16]. Unlike its enthalpic alternatives, entropic stabilization can occur with simple interactions and with higher flexibility, allowing crystals to deform as they grow and therefore reducing the likelihood of the kinetic

arrest of the crystal growth. Coincidentally, in this thesis we present results for a similar approach based on mobile DNA coated colloids, that allows the formation of the open CsCl lattice.

Therefore, we can see that entropy offers great potential for the design of new solid phases. This thesis aims to contribute a small brick to this field that is sure to grow in the future.

1.1.1 Thesis outline

This thesis is divided into a total of six chapters, with the core of the thesis being Chapter 2, 3, 4 and 5. For the remaining of this introduction, we present a short summary of some important concepts concerning 2D solids, since the most significant part of our work consisted in studying the phase diagram of polydisperse hard-disk. All these concepts will be used and applied in Chapter 3 and 4. In Chapter 2, we present our results for the entropic stabilization of 2D and 3D floppy lattices using mobile DNA coated colloids and it has been published in [17]. Chapter 3 and 4 shows our results for the effect of polydispersity in the equation of state of hard-disk and the effect that the polydispersity distribution shape can have on them and part of its content has been published in [18], with the rest to be published in the near future. In Chapter 5 we present preliminary results for adaptive bias Boltzmann generators, a new computational method to sample the potential energy surface of complicated multi-dimensional systems. Finally, Chapter 6 closes the thesis, summarizing our main results, and offers a perspective of future research paths.

1.2 Two dimensional solids

Particles arranged in a two dimensional space and displaying short range interactions can never display long range order. This was first proven to be true in the case of a magnetic 2D system using the following reasoning. If we approximate the spin-spin interaction between two neighbors using an harmonic expression and assume that each particle only interacts with a finite amount of neighbours, then the longest wavelength mode of the system implies an energy per particle of the order

of $(2\pi/L)^2$ (where L is the system length and we are considering that for the lowest wavelength mode a full spin rotation is shared between all spins in the system). Adding the energy of all spins leads to a total energy of the order of $L^d(2\pi/L)^2$, where d is the dimensionality of the system. It is easy to see that for one and two dimensional systems this energy does not diverge with the system size, and therefore, the energy cost per particle approaches zero as the system grows and the longest wavelength mode will be activated for any non zero temperature, destroying the long range order. However, this does not mean that all order is lost, since for a 2D system the deformation of the crystal correlations show a slow logarithmic increase with position, that is:

$$\left\langle \left[\vec{u}(\vec{R}) - \vec{u}(\vec{R}') \right]^2 \right\rangle \approx \log |\vec{R} - \vec{R}'| \quad (1.1)$$

where $\vec{u}(\vec{R})$ is the displacement of a particle and $\vec{R}$ and $\vec{R}'$ are two crystal sites. Therefore, although this system does not show long range order, the decay for the spatial correlations is slow enough to talk about a quasi-crystal and we can study the phase diagram of the system. This way, when analysing the structure factor for a two dimensional solid, one finds Gaussian peaks for the different crystallographic translations, rather than the discrete lines found for a real solid.

1.2.1 KTHNY theory

Contrary to 3D solids, microscopic theories of melting have been proposed to explain 2D melting. The most famous 2D melting mechanism is probably that proposed by the KTHNY theory (named after Kosterlitz, Thouless, Halperin, Nelson and Young); a topological theory of melting in which the transition is explained due to the formation of topological defects. This study of topological phase transitions was awarded with the 2016 Physics Nobel prize to both Thouless and Kosterlitz, due to the high expectancy on future applications in materials and electronics.

The main prediction of the theory is the existence of a temperature T_m at which solid phase (that presents quasi-long range positional order and long range orientational order), transits to the *hexatic* phase (which, unlike the fluid, still presents quasi-long range orientational order and is therefore a new phase specific of 2D solids). Furthermore, at a higher temperature T_i , all correlation is lost and a fluid

is formed[19, 20]. As we will see, each of these transitions is determined by the formation of an specific topological defect: free dislocations in the case of the solid-hexatic transition and disclinations in the case of the hexatic-fluid transition.

The key to the KTHNY theory lies in understanding what topological imperfections are and how they can appear through continuous transformations in a perfect crystal lattice. All further derivation will be referred for the case of an hexagonal lattice, since it is the densest possible for a 2D system and therefore the one observed most typically for 2D solids.

As we said, the KTHNY theory explains the melting due to the formation of free dislocations in the crystal lattice. Dislocations are topological defects formed by the insertion of half a crystal line in an otherwise perfect crystal lattice. Fig. 1.1 shows how one such dislocation looks like in a hexagonal lattice. From this figure we can also recognize the proper method to identify and characterize a dislocation. Starting from a certain lattice point, one draws a path that surrounds the dislocation, performing the same amount of steps along the two crystal directions, both forwards and backwards. While for a perfect lattice this gives a close loop, if a dislocation is contained by our path, the initial and final point will not be the same. Instead, the vector that joins the initial and final point in the path is called the Burgers vector ($\vec{b}$). This vector will always be a multiple of one of the three lattice vectors for the hexagonal lattice, due to the insertion of an integer number of parallel crystal half lines. Due to the high stress induced by a dislocation on the lattice, the KTHNY theory only considers the formation of dislocations with a Burgers vector with module equal to a single lattice unit.

From Fig. 1.1 we can also see that when performing the Voronoi construction to find the neighbours for every particle, the dislocation is marked by the presence of a five neighbour and seven neighbour particle pair, instead of the six neighbours per particle found in the perfect lattice. This way, the dislocation can be viewed as a defect pair joined by a vector $\vec{r}_{57}$, which is practically perpendicular to the Burgers vector.

There is still the question of how it is possible for half a crystal line to appear inside a perfect crystal through continuous transformations. In fact, this is not possible and the process by which dislocations appear in a hexagonal lattice is by the breaking of dislocation pairs. Dislocation pairs are bound dislocations that

FIGURE 1.1: Dislocation in a triangular lattice with an added crystal semi-row. The seven neighbour particle is represented in red while the five neighbour in green, the black arrows show the path around the dislocation and the red arrow shows the Burgers vector associated to the dislocation.

for the case of the hexagonal lattice take the form of a pair of five and another pair of seven neighbour particles arranged in a cross. They originate from the spontaneous fragmentation of a full crystal line into two halves, and can therefore appear through continuous transitions. They are present in hexagonal lattices at any non zero temperature and are compatible with quasi-long range order.

At high enough temperatures, dislocation pairs can unbind through a continuous transition to form two independent half crystal lines. As we mentioned before, the KTHNY theory predicts that the the melting temperature of the solid T_m will be that at which this unbinding occurs spontaneously. The consequent phase, however, is not a fluid, but rather still conserves quasi-long range orientational order. The system will then suffer another melting to the fluid that occurs due to the appearance of disclinations.

In the hexagonal lattice, disclinations appear as isolated particles with five or seven neighbours and their effect on the crystal lattice is a rotation of the crystal of $-\pi/3$ radians in the case of five neighbour particles and of $+\pi/3$ radians for seven neighbour particles. Therefore, just as it was the case for dislocations pairs and dislocations, the unbinding of a dislocation leads to the formation of two disclinations.

To calculate the T_m and T_i transition temperatures predicted by the KTHNY theory, we need to calculate the energy and entropy associated with the unbinding

of a dislocation pair and a dislocation, that will tell us at which temperatures this transitions are spontaneous.

1.2.2 Defect energies for a 2D triangular lattice

To calculate the energy associated to a certain topological defect we use a second order expression for the deformation of a 2D lattice. In this case, the elastic free energy F is expressed as:

$$F = \frac{\lambda}{2} u_{ii} u_{jj} + \mu u_{ij} u_{ij} \quad (1.2)$$

where we have used the Einstein notation for which u_{ij} expresses the deformation of the lattice along the i axis for every displacement along the j axis, and $u_{ij} = \frac{1}{2} (\delta_i u_j + \delta_j u_i)$. λ and μ are constants and known as the Lamé coefficients. When one decomposes the tensor matrix to obtain a new scalar and symmetric traceless tensor the equation 1.2 becomes:

$$F = \frac{B}{2} u_{ii}^2 + \mu \left(u_{ij} - \frac{1}{2} \delta_{ij} u_{kk} \right) \quad (1.3)$$

where δ_{ij} is the Kronecker delta and $B = \lambda + \mu$. The first term of the equation is related with area changes for the lattice, and therefore B is the bulk modulus. Consequently, the second term accounts for deformations in the lattice shape, and μ is the shear modulus. Relating the induced force with the responsible strain, one obtains an expression for the stress tensor.

$$\sigma_{ij} = \frac{\partial F}{\partial u_{ij}} = B u_{kk} \delta_{ij} + 2\mu \left(u_{ij} - \frac{1}{2} \delta_{ij} u_{kk} \right) \quad (1.4)$$

and inverting equation 1.4, one finds an expression of strain as a function of stress:

$$u_{ij} = \frac{1}{4B} \sigma_{kk} \delta_{ij} + \frac{1}{2\mu} \left(\sigma_{ij} - \frac{1}{2} \delta_{ij} \sigma_{kk} \right) \quad (1.5)$$

We can apply equation 1.5 to a 2D hexagonal lattice for which a force T is applied per unit length in opposite sides, with the force $f_y = T dl$ being applied to one wall of the lattice perpendicular to the lattice direction $\vec{e}_y$ and a similar force but with opposite direction $-\vec{e}_y$ is applied to the opposite lattice wall. In this case, it is easy to see that the strain tensor reduces to $\sigma_{yy} = T$, with all other terms in the matrix

being 0, and the resulting strain is:

$$u_{yy} = \left(\frac{1}{4B} + \frac{1}{4\mu} \right) T \quad (1.6)$$

$$u_{xx} = \left(\frac{1}{4B} - \frac{1}{4\mu} \right) T \quad (1.7)$$

$$u_{xy} = u_{yx} = 0 \quad (1.8)$$

This result allow us to express the force along the y axis in a Hooke's Law form ($\sigma_{yy} = K u_{yy}$), simply by inverting 1.6. Consequently, this K constant is the Young modulus of the system, and has value:

$$K = \left(\frac{1}{4B} + \frac{1}{4\mu} \right)^{-1} = \frac{4B\mu}{B + \mu} = \frac{4\mu(\lambda + \mu)}{\lambda + 2\mu} \quad (1.9)$$

In the same way, the minus ratio between the strain along the perpendicular direction to the strain and that along it, is known as the Poisson's ratio σ and can be related to the Lamé coefficients as:

$$\sigma = -\frac{u_{xx}}{u_{yy}} = \frac{B - \mu}{B + \mu} = \frac{\lambda}{\lambda + 2\mu} \quad (1.10)$$

Now, with these basic concepts defined, we can obtain the deformation energy associated for a certain dislocation or disclination for a 2D hexagonal lattice. For that, we need to take into account that each dislocation induces a displacement field $\vec{u}$ around it. And since in a system in equilibrium all forces at all points of the lattice should be zero, we must have that:

$$\lambda \frac{\partial u_{kk}}{\partial x_i} + 2\mu \frac{\partial u_{ij}}{\partial x_j} = 0 \quad (1.11)$$

As we said, the total effect of the displacement field around a dislocation reduces to a certain Burgers vector perpendicular to a lattice direction and with length equal to the lattice unit. We therefore must have that:

$$\oint \vec{u} = a_0 \vec{b}(\vec{r}) \quad (1.12)$$

This equation sets a constraint that allows us to solve equation 1.11 to obtain the displacement field induced by a single dislocation. This way, we obtain for a

dislocation at the origin and with $\vec{b} = b\vec{e}_x$ a displacement field[21]:

$$u_x = \frac{ba_0}{2\pi} \left(\phi + \frac{K}{8\mu} \sin 2\phi \right) \quad (1.13)$$

$$u_y = -\frac{ba_0}{2\pi} \left(\frac{\mu}{\lambda + 2\mu} \log r + \frac{K}{8\mu} \cos 2\phi \right) \quad (1.14)$$

Figure 1.2 shows the subsequent displacement vector field. From both the Figure and Equation 1.14 we can see that as we move farther away from the origin the Burgers vector diverges according to a logarithmic expression. Therefore, a single dislocation induces an infinite stress for the 2D hexagonal lattice and will therefore never be observed in nature. Indeed, when one calculates the stress energy associated with the displacement field, it is:

$$E = E_{el} + E_c = \frac{a_0^2 b^2 K}{8\pi} \log \frac{R}{a} + E_c \quad (1.15)$$

where E_c is the internal energy of the dislocation and the first term increases as the logarithm of the system size R .

To obtain the displacement field for a pair of dislocations we can simply add the vector fields of each of them, since equation 1.11 is linear. The subsequent displacement vector field is shown in Figure 1.3 and this time we can see that there is no stress divergence with size. Instead, for a dislocation pair separated by $\vec{r}_{12}$ and with opposite Burgers vectors ($\vec{b}_1 = -\vec{b}_2$), one obtains that its energy cost is equal to:

$$E = \frac{a_0^2 b^2 K}{4\pi} \log \frac{r_{12}}{q} = \frac{a_0 K}{4\pi} \log \frac{\vec{b} \cdot \vec{r}_{12}}{r_{12}^2} + 2E_c \quad (1.16)$$

and that therefore the two dislocations attract each other through a logarithmic potential.

Now, to consider the melting of the hexatic phase, we need to find an expression of the energy associated to a disclination pair, similar to 1.16. To do so, we consider the case of an arbitrary distribution of both dislocations and disclinations. As it was already mentioned, a disclination is characterized by a rotation of the crystal lattice, that in the case of the 2D hexagonal lattice can either be of $\pm\pi/3$. We therefore write a distribution of disclinations as:

$$s(\vec{s}) = \sum_{\alpha} s_{\alpha} \delta(\vec{r} - \vec{r}_{\alpha}) \quad (1.17)$$

FIGURE 1.2: Displacement field induced by a single dislocation on the 2D triangular lattice centred at the origin and Burgers vector $(a_0, 0)$

FIGURE 1.3: Displacement field induced by a two dislocations on the 2D triangular lattice centred at the origin, Burgers vectors $\vec{b}_1 = -\vec{b}_2 = (a_0, 0)$ and $\vec{r}_{12} = (a_0, 0)$

where s_α is the charge of the disclination and can take ± 1 values. Then, to write the distribution of dislocations we make use of the fact that they can be expressed as a pair of disclinations with opposite rotations.

$$\vec{b}(\vec{r}) = \sum_{\alpha} \left(b_{\alpha x} \vec{i} + b_{\alpha y} \vec{j} \right) \delta(\vec{r} - \vec{r}_\alpha) \quad (1.18)$$

including equation 1.18 as part of 1.17 and applying a Fourier transform, we obtain an expression for the total disclination field.

$$\bar{s}(\vec{q}) = \frac{\pi}{3} + ia_0 [q_y b_x(\vec{q}) - q_x b_y(\vec{q})] \quad (1.19)$$

and now we can solve equation 1.11 by applying the contour conditions:

$$\oint_{\Gamma} du_i = a_0 \sum_{\alpha} b_{\alpha i} \quad (1.20)$$

$$\oint_{\Gamma} d\theta = \frac{\pi}{3} \sum_{\alpha} s_{\alpha}$$

where Γ is a path surrounding the dislocation or disclination. When one calculates the energy of the resulting deformation field, it is [21]:

$$H_{DD} = \frac{K}{2} \left(\frac{\pi}{3} \right)^2 \int \frac{d^2 q}{(2\pi)^2} \bar{s}(\vec{q}) \bar{s}(-\vec{q}) + E_c \sum_{\alpha} b_{\alpha}^2 + E_s \sum_{\alpha} s_{\alpha}^2 \quad (1.21)$$

For the case of the hexatic phase for the 2D hexagonal lattice, that according to the KTHNY theory presents stable dislocations, one can further simplify this expression by only considering the disclination-disclination interactions, with the abundant dislocations only acting as a mean field that affects these pair forces. In doing so, one finds an equation similar to 1.21, but with an effective Young's modulus K_A with value:

$$K_A = \frac{2E_c a^2}{a_0^2} \quad (1.22)$$

where it can be seen than this constant only depends on the core energy of the dislocation E_c and its core size a . Therefore, for the direct space one finds an effective deformation energy based purely on disclinations [21]:

$$H_{\text{Disc}} = -\frac{K_A}{2\pi} \left(\frac{\pi}{3} \right)^2 \frac{1}{2} \sum_{\alpha \neq \alpha'} s_{\alpha} s'_{\alpha} \log \frac{R_{\alpha, \alpha'}}{a_s} + E_s \sum_{\alpha} s_{\alpha}^2 \quad (1.23)$$

So it can be seen that both dislocation pairs and disclination pairs show a logarithmic increase in energy as the distance between the pair increases.

$$\begin{aligned}\frac{U(r)}{k_B T} &= c \log \frac{R}{a} \\ c &= \frac{K a_0^2}{4\pi k_B T} \quad \text{for a dislocation pair} \\ c &= \frac{K_A \pi}{18 k_B T} \quad \text{for a disclination pair}\end{aligned}\tag{1.24}$$

This expression allows us to determine the point at which the solid-hexatic and the hexatic-fluid transitions occur for the 2D system, according to the KTHNY theory. As already mentioned, hexatic phase forms when dislocation pairs unbind, destroying translational order; while the fluid will form at higher temperature with the unbinding of the free dislocations to form disclinations, leading to a fully isotropic system. To calculate the temperature at which this will occur, we calculate average distance between dislocation and disclination pairs, according to equation 1.24.

$$\langle r^2 \rangle = \frac{\int r^2 \exp -\frac{U(r)}{k_B T} d^2 r}{\int \exp -\frac{U(r)}{k_B T} d^2 r} = \frac{2-c}{4-c} a^2\tag{1.25}$$

where the expression only gives physical results for $c > 4$. In the limit $c \rightarrow 4$ the expression diverges, therefore marking the unbinding of the dislocations/disclinations pairs. So by simply calculating the temperature values in equation 1.24 for the $c \rightarrow 4$ limit, one obtains an expression for the solid-hexatic and hexatic-fluid transitions predicted by the KTHNY theory.

$$\begin{aligned}k_B T_m &= \frac{K a_0^2}{16\pi} \\ k_B T_i &= \frac{K_A \pi}{72}\end{aligned}\tag{1.26}$$

So finally, we have arrived to an expression for the transitions temperatures for both the solid-hexatic (T_m) and hexatic-fluid melting (T_i), based purely on the elastic constants of the 2D system under consideration. This is a remarkable achievement of the KTHNY theory and the main reason for its popularity.

From this section we can also conclude that the KTHNY predicts a continuous transition for both the solid-hexatic and the hexatic-fluid transition, as the distances between dislocation and disclination pairs continuously increase until the system becomes plastic. This is the main point of disagreement between the KTHNY theory and other competing models of 2D melting, that rather predict first order transitions, with the fluid phase growing in the bulk of the solid. The existence of these competing melting scenarios makes it impossible to build a general phase diagram with 2D solids, with the actual scenario being case specific and determined by the nature of the interaction potential.

1.2.3 Correlation functions

We end our theoretical introduction to the 2D triangular lattice by describing the behaviour predicted for the translational and rotational correlation functions for the quasi-solid, hexatic and fluid phases. This is fundamental for the computational study of 2D solids, as it supposes the basis for discernment between the three phases.

As equation 1.1 shows, 2D solids show a divergence of the mean square displacement from crystallographic positions $\langle |\vec{u}(R)|^2 \rangle$, and consequently no 2D system is able to have real long range positional correlations. This way, while for the 3D solid Bragg's peaks in the structure appear as δ -functions centred at specific $\vec{G}$ reciprocal lattice vectors, for the 2D quasi-solid they are substituted by exponential singularities $|\vec{q} - \vec{G}|^{-2+\eta}$ where η is given by:

$$\eta = \frac{KT}{\pi\mu a_0^2} \frac{3\mu + \lambda}{2\mu + \lambda} \quad (1.27)$$

For the case of the pair correlation function, the solid presents a power law decay of its envelope function, of the type $e^{-R/\chi}$. However, for temperatures $T > T_m$ the translational correlation function shows an exponential decay typical of fluids. This is therefore a valid method to discern whether a certain 2D arrangement of points corresponds to a quasi-solid, and it was in fact the method used in all our work concerning the polydisperse hard-disk system.

On the other hand, the rotational order for a 2D triangular lattice is usually measured by the correlation function $g_6(r)$.

$$g_6(r) = \langle \psi_6(\vec{r}) \psi_6^*(\vec{r}_0) \rangle \quad (1.28)$$

where $\psi_6(\vec{r})$ is defined as:

$$\psi_6(\vec{r}_l) = \frac{1}{n_l} \sum_{i=1}^{n_l} e^{i6\theta_{li}(\vec{r}_l)} \quad (1.29)$$

where $\vec{r}_l$ is the position of a certain particle l , with n_l neighbours (usually determined through a Voronoi construction) and $\theta_{li}(\vec{r}_l)$ is the angle formed by the line linking l to a neighbour and the x axis.

For the quasi-solid phase and for high $\vec{r}$, this function simply takes a constant value:

$$\log g_6 = -\frac{9}{8\pi} \frac{k_B T}{\mu a_0^2} \quad (1.30)$$

Meanwhile, the hexatic phase shows an algebraic decay as $r^{-\eta_6}$, with the exponent being:

$$\eta_6 = \frac{18k_B T}{\pi K_A} \quad (1.31)$$

where, as before, K_A is the effective Young's modulus describing the interaction between disclinations mediated by dislocations.

Finally, unlike the hexatic phase, the fluid presents an exponential decay for the g_6 orientational correlation function. As before, this provides us with a simple method to differentiate between the hexatic and the fluid phase.

1.2.4 Phase diagram of the hard disk system

As we already mentioned, there are alternative microscopic melting theories for 2D solids and it is therefore not possible to describe a general phase diagram for two dimensional solids. For example, it was proposed by Chui [22] that dislocations could arrange forming grain boundaries that surround the coexisting fluid and solid phase. Alternatively, the melting can also be induced by phase coexistence [23], so that the two phases are in equilibrium separated by an interface, which length is minimized to minimize the surface tension free energy. Therefore, which mechanism

describes the melting of a certain system will depend on different system properties, such as the nature of the interaction between particles particles[24, 25], the particle shape[26], vacancies[27], out of plane fluctuations [28] and several others[29, 30].

In the case of hard disks, probably the simplest model computational physics can consider and one that has been thoroughly studied since the very beginning of the field, it has been surprisingly difficult to establish its melting mechanism, due to the drastic system size dependency of the results. Although for the first numerical study of the system by Alder and Wainwright the melting was found to occur through a first order transition[31], the formulation of the KTHNY theory and the consequent search for experimental systems that follow the mechanism lead to further study of hard disks, with contradictory results. Evidence for both a single first order transition[32, 33] and two continuous transitions was found[34], with many other publications being unable to distinguish between the two scenarios[35, 36].

This debate appears now to be settled, thanks to the work by Bernard and Krauth in 2011[37]. Identifying the cause of this disagreement between different numerical calculations to be the drastic system size effect in the equation of state, the authors were able to equilibrate unprecedented configurations with 1024^2 particles, thanks to the use of the Event Chain Monte Carlo algorithm, that will be presented later in this Introduction section and has proven to improve equilibration times for high packing fraction configurations of hard-disk of all of its competitors by at least one order of magnitude[38]. Thanks to this considerable system size, the authors were able to observe the coexistence of an ordered and disordered phase directly, of packing fractions $\eta \sim 0.700$ and $\eta \sim 0.716$ respectively. They were also able to identify the typical $\sqrt{N}$ system size dependence for the over-compression work calculated by a Mayer-Wood Loop construction of the equation of state, typical of first-order transitions in 2D, for which the interfacial energy is directly proportional to the length of the inter-phase.

Then, by studying the orientational correlation of this ordered coexisting phase, they determined it to be the hexatic and not the solid, therefore confirming that hard-disks present a stable hexatic phase. However, the continuous transition to the solid occurs at packing fraction $\eta \sim 0.720$, and consequently they found the resulting single phase packing fraction stability range of the hexatic phase to only be $\eta \in [0.716, 0.720]$. Although this range was larger than the packing fraction

fluctuations they observed for the coexisting ordered phase, Russo *et al.* [39] proved that this stability range is so small that adding an insignificant concentration (less than 1%) of a smaller particle in the system leads to the complete disappearance of the hexatic phase, with the solid melting directly to the fluid through a first order transition. This result begs the question of whether it would be possible to observe experimentally a phase stable only in such specific conditions.

In this respect, the first experimental observation of melting for particles interacting through hard core repulsion was performed by Lube *et al.* [40], although in this case the particles also presented a small exclusion square attractive potential. A hexatic phase was clearly identified, but in later work by Karnchanaphanurach[41], using a colloidal system with perfect hard-core interaction, this result was not reproduced. However, more relevant to this work is the recent experimental study of the hard-disk phase diagram by analysing the sedimentation of spherical hard colloids in a tilted plane, performed by Dullens *et al.*[42]. In this case the assembling colloids were synthesized by forming a dispersion of melamine formaldehyde droplets in a water-alcohol solution. The colloidal dispersion was then packed in a glass cell and allowed to sediment for several weeks, with the main controlling parameter of the experiment being the angle α at which the glass cells were tilted. Once an equilibrium configuration was reached, using imaging techniques, they were able to obtain the equation of state as a function of pressure and classify the different regions of the formed sediment according to the behaviour of its positional and orientational order parameters. Their results confirmed the existence of stable liquid, hexatic and solid phases, with a hexatic-liquid coexistence region of $\eta \approx [0.68 - 0.70]$. More interestingly, the single phase stability packing fraction range for the hexatic phase was found to be a surprising $\eta \approx [0.70 - 0.73]$, an order of magnitude higher with respect to theoretical calculations. This disagreement between theory and experiment is what lead to our study of the polydisperse hard-disk phase diagram, to determine whether the particle size polydispersity (unavoidable in any experimental set-up) could be responsible for this surprising stabilization of the hexatic phase.

1.3 Methods

1.3.1 Free energy calculation in solids

In Chapter 2 of this thesis we calculate the free energy of the crystal phases formed by the self assembly of mobile DNA coated colloids. Therefore in this section we offer a small description of how the free energies of solid phases are calculated in computational mechanics. Generally speaking, the determination of free energies always offers a challenge, since unlike some other thermodynamic functions of a system (such as its volume, pressure, internal energy, etc,) it is not possible to calculate the free energy of a system just knowing its configuration and the potential that describes the interactions between its particles. This is because the calculation of free energy requires the determination of entropy and, according to Boltzmann Law, entropy depends on the partition function of the system that we are studying, which requires the calculation of a very complex multidimensional integral for all space with as many variables as degrees of freedom has our system ($3N$). It can therefore only be calculated analytically for some very specific cases, in which particles are completely uncorrelated to each other, so that one can apply separation of variables.

Fortunately, thermodynamic relations give expressions for the derivative of free energy expressed as a function of other magnitudes that can be easily measured. For example, from the general expression $dF = PdV - SdT$ and the fact that the Helmholtz Free Energy is an exact differential, it follows that $\left(\frac{\partial F}{\partial V}\right)_T = P$. This allows us to calculate the difference in free energy between two systems by performing thermodynamic integration, in which we transition from one system to the other by changing a certain control parameter in several steps, in order to perform a numerical integration of the free energy change. For example, using thermodynamic integration we can calculate the difference in the Helmholtz free energy between two different density gases by progressively changing the system volume while keeping temperature constant, so that we can calculate $\Delta F = \int PdV$.

Consequently, in order to obtain absolute values for the free energy of a system we just need to perform thermodynamic integration to a reference system for which its partition function is known. In the case of gases, this reference is usually taken to be an ideal gas configuration, for which entropy is simply $s = k_B T \log V$.

Alternatively, if the system we want to study is not very compact, one can directly apply the Widom insertion method[43], that calculates the chemical potential as a function of the average interaction energy of the system when one tries to introduce a new particle in it. However, neither of these strategies is useful when calculating the free energies of solid phases. First, it is not possible to perform thermodynamic integration to the ideal gas because the transformation from the solid to the gas would involve a first order transition, during which the derivative of the free energy diverges, making it impossible to perform any numerical integration. And neither it is possible to use the Widom insertion method, since the high packing fraction of solids make the insertion of a new particle into its structure extremely unlikely, and unrealistic simulation times would be required to sample the average interaction potential precisely.

Fortunately, there is an alternative reference system to calculate free energies of solids: the Einstein crystal[44]. An Einstein Crystal is a crystal lattice stabilized merely by the effect of independent harmonic oscillators that bind each particle to its corresponding lattice position. In this case, the partition function of the system can be calculated by considering each particle a quantum harmonic oscillator and calculating all possible arrangements of the energy quanta.

To perform the thermodynamic integration between our solid and the Einstein crystal, the method uses a switching parameter λ , that will gradually transform the potential energy function of the real solid into that of an Einstein crystal. It will therefore take the form:

$$\tilde{U}(\mathbf{r}^N) = U(\mathbf{r}^N) + (1 - \lambda) [U(\mathbf{r}^N) - U(\mathbf{r}_0^N)] + \lambda \sum_{i=1}^N \alpha_i (\mathbf{r}_i - \mathbf{r}_{0,i})^2 \quad (1.32)$$

where $\mathbf{r}_{0,i}^N$ is the lattice position of particle i , $U(\mathbf{r}_0^N)$ is the static contribution to the potential energy, α_i is the Einstein crystal spring constant binding the particle i to its lattice site, and, as mentioned, λ is the switching parameter. It can be seen that if $\lambda = 0$, the system behaves normally, while if $\lambda = 1$ it becomes a perfect Einstein crystal. The method then calculates the free energy difference between the Einstein crystal and the real system as:

$$\begin{aligned}
F &= F_{\text{Ein}} + \int_{\lambda=1}^{\lambda=0} d\lambda \left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_{\lambda} \\
&= F_{\text{Ein}} + \int_{\lambda=1}^{\lambda=0} \left\langle \sum_{i=1}^N \alpha_i (\mathbf{r}_i - \mathbf{r}_{0,1})^2 - [U(\mathbf{r}^N) - U(\mathbf{r}_0^N)] \right\rangle_{\lambda}
\end{aligned} \tag{1.33}$$

Finally, one can calculate the absolute free energy of an Einstein crystal according to expression:

$$F_{\text{Ein}} = U(\mathbf{r}_0^N) - \frac{d}{2\beta} \sum_{i=1}^N \ln \left(\frac{\pi}{\alpha_i \beta} \right) \tag{1.34}$$

where d is the dimension of the system and β is $(k_B T)^{-1}$.

This implementation of the Einstein crystal method is not, however, the one we use for the calculations in Chapter 2. This is because the version of the method we just described is only useful when the potential energy function of the system we want to study is continuous, so that it can be gradually switched off and transformed into the pure Einstein crystal expression.

Since our mobile DNA colloids interact (in part) as hard spheres, we need an alternative method to calculate the free energy. We therefore apply an alternative version of the Einstein crystal method[45], in which instead of gradually switching off the potential, the method increases gradually the spring constant λ of the system while maintaining the original potential, such that:

$$U(\lambda) = U_0 + \lambda U = U_0 + \lambda \sum_{i=1}^N (\mathbf{r}_i - \mathbf{r}_{0,i})^2 \tag{1.35}$$

In this case, the free energy difference between the system with λ parameter and the original system can be expressed as:

$$F_{\text{HS}} = F(\lambda_{\text{max}}) - \int_0^{\lambda_{\text{max}}} d\lambda \langle U(\mathbf{r}^N, \lambda) \rangle_{\lambda} \tag{1.36}$$

For sufficiently large λ_{max} , all other interactions between particles will be negligible and the system will behave as a pure Einstein crystal. Therefore, the challenge of this approach, and the reason it is not the one preferred for continuous potentials,

is finding a proper value for $\lambda_{\max}$, large enough so that the system behaves as an Einstein crystal, but not too large, as this would increase the error involved in the numerical integration.

1.3.2 Event chain Monte Carlo

For the study of the polydisperse hard-disk phase diagram and the effect of the distribution shape in it (Chapters 3 and 4), we required an efficient method to equilibrate high packing fraction configurations of the hard-disk system. In our work we decide to use the relatively new event chain Monte Carlo algorithm (ECMC), that has consistently proven to significantly improve equilibration computational times when compared with other alternative sampling techniques [38]. The main characteristic of the method, and what differentiates it from local Monte Carlo (LMC), is that it is configuration rejection free: that is, all configurations generated during the algorithm will be accepted.

Although a general formulation of event chain Monte Carlo has already been proposed [46], the method was first described for hard spheres and hard disks [47], where its implementation is much simpler. In this case, the method occurs as a deterministic chain of events, each event consisting in the propagation of a single disk along a certain direction until a collision with another particle occurs. At this moment, the initial particle will stop and the one responsible for the collision will be propagated in the same manner. This process will continue until the total displacement for the chain reaches a certain threshold l , that is the only controlling parameter of the algorithm, which value needs to be optimized to guarantee the fastest possible equilibration. Therefore, the working of the method is almost deterministic, with only the initial particle of the chain and the direction along which it will be propagated (θ) being randomly determined at beginning of each iteration.

The high performance of ECMC is not only caused by its coordinated particle movement (the lack of which makes LMC so inefficient in simulations of high packing fraction hard-disks), but also the fact that the algorithm allows the breaking of detailed balance, while still maintaining global balance. Algorithms that break detailed balance allow probability flows in the configuration space and can lead

to faster equilibration[48]. In the case of hard disks, this break of detailed balance is achieved by limiting the propagation direction sampling to a smaller range than the full unitary circle. The smallest range that guarantees two independent propagation directions and therefore global balance and ergodicity is that given by the first quadrant $\theta = [0, 2\pi]$. This version of the ECMC algorithm is the one we implemented for our polydisperse hard-disk simulations.

Finally, another useful feature of ECMC is that it allows the calculation of the system's pressure on the go, using only information that is obtained with no additional cost during simulation. The used expression is shown in equation 1.37 and is based on the ratio between the total displacement occurred during an chain of events ($x_{\text{final}} - x_{\text{init}}$) caused both by the movement of the particles and the advancements occurred when a lift to a new event occurs, due to the radius of the particles; divided by the total propagation caused by the event chain, that is always l . This magnitude needs to be sampled over a large quantity of chains, to obtain the average for all propagation directions and regions of the system.

$$\beta P = \rho \left\langle \frac{x_{\text{final}} - x_{\text{init}}}{l} \right\rangle \quad (1.37)$$

It can be seen in Eq. 1.37 that for the ideal gas case, when particles are very far from each other and never collide, no lifts occur and therefore the two quantities become equal and the general equation of state of ideal gases is returned, proving the consistency of the equation.

This equation was used to obtain pressure for all our simulations for the polydisperse hard-disk system in Chapters 3 and 4.

Chapter 2

Entropy stabilization of floppy crystals of mobile DNA-coated colloids

2.1 Introduction

Nucleic acids are ubiquitous in nature because of their capability of encoding large amounts of information via canonical Watson-Crick base-pairing interactions [49]. With the help of chemical methods to make synthetic oligonucleotides of arbitrary sequences, one can use specific binding interactions between single-stranded DNA (ssDNA) chains to program selective interactions between different colloidal particles. For example, one can graft DNA linkers to the surface of colloidal particles with open ends of ssDNAs. These ssDNA tails serve as “sticky ends” that bind specifically to other colloids coated with ssDNA tails of complementary sequence, which offers a novel way of manipulating the self-assembly of colloidal particles [50, 51]. By using DNA-coated colloids (DNACCs), a number of ordered crystals [52–58] and self-assembled “colloidal molecules” [59] have been obtained in experiments, while the self-assembly mechanism of DNACCs is still not well understood [57, 60, 61]. For example, the diffusionless transformation from a floppy crystal to the other compact crystal has been observed in experimental systems while the underlying mechanism remains not fully resolved [61].

Recently, a novel system of mobile DNA-coated colloids (mDNACCs) was introduced that displays qualitatively new properties [62, 63]. Compared with immobile DNA-coated colloidal systems, mDNACCs have a broader temperature window for self-assembly, and therefore allow the better control over the assembly process [62]. Mobility of DNA linkers also allows particles to more easily roll around each other and rearrange [62, 64], without grafting of very high density. Moreover, unlike colloids with patches in specific locations [59, 64], the interaction in mDNACCs is intrinsically a many-body potential, which could be employed to control the “valency” of particles without patches by tuning nonspecific repulsions between the particles [1, 65]. However, despite these novel properties and potential applications, the principles determining the collective self-assembly of mDNACCs remain unclear. To this end, we study the equilibrium self-assembly in binary systems of mDNACCs with complementary sequences. We construct the phase diagrams for systems in both two (2D) and three dimensional (3D) spaces. At low pressure, we find floppy square and CsCl crystals in 2D and 3D systems, respectively, which are more stable than the corresponding compact hexagonal and CuAu crystals. This behavior holds for a large range of binding strengths, even in the infinite binding strength limit. We demonstrate that these floppy crystals are stabilized by vibrational entropy, and “floppy” modes play an important role in the infinite binding strength limit.

2.2 Finite binding energy mDNACC

We consider a binary system of N colloids A and B , coated with mobile DNA linkers. Each linker terminates in a short ssDNA sequence, and particles of type A and B are coated with ssDNA linkers of complementary sequences. The parameters are chosen to be the same as for the system without nonspecific repulsions in Ref. [65]: the system has equal numbers of A and B particles $N_A = N_B = N/2$; each colloid is modelled as a hard sphere with diameter $\sigma = 200 \text{ nm}$, on which $n = 70$ double-stranded DNA (dsDNA) linkers of length $L = 20 \text{ nm}$ terminating in a short ssDNA sequence are grafted. The effective interaction energy βU consists of an attraction part coming from the binding of the linkers βF_{att} , and a repulsive part due to the excluded volume interaction βF_{rep} , where $\beta = 1/kT$ with k and T being the Boltzmann constant and temperature of the system, respectively. Using

a mean-field approach [65–67], βF_{att} can be written as

$$\beta F_{\text{att}} = \sum_{i=1}^N n[\ln p_i + (1 - p_i)/2]. \quad (2.1)$$

Here p_i is the probability that a linker on particle i is unbound, satisfying the following set of equations

$$p_i + \sum_j p_i p_j e^{-\beta \Delta G_{ij}} = 1. \quad (2.2)$$

$\beta \Delta G_{ij}(\mathbf{r}_i, \mathbf{r}_j)$ is the free energy for the formation of a bond between a pair of particles $i - j$, which can be written as

$$\beta \Delta G_{ij}(\mathbf{r}_i, \mathbf{r}_j) = \beta \Delta G_0 + \beta \Delta G_{\text{cnf}}(\mathbf{r}_i, \mathbf{r}_j), \quad (2.3)$$

where $\beta \Delta G_0$ is the binding strength (hybridization free-energy) of two complementary ssDNAs in solution, depending on the DNA sequence and being a function of temperature and salt concentration, and $\mathbf{r}_{i/j}$ is the position of particle i/j . The smaller the value of $\beta \Delta G_0$ implies, the stronger the binding strength. $\beta \Delta G_{\text{cnf}}$ is the configurational cost for bond formation that can be calculated analytically for $L \ll \sigma$ [65]. The repulsion originates from excluded interactions between DNA linkers and hard-sphere cores, and it is of general form [65]

$$\beta F_{\text{rep}} = -\ln \left(\frac{\Omega(\{\mathbf{r}_i\})}{\Omega_{\text{free}}} \right), \quad (2.4)$$

where the partition function $\Omega(\{\mathbf{r}_i\})$ counts all accessible states of linkers given the positions of the colloids, and Ω_{free} is the value of $\Omega(\{\mathbf{r}_i\})$ when colloids are separated from each other by an infinite distance. Then using $\beta U = \beta F_{\text{att}} + \beta F_{\text{rep}}$ above, we perform extensive Monte Carlo simulations for systems of $N \approx 500$ mDNACCs; and by free energy calculations [68], we construct the phase diagrams in both 2D and 3D. In our 2D system of mDNACCs, colloids are moving in a 2D plane while DNA linkers rotate in a 3D space, which can model the self-assembly of mDNACCs at the bottom of an experimental chamber [61, 62] or the liquid-liquid interface [69].

Figure 2.1 shows phase diagrams in the area/packing fraction η - binding strength $\beta \Delta G_0$ representation for mDNACC systems in 2D and 3D, respectively. Since the

FIGURE 2.1: Phase diagrams for mobile DNA-coated colloids in two and three dimensional space, in the area/packing fraction η - binding strength $\beta\Delta G_0$ plane. Black circles mark the position of the triple points obtained by extrapolation. The dashed line is an estimate of the transition from the hexagonal phase to the disordered hexagonal phase. Snapshots are shown for the ordered crystal phases. Gray areas represent the coexistence regions.

numbers of complimentary linkers on both A and B colloids are the same, we do not consider crystals of asymmetric stoichiometry. One can see that at low density, because of entropy, the systems remain in a disordered fluid phase [70, 71], but when increasing the density, ordered crystals form. For weak ssDNA bindings, for which βF_{rep} does not allow any linkers bond to form, mDNACCs crystallize into disordered crystals, i.e. disordered hexagonal crystal in 2D and face-centered cubic crystal in 3D, in which particles are located on ordered lattices but types of particles are random. This transition is purely determined by the interactions between the hard colloids and is therefore similar to those of hard-disks and hard spheres respectively. However, if increasing the ssDNA binding strength the AB colloids start binding with each other and when compressed the fluid arranges into ordered crystals phases. As shown in Fig. 2.1, for $\beta\Delta G_0 \lesssim -10.5$, when increasing the density of 2D mDNACC systems, a floppy ordered square crystal first crystallize from the fluid (Fig. 2.1a), and similarly in 3D systems for $\beta\Delta G_0 \lesssim -9$, a floppy CsCl crystal forms at relatively low density (Fig. 2.1b). When further increasing the density, the compact crystals become stable because of their high packing efficiency, i.e. hexagonal crystals in 2D, CuAu and tetragonal crystals in 3D. From the 2D phase diagram we can also see that there is no equilibrium fluid phase for binding energies stronger than $\beta\Delta G_0 \approx -10.75$. This is because for larger binding energies the system arranges to the open square lattice even at zero pressure, which demonstrates how fast is the switch on of the mDNACCs interactions, due to its many body nature.

The phase diagram of 3D mDNACC system is qualitatively similar to that of oppositely charged colloids [72], however, as we shall see, the stabilization of the CsCl phase occurs for completely different reasons than for charged colloids, for which the repulsion between same charged particles favours enthalpically the less compact phase.

As we explained in the introduction, stabilization of floppy crystals can be due to either enthalpy or entropy. For a single component system of DNA-coated nanoparticles [73], including the configurational entropy of the linkers in the effective potential (as adopted in this work), the b.c.c. crystal is believed to be favored over the f.c.c. crystal due to its lower enthalpy [73, 74]. And for a binary system of colloids coated with very short DNA linkers, it has been mentioned that the

CsCl structure is favored over the CuAu structure by virtue of its higher vibrational entropy [54]. Despite these examples, stabilization mechanisms for floppy crystals are largely undistinguished for DNACCs [57, 61, 71, 75]. For this case, to explore the stabilization mechanism for the floppy crystals in mDNACCs, we compare the potential energy and vibrational entropy at coexistence. For simplicity, we use $S = S_{\text{vib}}$ and neglect other contributions to the entropy since they are the same for phases being compared. The effective potential energy βU is calculated directly in Monte Carlo simulations. The vibrational entropy per particle is obtained as $S_{\text{vib}}/kN = \beta U/N - \beta F/N$, in which the Helmholtz free energy βF is calculated by thermodynamic integrations to the reference Einstein crystal. As shown in Fig. 2.2, we see that the floppy square/CsCl crystal has higher vibrational entropy than the compact hexagonal/CuAu crystal, while the effective energy for the floppy square/CsCl crystal is even slightly higher. This suggests that the vibrational entropy stabilizes the observed floppy crystals in mDNACCs.

Finally, a remarkable feature of these phase diagrams is the wide range of binding strengths for which floppy square and CsCl crystals are stable. It would be of interest to determine how wide this stability range actually is, as in experiments the hybridization free-energy $\beta\Delta G_0$ is very sensitive to external conditions, such as the salt concentration and temperature [1], and tuning $\beta\Delta G_0$ precisely can be challenging. Therefore, when the crystallization of DNACCs is observed, the hybridization free-energy is usually very strong, when almost all linkers are bound. Therefore, to examine whether entropy stabilized floppy crystals of mDNACCs exist at conditions close to experiments, we simulate mDNACCs at $\beta\Delta G_0 \rightarrow -\infty$, which are essentially the very bottom of the phase diagrams in Fig. 2.1.

2.3 Infinite binding energy system

The above method for simulating mDNACCs can only be used to simulate mDNACCs with moderate DNA hybridization free-energies and many unbound ssDNA linkers [65–67]. For the limit $\beta\Delta G_0 \rightarrow -\infty$, Eq. 2.1 leads to a divergent βF_{att} . This is due to the use of the Stirling approximation, that does not hold when the amount of free linkers on a colloid approaches 0 ($x_\alpha \approx 0$). We therefore modify the original method to consider the very low temperature region, when all linkers are bound and the system can therefore be considered athermal. The derivation

FIGURE 2.2: Interaction potential $\beta U/N$ and vibrational entropy S_{vib}/kN versus the binding strength $\beta \Delta G_0$, at the square - hexagonal phase coexistence in 2D, and CsCl - CuAu phase coexistence in 3D.

of the new method starts with the expression of the partition function accounting for the binding of the linkers in the mDNACCs system used by Angioletti-Uberti:

$$\begin{aligned}
 Z_{\text{bind}} &= \sum_{\{x_{\alpha\beta}\}} W(\{x_{\alpha\beta}\}) \prod_{\alpha < \beta} \Theta_{\alpha\beta}^{x_{\alpha\beta}} \\
 W(\{x_{\alpha\beta}\}) &= \prod_{\alpha} \frac{n_{\alpha}!}{\binom{n_{\alpha} - \sum_{\beta} x_{\alpha\beta}}{n_{\alpha}}} \prod_{\alpha < \beta} \frac{1}{x_{\alpha\beta}!}
 \end{aligned} \tag{2.5}$$

where the $W(\{x_{\alpha\beta}\})$ accounts for all possible arrangements that lead to $x_{\alpha\beta}$ DNA bonds between two colloids and $\Theta_{\alpha\beta}^{x_{\alpha\beta}}$ is the binding strength. We obtain the effective binding strength between two colloids by calculating the average binding

FIGURE 2.3: Simulation results for mobile DNA-coated colloids at the infinite binding strength limit $\beta\Delta G_0 \rightarrow -\infty$. (a,b) Probability density distribution of area/packing fraction η of the system versus pressure P in 2D (a) and 3D (b). (c) Interaction potential $\beta U/N$ and enthalpy $\beta H/N$ versus the pressure P . Solid curves are obtained by data fitting.

free energy over the surface of the two colloids.

$$\begin{aligned} \Theta_{\alpha\beta}(\vec{R}_A, \vec{R}_B) &= \langle \exp(\beta\Delta) \rangle_{|\vec{R}_A, \vec{R}_B} = \frac{\int_{S_A, S_B} \exp[-\beta\Delta G_{\alpha\beta}] d\vec{r}_\alpha d\vec{r}_\beta}{S_A S_B} \\ &= \exp[-\beta\Delta G_0] \frac{\int_{S_A, S_B} \exp[-\beta\Delta G_{\text{cnf}}(\vec{r}_\alpha, \vec{r}_\beta)] d\vec{r}_\alpha d\vec{r}_\beta}{S_A S_B} = \Theta_0 \Theta_{\alpha\beta}^*(\vec{R}_A, \vec{R}_B) \end{aligned} \quad (2.6)$$

This average is calculated considering that all colloids are fixed at their configuration positions and runs through all possible $\vec{r}_\alpha$ and $\vec{r}_\beta$ grafting positions of linkers in each colloid. We also express the binding energy as the sum of a linkers binding ΔG_0 and a configuration contribution $\Delta G_{\text{cnf}}(\vec{r}_\alpha, \vec{r}_\beta)$.

If we substitute 2.6 into 2.5 and consider the case when all linkers are bound ($n_\alpha = \sum_\beta x_{\alpha\beta}$), we obtain the expression for the partition function in the infinitely strong binding regime:

$$Z_{\text{bind}} = \sum_{\{x_{\alpha\beta}\}} \prod_{\alpha} n_{\alpha}! \prod_{\alpha < \beta} \Theta_0^{x_{\alpha\beta}} \prod_{\alpha < \beta} \frac{1}{x_{\alpha\beta}!} \Theta_{\alpha\beta}^{*x_{\alpha\beta}} \quad (2.7)$$

Since we take a fixed amount of linkers per colloid, the prefactor $Z_0 = \prod_{\alpha} n_{\alpha}! \prod_{\alpha < \beta} \Theta_0^{x_{\alpha\beta}}$ will be constant for all configurations. Which simplifies the partition function to just:

$$Z_{\text{bind}} = Z_0 \sum_{\{x_{\alpha\beta}\}} \prod_{\alpha < \beta} \frac{1}{x_{\alpha\beta}!} \Theta_{\alpha\beta}^{*x_{\alpha\beta}} \quad (2.8)$$

Now we calculate the effective binding energy from the partition function, applying the Stirling approximation.

$$\begin{aligned} Z &= \sum_{\{x_{\alpha\beta}\}} \exp(-\beta U_{\text{bind}}(\{x_{\alpha\beta}\})) \\ \beta U_{\text{bind}}(\{x_{\alpha\beta}\}) &= \sum_{\alpha < \beta} x_{\alpha\beta} (\log x_{\alpha\beta} - 1 - \log \Theta_{\alpha\beta}^*) - \log Z_0 \end{aligned} \quad (2.9)$$

This use of the Stirling approximation does not imply any risk of divergence of the potential (that would occur if $x_{\alpha\beta}$ approached zero), since the Lagrange multiplier enforces that all linkers are bound at every moment, independently of the βF_{rep} associated with the bonding of colloids α and β . The only condition required will be therefore for each colloid to have enough linkers so $x_{\alpha\beta}$ will be larger for all its neighbours, which will be true for high grafting densities. As for the low binding energy case, to this binding energy we need to subtract that due to the excluded volume interaction of the DNA links, that destabilizes the crystal with respect the fluid phase:

$$\beta U_{\text{rep}} = - \sum_{\alpha} n_{\alpha} \log \frac{\Omega_{\alpha}}{\Omega_0} \quad (2.10)$$

where Ω_{α} is the volume available to the linkers grafted to α in the configuration, and Ω_0 that for an isolated colloid. And finally, the configuration binding strength $\Theta_{x_{\alpha\beta}}$ is given by $\Xi_{\alpha\beta} = \Omega_{\alpha\beta} / \Omega_{\alpha} \Omega_{\beta}$, where $\Omega_{\alpha\beta}$ is the volume available to the two linked DNA strands. This gives us a general expression for the potential energy of the system.

$$\beta U_{\text{inf}}(\{x_{\alpha\beta}\}) = \sum_{\alpha < \beta} x_{\alpha\beta} (\log x_{\alpha\beta} - 1 - \log \Xi_{\alpha\beta}) \quad (2.11)$$

When analysing the Hessian of the potential energy function $H_{ij} = \partial^2 U_{\text{inf}} / \partial x_{\alpha\beta} \partial x_{\alpha'\beta'}$ one finds that it is diagonal and positive definite, and therefore βU is a convex function with a single minimum. Therefore, minimizing this energy with respect $\{x_{\alpha\beta}\}$ guarantees obtaining the global minimum energy for the mDNACCs given the infinite binding energy constraints.

The equilibrium linker distribution is therefore given by minimizing the Lagrange function $\mathcal{L} = \beta U_{\text{inf}} + \sum_{\alpha} \lambda_{\alpha} (n_{\alpha} - \sum_{\gamma} x_{\alpha\gamma})$ via

$$\frac{\partial}{\partial x_{\alpha\beta}} \left[\beta U_{\text{inf}}(\{x_{\alpha\beta}\}) + \sum_{\alpha} \lambda_{\alpha} \left(n_{\alpha} - \sum_{\gamma} x_{\alpha\gamma} \right) \right] = 0 \quad (2.12)$$

Substituting Eq. 2.11 and $x_{\alpha\beta} = x_{\beta\alpha}$ into Eq. 2.12, we get

$$x_{\alpha\beta} = e^{\lambda_{\alpha} + \lambda_{\beta}} \Xi_{\alpha\beta}, \quad (2.13)$$

where coefficients $\{\lambda_{\alpha}\}$ satisfy the constraint of all linkers being bound, and can be solved by, e.g., self-consistent iterations with

$$e^{\lambda_{\alpha}} = \frac{n_{\alpha}}{\sum_{\gamma} e^{\lambda_{\gamma}} \Xi_{\alpha\gamma}}, \text{ for all } \alpha. \quad (2.14)$$

It can be proven that Eq. 2.13 gives the global minimum of βU_{inf} under the constraint set by all linkers being bound [17]. We then use $\beta U = \beta U_{\text{inf}} + \ln Z_0$ with the same constraint to perform Monte Carlo simulations in the NPT ensemble for both 2D and 3D mDNACC systems. Results are shown in Fig. 2.3 for a system of $N = 100$ in 2D and $N = 256$ in 3D. From the probability density distributions of packing fraction of the systems in Fig. 2.3a,b, we see that floppy crystals do exist in the low pressure region. Coexistence of floppy and compact phases occurs at about $P \simeq 1.0kT/\sigma^2$ and $3.0kT/\sigma^3$ for 2D and 3D systems, respectively.

This coexistence region was confirmed by the presence of two maxima in the packing fraction probability distribution of the system. This can be more clearly seen from Fig. 2.4, that is similar to Fig. 2.3.a but includes the average packing fraction of the system as a function of time, as well as the average packing fraction for the square and hexagonal phases, that are obtained by fitting the distribution to the sum of two Gaussian curves of the form $c \exp(-(\eta - \eta_0)^2/2\nu^2)$. From this figure we can see that it is not possible to determine whether the transition is first order or continuous from the equation of state of the full system (red line in Fig. 2.4), since the curve is not smooth enough due to the large fluctuations of the simulation box volume. However, when we plot η_0 for each of our fit Gaussian curves, it can be seen that this lack of definition in the equation of state is caused by the transition of the system between two stable phases (the open square and the compact hexagonal), which in turn confirms the phase transition to be first order.

FIGURE 2.4: Probability distribution function for packing fraction for the 2D infinite binding mDNACCs system, showing the average of the distribution (red line) and the center of two fit Gaussian distributions that correspond to the square (green line) and the hexagonal phase (blue line). A clear coexistence region can be distinguished for pressures between $1.0kT/\sigma^2$ and $1.5kT/\sigma^2$.

Unlike our previous case, the method described in section 1.3.1 to calculate the free energy of the solid phases will not be applicable in the infinite binding limit. This is because the interaction between colloids being infinitely strong, it will never be the case that the potential energy function of the system will be approximately equal to that of a Einstein crystal, independently of how high the λ value used for equation 1.36. Consequently, we can only use thermodynamic functions that do not depend on the system's partition function to study the stabilization of floppy crystals. Therefore, we first compare the internal energy per particle $\beta U/N$ for both floppy and compact crystals, as shown in Fig. 2.3c, near the phase coexistence, and find them to be very similar in the case of the hexagonal and square crystals, and that the CuAu crystal has a slightly lower potential energy than the CsCl crystal. Since the floppy and compact phases have equal chemical potential $\beta(\mu_A N_A + \mu_B N_B)/N = \beta U/N + \beta PV/N - S_{\text{vib}}/kN$ at the coexistence, to compare the vibrational entropy of the floppy and compact phases, we calculate the enthalpy per particle $\beta H/N = \beta U/N + \beta PV/N$. Figure 2.3c shows that enthalpy for the floppy phase is higher than that of the corresponding compact phase in both 2D and 3D near coexistence. This implies that the floppy crystals have higher vibrational entropy near phase coexistence. Thus vibrational entropy also stabilizes the

floppy phases at the limit $\beta\Delta G_0 \rightarrow -\infty$.

In other works it has been shown that a vibrational mode analysis reveals that “floppy” modes, namely collective motions that do not change the interaction energy, play an important role for the diffusionless CsCl-CuAu transition in a system of DNACCs with short linkers [60]. For mDNACCs in the infinite binding limit, we observe that distances between neighboring $A - B$ pairs are very short ($d_{AB} \simeq 1.04\sigma$), and that the effective interaction energy for the floppy and compact crystals are very similar at coexistence, especially for the 2D system. These motivate us to explore the role of “floppy” modes in the stabilization of floppy crystals. We tried a vibrational mode analysis within the harmonic approximation, to evaluate the entropy [14, 15] and count the floppy modes [60]. We calculated approximately the dynamical matrix by measuring displacement correlations between particles [76]. However, for system sizes we were able to simulate, we observe large volume fluctuations which invalidate the harmonic approximation.

Instead, we approximate the system by a *sticks and balls* model: every colloid is bonded with a fixed number of rigid colloids of the opposite type and distances between all $A - B$ pairs are fixed. This approximation is exact for mDNACC systems with $\beta\Delta G_0 \rightarrow -\infty$ and linker length $L \rightarrow 0$, for which the effective potential βU_{inf} is a constant and the system allows only “floppy” moves. We conducted NPT simulations for the model with $N = 100$ in 2D. To illustrate how the model works, in Fig. 2.5a we show two consecutive moves from an initial square lattice. The probability density distribution of area fraction of the system is shown in Fig. 2.5b for different pressure. We see that, for pressure below $P \simeq 4kT/\sigma^2$, “floppy” moves favour the low density square crystal over the dense hexagonal crystal. This result supports that vibrational entropy associated with “floppy” modes play an important role in the stabilization of observed floppy phases in the infinite binding limit. This is related to the fact that mechanically floppy networks can become rigid (stable) when thermal effects are present [77].

Unfortunately, we cannot use this same approach to study the transition between the CsCl and CuAu phase for the 3D system. This is because the CsCl forms by the piling of square lattice planes of A and B colloids in such a way that each A colloid occupies the hole left by four B colloids and vice-versa; while the CuAu forms by the piling of hexagonal lattice planes of A and B colloids in such a way that each

FIGURE 2.5: (a) Two consecutive “floppy” moves of the sticky-sphere model from an initial square lattice. In simulations only particles in a single column or row are allowed to move relatively to other particles at a time. Here we show a column move in the middle (top) followed by a row move in the third row (bottom). (b) Probability density distribution of area fraction η of the system versus pressure P for the sticky-sphere model in 2D, for which distances between neighboring A and B particles are fixed at $d_{AB} = 1.04\sigma$.

A colloid occupies the hole left by three B colloids, with the overall piling following an ABC order. It is not possible to transition between these two crystal structures without breaking any bonds and consequently no *sticks and balls* model will be able to simulate this process. This does not represent a problem for our model for mDNACCs in the infinite binding energy limit, because the Lagrange multiplier only enforces that all linkers are bonded at all moments and this constraint does not stop a particle from changing its neighbours.

2.4 Conclusions

In conclusion, we have studied the self-assembly of a binary system of colloids coated with complementary mobile DNA linkers. We construct the phase diagrams for both 2D and 3D mDNACC systems, in which we observe stable floppy square/CsCl crystals at strong enough ssDNA binding strength. We also derive an effective potential for the system in the infinite ssDNA binding strength limit, and

formulate a Monte Carlo method to simulate the effect of entropy for the system in this limit. Our results show that even for the infinitely strong ssDNA binding limit, floppy crystals are still more stable than compact ones at low pressure because of their higher vibrational entropy. This suggests that the strong ssDNA binding limit of mDNACC systems is different from the conventional atomistic or molecular systems, in which at strong interaction limit, i.e. zero temperature limit, the effect of entropy vanishes. Although we simulated the system of colloids coated with mobile DNA linkers, the observed special effect of vibrational entropy can also explain the physics of forming floppy CsCl crystal of colloids coated with long flexible DNA linkers at experimental conditions, in which the binding sites of DNA linkers can cover the whole surface of DNA corona [75]. This opens up new possibilities for designing mDNACC systems to utilize the effect of entropy for fabrication of novel functional colloidal materials. For example, in asymmetric binary mixtures of mDNACCs, more open structures are expected, especially when nonspecific repulsions are introduced [65], of which the self-assembly mechanism can be very different from that of conventional patchy particle systems [14, 15, 78]. Moreover, in experimental systems of mobile DNA coated liposomes, the combination of vibrational entropy and deformability are expected leading to the formation of more interesting structures [79–82]. Additionally, the method developed for the infinite binding strength limit could also be modified to simulate other network systems, e.g. vitrimers [83, 84], in which the relaxation is not driven by energy change but a result of entropy maximization.

Chapter 3

Melting and re-entrant melting of polydisperse hard-disks

3.1 Introduction

As we mentioned in Chapter 1, the melting of two-dimensional solids has been heavily discussed since it was proven that long-wavelength thermal fluctuations prevents long-range positional order in 2D systems [85–87]. A significant attempt to reach a general description of 2D melting was the Kosterlitz-Thouless-Halperin-Nelson-Young (KTHNY) theory, that predicts a new phase of matter, i.e., the hexatic phase, which has quasi-long-range orientational order and short-range positional order. The hexatic phase is predicted to be interposed between the usual solid and fluid phases. Therefore, the melting in 2D was predicted to undergo two continuous transitions from solid to hexatic and hexatic to fluid, respectively [88–92]. However, the KTHNY theory does not rule out any first-order transition due to other effects [93], e.g., the grain-boundary induced melting [22, 94]. In arguably the simplest benchmarking model system in 2D, i.e., the system of monodisperse hard-disks, the physics of melting transition has been debated for long [27, 41, 95–101]. It was recently settled that the melting of solids in the system of monodisperse hard-disk undergoes a two-stage process consisting of a continuous solid-hexatic transition followed by a first-order hexatic-fluid transition [37, 38], and the shape and softness of particles also play important roles in the 2D melting [24, 26, 102].

It was found that pinning a small fraction of particles can change the melting transition in hard-disk systems significantly [29, 30]. Moreover, simulations of binary hard-disk mixtures showed that the presence of tiny amounts of small particles can eliminate the stability of hexatic phase [39]. These highlight that the melting transition in 2D is subtle. A recent experiment with colloidal hard spheres in 2D [42] appears to support the two-stage melting found in simulation [37, 38]. However, most of experimental systems possess certain degree of (continuous) particle size polydispersity, and polydisperse hard-disks have also been widely employed as a model system to investigate the glass transition [103–105]. Yet the effect of polydispersity on the nature of phase transitions in 2D remains unknown. To this end, we investigate the equilibrium phase behaviour of a 2D polydisperse hard-disk system (PHDS) with Gaussian-like particle size polydispersity. We find that with increasing the polydispersity, the first-order hexatic-fluid transition becomes weaker, and completely switches to be continuous following the KTHNY scenario in PHDS with around 7% size polydispersity. Furthermore, the packing fraction range for stable hexatic phase increases significantly by one order of magnitude compared to that in monodisperse hard-disk systems. More surprisingly, in PHDS with slightly higher polydispersity, we observe re-entrant solid-hexatic and hexatic-fluid transitions at high density, which were proven impossible in 3D polydisperse hard-sphere systems [106].

Finally, we find that this re-entrant melting is accompanied by a counter-intuitive decrease in packing fraction under pressurization and confirm that this behaviour is intrinsic to the Gaussian distribution used to describe the system’s polydispersity, so that it can be observed for any phase (fluid, hexatic or solid) and has no relation with the phase transition. This final result led us to study the effect of the particle size distribution shape in the phase diagram of polydisperse hard-disks, that we present in Chapter 4 of this work.

3.2 Model

To model the effect of polydispersity, we consider a 2D system of volume V containing N polydisperse hard-disks based on the semigrand canonical ensemble ($NVT - \Delta\mu$), in which the chemical potential difference between particles of different size is fixed [107–110]. In this work, we use the following function for the

FIGURE 3.1: **Phase behaviour of polydisperse hard-disks.**(a): Equation of state (EOS) for polydisperse hard-disk systems (PHDS) with various polydispersity parameter $\nu/\sigma_0 = 0.005$ to 0.08 in the representation of $(P - P^*)\sigma_0^2/k_B T$ vs $(\rho^{-1} - \rho_{hex}^{-1})\sigma_0^{-2}$, where P^* and ρ_{hex} are the pressure and density of the hexatic phase at the fluid-hexatic transition, respectively, and the solid lines are fits of the EOS using 5th order polynomials. (b): $|\langle\Psi_6\rangle|$ as functions of $(\rho^{-1} - \rho_{hex}^{-1})\sigma_0^{-2}$ for systems with $\nu/\sigma_0 = 0.07$ and 0.08 . P and $\rho = N/V$ are the pressure of the system and the density of particles in the system, respectively. (c): EOSs for PHDS of various numbers of particles $N = 64^2 \sim 512^2$ at $\nu/\sigma_0 = 0.07$. The solid symbols are the obtained fluid-hexatic transition points.

FIGURE 3.2: **Phase behaviour of polydisperse hard-disks.**(a): Phase diagram of the PHDS in the representation of $\rho\sigma_0^2$ and $s/\langle\sigma\rangle$, in which the dashed lines are the interpolated phase boundaries for re-entrant melting of solid and hexatic phases. The phase boundaries are obtained from $NVT - \Delta\mu$ MC simulations for PHDS with $\nu/\sigma_0 = 0.005$ to 0.0835 . Inset: the enlarged view of the region of phase diagram at $0 \leq s/\langle\sigma\rangle \leq 0.02$. The dotted lines connect the re-entrant transitions at $\nu/\sigma_0 = 0.08, 0.0805, 0.081, 0.082, 0.083$, and 0.0835 from left to right. and (b): same phase diagram but including the equations of state along the re-entrant transitions for $\nu/\sigma_0 = 0.08, 0.0805, 0.081, 0.082, 0.083$, and 0.0835 . Figure b) is mostly referenced in Chapter 4 of this work.

distribution of chemical potential difference

$$\Delta\mu(\sigma) = -k_B T \frac{(\sigma - \sigma_0)^2}{2\nu^2}, \quad (3.1)$$

where σ is the particle diameter changing from 0 to ∞ with k_B and T the Boltzmann constant and temperature of the system, respectively. ν is the polydispersity parameter, and in the ideal gas limit Eq. (3.1) gives a Gaussian-like particle size distribution centered around σ_0 with the standard deviation ν . This models PHDS in equilibrium with the dilute reservoir of Gaussian-like particle size distribution, e.g., the very top region in sedimentation experiments [42].

FIGURE 3.3: **Stabilization of hexatic phase.** (a): The fractions of topological defects and dislocations in the hexatic phase as functions of $(\rho - \rho_{hex})\sigma_0^2$ for systems of various polydispersity parameter ν . Here ρ and ρ_{hex} are the density of the system and the lowest density of stable hexatic phase, respectively. The vertical dashed lines indicate the hexatic-solid transition points. (b): Low pressure phase diagram of polydisperse hard-disks in the representation of η vs $s/\langle\sigma\rangle$, where η is the packing fraction of the system with σ_i the diameter of particle i . The state points obtained from simulations at each ν/σ_0 from 0.005 to 0.08 are shown as the symbols, which are color coded with $\langle\sigma\rangle/\sigma_0$. The error bars are smaller than the symbols.

3.3 Results

3.3.1 Phase diagram

By performing Monte Carlo (MC) simulations, we calculate the equation of state (EOS) for a system of $256^2 = 65,536$ hard-disks with various ν from 0.005 to $0.08\sigma_0$. As shown in Fig. 3.1a, one can see that when the polydispersity parameter is small, i.e., $\nu/\sigma_0 \leq 0.05$, there is clearly a Mayer-Wood loop in the EOS [111] implying a first-order transition from fluid with increasing the density of the system. Similar to the monodisperse hard-disk system, i.e., $\nu = 0$, the coexisting phase with fluid in PHDS is a hexatic phase [37, 38, 110]. As we explained in Chapter 1, to characterize the structural difference between fluid and hexatic phases, we calculate the six-fold bond orientation order parameter $\langle \Psi_6 \rangle = \left\langle \frac{1}{N} \sum_{k=1}^N \psi_6(\mathbf{r}_k) \right\rangle$ with $\psi_6(\mathbf{r}_k) = \frac{1}{N_k} \sum_{j=1}^{N_k} \exp(i6\theta_{kj})$, where θ_{kj} is the angle between the vector connecting particle k with its neighbour j and a chosen fixed reference vector. N_k is the number of first neighbours for particle k based on the Voronoi tessellation of the system. As shown in Fig. 3.2a, the density difference between the coexisting hexatic and fluid phases becomes smaller with increasing ν , and when $\nu/\sigma_0 \geq 0.07$, the density jump from fluid to hexatic phase disappears, while $|\langle \Psi_6 \rangle|$ increases significantly at ρ_{hex} (Fig. 3.1b). Here we note that $|\langle \Psi_6 \rangle|$ in an infinitely large system, should be zero for the hexatic phase, while in our large finite system it can be a positive number indicating the existence of some orientational order in the system. However, we do not use $|\langle \Psi_6 \rangle|$ to identify the existence of the hexatic phase, for which we always check the change of correlation functions to confirm. This suggests that the first-order fluid-hexatic transition becomes weaker with increasing ν , and changes to be continuous at high polydispersity, i.e., $\nu/\sigma_0 \geq 0.07$, following the celebrated KTHNY scenario [88–91]. This, to some extent, is similar to the melting of soft spheres [24]. Because with increasing the particle size polydispersity, the distribution of distance between the nearest neighbours becomes wider, which is similar to introducing a “soft” repulsion between particles, and it has been also found in the melting of soft spheres in 2D, the transition type can switch from a first order transition to a continuous KTHNY scenario with increasing the “softness” of the potential [24].

FIGURE 3.4: **Re-entrant melting transitions.** (a,b): The pair correlation function $g(x,0) - 1$ along the major axis of the system (a) and the six fold orientation correlation function $g_6(r)$ (b) for polydisperse hard-disk systems (PHDS) with $\nu/\sigma_0 = 0.08$ at various densities. (c): $|\langle\Psi_6\rangle|$ and $s/\langle\sigma\rangle$ as functions of density $\rho\sigma_0^2$ for PHDS with $\nu/\sigma_0 = 0.082, 0.083$, and 0.0835 . (d): Contour plot of the probability density distribution of particle size σ/σ_0 for systems of different densities at $\nu/\sigma_0 = 0.0835$. (e): Mean square displacement $\langle\Delta r^2\rangle/\sigma_0^2$ in PHDS of $\nu/\sigma_0 = 0.0835$ at densities $\rho\sigma_0^2 = 1.592, 2.292$, and 3.565 as marked in the inset, where $\tau = \sigma_0\sqrt{m/k_B T}$ is the time unit of molecular dynamics simulations with m the mass of particles. Inset: the diffusion coefficient D as a function of density in PHDS with $\nu/\sigma_0 = 0.0835$.

3.3.2 Finite size effect on the melting in polydisperse hard-disks

It was shown that the system finite size effect is important in studying 2D melting [37, 112]. Therefore, we perform MC simulations for PHDS of various numbers of particles from $N = 64^2$ to 512^2 at $\nu/\sigma_0 = 0.07$, and the EOSs for different system sizes are shown in Fig. 3.1c. When the system size is small, i.e., $N = 64^2 = 4096$, there is a pronounced Mayer-Wood loop in EOS, signaturing a first-order fluid-hexatic transition consistent with Ref [110]. Interestingly, with increasing the system size, the first-order fluid-hexatic transition becomes weaker, and changes to be continuous at $N \geq 256^2$. Further increasing N does not change EOS significantly, which suggests the finite size effect negligible in our system of $N = 256^2$ particles. The change from a first order like transition to a continuous transition with increasing system size can be interpreted as following. A finite 2D system with periodic boundary conditions can be seen as a system wrapped on a torus, and this increases the cooperation between neighbouring particles as a result of extra “communication” via paths encircling the torus, which was shown being able to change the transition type from continuous to first order like in small systems [113]. By contrast, this effect does not appear in systems with open free boundaries, while in real experiments, the open boundary walls can induce order in the fluid. Moreover, for $\nu/\sigma_0 < 0.07$, we ensure that most of the coexisting densities do not change significantly with further increasing the system size to $N = 512^2$, while the only exception is for systems at $\nu/\sigma_0 = 0.05$ (as shown in Figure 3.5), of which the exact boundary was out of the reach of our computation capability.

3.3.3 Stabilization of the hexatic phase

Our second finding for the phase diagram of polydisperse hard-disk is that polydispersity can significantly increase the single phase packing fraction stability range of the hexatic phase. This is determined by locating the hexatic-solid transition point. As we explained in Chapter 1, this transition can be identified by analysing the correlation function $g(x, 0) - 1$ along the major axis of system, that switches from an exponential decay to a power law decay with exponent $-1/3$ upon solidification [37].

FIGURE 3.5: The equation of state for polydisperse hard-disk systems (PHDS) of various system sizes with $\nu/\sigma_0 = 0.005, 0.01, 0.02$ and 0.05 . The solid lines are fits of the EOS using 5th order polynomials, and solid symbols are the obtained coexisting densities from the Maxwell construction.

Although for the monodisperse and low polydispersity systems ($\nu/\sigma_0 < 0.07$) we find a good agreement to the expected power law decay for the correlation function of the solid, as one can see from Fig. 3.4.a (top) that represents the function for several densities close to the calculated hexatic-solid transition for the ν/σ_0 system, in this case the decay is slightly faster than expected. We believe this is because of the effect that the particle size polydispersity has in the positional correlations, that introduces an error that is not considered in the general derivation for the KTHNY solid. The reason we believe this is so is because, as it can be seen from the figure, further compression of the system does not affect the decay of $g(x, 0) - 1$. To our knowledge, the only previous study of the effect of polydispersity in the phase diagram of 2D hard-disk was that by Pronk and Frenkel in 2004[110], but in that case no study of the $g(x, 0) - 1$ function decay was performed, as their focus was to determine whether polydispersity affected the hexatic-fluid transition. Therefore, we cannot confirm if polydispersity is playing any part in this result.

However, a recent paper by Li and Ciamarra[114] presented a new technique to

determine the orientation of the lattice symmetry axis used to project the $g(x, 0) - 1$ function, and cited our work [18] among other [115] as examples of possible misidentification of this axis, due to the faster than expected decay. During our data analysis we chose our projection axis to be the one that led to the lowest possible decay for this pair correlation function, so we find this to be unlikely. However, implementing this method to determine the hexatic-solid transition is a possible path for future development of our work.

Once we have determined all transition points, we build the phase diagram of our system, that is shown in the representation of $\rho\sigma_0^2$ vs $s/\langle\sigma\rangle$ vs in Fig. 3.2a, where $s = \sqrt{\langle\sigma^2\rangle - \langle\sigma\rangle^2}$ is the absolute particle size polydispersity in the system and is divided by the average particle as this is a more meaningful measure of the actual polydispersity of our system.

However, in experiments, one of the most relevant parameters is the packing fraction of the system. Thus we plot the low pressure phase diagram of PHDS in the representation of $\eta = \frac{\pi}{4}\langle\sum_i\sigma_i^2/V\rangle$ vs $s/\langle\sigma\rangle$ in Fig. 3.3b. Here the low pressure means the range of pressure close to the fluid→hexatic and hexatic→solid transitions below the re-entrant transitions. One can see that in systems of larger size polydispersity, the average particle size is smaller, and the actual packing fraction range for stable hexatic increases from $0.2 \sim 0.3\%$ in the monodisperse hard-disk system to about $2 \sim 3\%$ in the PHDS with $s/\langle\sigma\rangle \simeq 0.07$.

To understand the physics behind this enhanced stability of the hexatic phase in PHDS, we calculate the fraction of topological defects in the system using the method in Ref [26]. Particle k is identified as a topological defect if its disclination charge $q_a = N_k - 6 \neq 0$. These topological defects form clusters which can be classified based on the total disclination charge and Burgers vector [26]. Defect clusters having nonzero Burgers vectors and zero disclination charges are called dislocations, and the dislocations as well as the defect clusters with nonzero disclination charges can destroy the bond orientational order in 2D solids. Although the total amount of topological defects is not directly related to the stability of the hexatic phase, the dislocations are evidence for the hexatic phase along with few defect clusters with nonzero disclination charges [26]. The fraction of particles in defect clusters with nonzero disclination charge is much smaller, i.e., more than one order of magnitude smaller, than that of dislocations in the hexatic phase found

in our simulations, although the changing trends of these two types of defects are very similar (Figure 3.6).

The calculated fractions of topological defects and dislocations in our simulations with various ν/σ_0 are plotted as functions of $(\rho - \rho_{hex})\sigma_0^2$ in Fig. 3.3a, where ρ_{hex} is the lowest density of stable hexatic phase, and the vertical dashed lines locate hexatic-solid transitions. The representation as a function of ρ instead of η is done only to simplify the figure, as our only purpose is to show how the concentration of topological imperfections varies between the lowest density hexatic and solid phases as ν/σ_0 changes. In this way, one can see that typically when the fraction of dislocations decreases to below about 1 ~ 1.5%, the system solidifies. When the polydispersity is small, i.e., $\nu/\sigma_0 \leq 0.05$, the dependence of both fractions of defects and dislocations on density do not change significantly with increasing ν , and the density ranges for stable hexatic phase below the dashed lines in Fig. 3.3a are almost the same. However, at $\nu/\sigma_0 \geq 0.07$, the fraction of topological defects increases to around 10% at ρ_{hex} , and also the fraction of dislocations increases to about 4.5% at ρ_{hex} . Along with increasing the density at $\nu/\sigma_0 = 0.07$ and 0.08, the fraction of dislocations decreases to below the threshold for solidification at much higher density compared with the system of $\nu/\sigma_0 \leq 0.05$. The fraction of defect clusters with nonzero disclination charges also changes similarly in systems of different polydispersity (Figure 3.6). Therefore, this suggests that the size polydispersity of hard-disks creates more topological defects, which could subsequently increase the fraction of dislocations as well as the defect clusters with nonzero disclination charges in the system. This destroys the quasi-long-range positional correlation in the solid driving the formation of hexatic phase.

Finally, it can be observed that at fixed ν , increasing the density of the system along with the fluid-hexatic and hexatic-solid transitions decreases the actual polydispersity $s/\langle\sigma\rangle$ in the system, which is due to the formation of more ordered structures eliminating the particle size fluctuation in the system. Increasing the density also leads to a decrease in the average particle size $\langle\sigma\rangle$. Therefore, the packing fraction of the system η does not linearly increase with the density. As we will show in the next section, this leads to a surprising finding in the phase diagram of the system. However, this fractionation does not affect the qualitative behaviour of the particle size distribution, that remains very close to a Gaussian-like distribution centred around $\langle\sigma\rangle$ (Figure 3.7).

FIGURE 3.6: The fraction of particles in topological defects clusters with nonzero disclination charges in the hexatic phase as functions of $(\rho - \rho_{\text{hex}})\sigma_0^2$ for systems of various ν . The representation as a function of ρ rather than η was chosen to simplify the figure. This is however not important and the main purpose of Fig. The vertical dashed lines indicate the hexatic-crystal transition points.

FIGURE 3.7: Probability density distribution of particle size σ in polydisperse hard-disk systems (PHDS) with various ν and ρ , in which the solid lines are fitted Gaussian distribution.

3.3.4 Re-entrant melting

Another interesting feature of the phase diagram in Fig. 3.2a is that when the size polydispersity of hard-disks is around 0.08 to 0.1, at very high density, increasing density triggers re-entrant melting transitions from solid to hexatic and hexatic to fluid. As shown in Fig. 3.4a, for PHDS with $\nu/\sigma_0 = 0.08$, with increasing the density from $\rho\sigma_0^2 = 1.6$ to 1.6424, the system transforms from a hexatic phase with the short-range positional correlation, i.e., an exponential decay $g(x, 0) - 1 \sim \exp(-x)$, to a solid with quasi-long-range positional order, i.e., a power law decay $g(x, 0) - 1 \sim x^{-\alpha}$ with $\alpha \leq 1/3$, where the six fold orientation correlation function $g_6(r) = \langle \psi_6^*(\mathbf{r}' + \mathbf{r}) \psi_6(\mathbf{r}') \rangle$ [116] remains almost the same (Fig. 3.4b). At much higher density, increasing the density from $\rho\sigma_0^2 = 3.64$ to 3.67, $g(x, 0) - 1$ changes again from a power law decay to an exponential decay with almost the same $g_6(r)$ (Fig. 3.4b), which suggests a re-entrant transition from solid to hexatic phase. For hard-disk systems with higher polydispersity, i.e., $\nu/\sigma_0 = 0.082 \sim 0.0835$, we plot $|\langle \Psi_6 \rangle|$ as functions of $\rho\sigma_0^2$ in Fig. 3.4c, and one can see that with increasing ρ , $|\langle \Psi_6 \rangle|$ first increases from 0 to around 0.8 indicating the formation of an ordered phase, and further increasing ρ leads to the drop of $|\langle \Psi_6 \rangle|$ to 0 implying a re-entrant melting transition into a disordered phase. By checking the scaling of $g(x, 0) - 1$ in the system, we ensure that the ordered phase formed is a hexatic phase.

To understand the mechanism behind the re-entrant melting of the hexatic phase, we plot $s/\langle \sigma \rangle$ as functions of $\rho\sigma_0^2$ in the lower panel of Fig. 3.4c, from which one can find a clear correlation between the change of $|\langle \Psi_6 \rangle|$ and $s/\langle \sigma \rangle$. Here we ensure that our simulations at high density have been well equilibrated and independent with the initial configurations (Figure 3.8). Actually, it has been shown that the equilibration in systems of continuously polydisperse particles is much faster than that in monodisperse systems [117–119]. At high density, during the decrease of $|\langle \Psi_6 \rangle|$, $s/\langle \sigma \rangle$ increases significantly to the value even higher than that in the low density fluid.

This suggests that by increasing the density of the system, the standard deviation of particle size increases, and the combinatorial entropy in the system associated with the variation of particle size increases, which stabilizes the amorphous structures against the ordered ones in the system.

FIGURE 3.8: Six-fold bond orientation order parameter of the system $\langle \Psi_6 \rangle$ as functions of Monte Carlo (MC) sweeps in simulations from perfect crystal and random fluid at $\rho \sigma_0^2 = 1.528, 2.292$ and 3.565 with $\nu/\sigma_0 = 0.0835$. From these, one can see that the correlation time is around 10^7 MC sweeps, which is smaller than our simulation of around 10^8 MC sweeps, and the our simulations have been well equilibrated.

FIGURE 3.9: Equation of state for polydisperse hard-disk system with $\nu/\sigma_0 = 0.082, 0.083$ and 0.0835 .

Moreover, along with the decrease of $|\langle\Psi_6\rangle|$ with density at high pressure, we do not observe any Mayer-Wood loops in EOS (Figure 3.9). This implies that the re-entrant melting of hexatic phase at high density is continuous, which is similar to that in soft particle systems [120, 121]. Furthermore, it was shown that the strong particle size fractionation in polydisperse hard spheres eliminates the possibility of re-entrant melting in 3D polydisperse hard-sphere solids [106]. Therefore, in Fig. 3.4d, we plot the probability density distribution of particle size $p(\sigma)$ for PHDS of various density with $\nu/\sigma_0 = 0.0835$. One can see that within the whole density range, there is always a single peak in $p(\sigma)$ at fixed ρ , indicating no particle size fractionation in the system. This also suggests that re-entrant melting can indeed exist in 2D polydisperse particle systems, if there is no particle size fractionation [122]. Here we note that although all our simulations starting with different particle size distributions converge to the same result at fixed polydispersity and density, there still exist a possibility that the system may possess an equilibrium multimodal particle size distribution which is not accessible with direct simulations [122].

In addition, to check whether the amorphous phase formed at high density is a fluid or glass, we perform event driven molecular dynamics (EDMD) simulations starting from the equilibrated configurations obtained from our MC simulations at $\nu/\sigma_0 = 0.0835$. In the EDMD simulation, the system evolves via a time-ordered sequence of elastic collision events, which are described by Newton's equations of motion. The spheres move at constant velocities between collisions, and the velocities of the respective particles are updated when a collision occurs. All collisions are elastic and preserve energy and momentum [123]. The calculated mean square displacement (MSD) for various densities are shown in Fig. 3.4e. One can see that not only the diffusion coefficient increases by nearly two orders of magnitude along with the re-entrant melting of hexatic phase (Fig. 3.4e inset), but also the plateau in MSD almost disappears. This suggests that the high density amorphous phase is a diffusive fluid, and in the equilibrium sedimentation of such PHDS of Gaussian-like particle size distribution, there should be a "floating hexatic phase" sandwiched in two fluids. For PHDS with $\nu/\sigma_0 \geq 0.084$, with increasing the density, we do not observe any ordered phase in our simulations.

Finally, it is worth asking what is causing this increase in diffusivity after re-entrant melting. Although we originally believed that the increase in polydispersity was

FIGURE 3.10: $s/\langle\sigma\rangle$ (a) as well as η and $D\tau/\sigma_0^2$ (b) as functions of density $\rho\sigma_0^2$ for a system with $\nu/\sigma_0 = 0.0835$.

FIGURE 3.11: Packing fraction η as a function of density $\rho\sigma_0^2$ for Gaussian-like polydisperse hard-disks with $\nu/\sigma_0 = 0.1$.

somehow helping the movement of the particles, the answer is in fact much more simple. As it can be seen from Fig. 3.10, that represents η as a function of density $\rho\sigma_0^2$ along with the diffusion coefficient D for the $\nu/\sigma_0 = 0.0835$ system, the increase in diffusivity due to the re-entrant melting is accompanied by a decrease in the system's packing fraction, that occurs because of a decrease in the average particle size. Figure A.1, in the Appendix A of this work, shows the equations of state in the η VS $\rho\sigma_0^2$ for the $\nu/\sigma_0 = 0.082$ and 0.083 system, which follow a similar trend. This decrease in packing fraction explains the higher mobility of the particles, but supposes a very surprising result. But even more interestingly, even in the hexatic phase, the packing fraction η does not increase monotonically with increasing density, but instead reaches a maximum value at about $\rho\sigma_0^2 \simeq 2.7$.

Therefore, to check whether the non-monotonic dependence of η on ρ is a general feature for Gaussian-like polydisperse hard-disks at large ν or if it is related with the phase transition, we perform simulations for Gaussian-like polydisperse hard-disks with $\nu/\sigma_0 = 0.1$, in which, according to the phase diagram of the system (Fig. 3.2.a), there is no phase transition with increasing ρ , and the obtained η as a function of $\rho\sigma_0^2$ is shown in Fig. 3.11. One can see that with increasing ρ , the packing fraction of the Gaussian-like polydisperse hard-disks first increases and then decreases with a maximum at around $\rho\sigma_0^2 \simeq 3$. This suggests that the non-monotonic behaviour of packing fraction is indeed a general feature of highly polydisperse Gaussian-like hard-disks, including the polydispersity range where the re-entrant melting transition occurs.

3.3.5 Non monotonic packing fraction behaviour for the ideal gas phase

Our study of the counter-intuitive non monotonic dependence between η and ρ for the Gaussian polydisperse hard-disk system for high polydispersity parameters led us to find a simple ideal gas system that qualitatively reproduces this behaviour. To understand the derivation of this model, we first look at Eq. 3.1 that, as we explained, gives us the chemical potential cost of exchange one of the system's disk with a different sized one. However, in the $NVT - \Delta\mu$ simulations, the probability of this exchange occurring will not be equal to $\exp[(\Delta\mu(\sigma_{\text{old}}) - \Delta\mu(\sigma_{\text{new}}))/k_B T]$, because many exchanges to larger disks will be rejected because of particle overlap. Therefore, there is an excess term in the chemical potential that determines the probability of a certain exchange to occur:

$$\Delta\mu^*(\sigma) = \Delta\mu_{\text{Gauss}}(\sigma) + \Delta\mu_{\text{excess}}(\sigma) \quad (3.2)$$

where $\Delta\mu_{\text{Gauss}}(\sigma)$ is the chemical potential determined by Eq. 3.1 and $\Delta\mu_{\text{excess}}(\sigma)$ is the excess chemical potential determined by the probability of an overlap occurring because of the insertion of the disk. This excess chemical potential can be familiar to the reader, as it is the one calculated by the Widom insertion method. This way, the probability of a certain particle exchange occurring during the $NVT - \Delta\mu$

FIGURE 3.12: a) Average particle size ($\langle \sigma \rangle / \sigma_0$) and b) absolute polydispersity (s / σ_0) VS density ($\rho \sigma_0^2$) for a system of Gaussian PDHD with $\nu / \sigma_0 = 0.1$ (black) and for an ideal gas configuration of the Gaussian polydisperse hard-disk system under equal pressure (red).

simulation is actually:

$$p^*(\sigma_{\text{old}} \rightarrow \sigma_{\text{new}}) = \exp \frac{\Delta\mu^*(\sigma_{\text{new}}) - \Delta\mu^*(\sigma_{\text{old}})}{k_B T} \quad (3.3)$$

For the case of an ideal gas, $\Delta\mu_{\text{excess}}(\sigma)$ has a simple expression: it is equal to the work required to deform the gas so that we create an empty region of area $\pi\sigma^2/4$ in which the hard-disk will be inserted. Therefore, assuming that the pressure of the ideal gas will not be considerably affected during this small deformation, the excess chemical potential is simply $\Delta\mu_{\text{excess}}(\sigma) = P\pi\sigma^2/4$, where P is the pressure of the system. Now, if we introduce this expression in equation 3.3 and find the distribution determined by this new chemical potential, we find it to be another Gaussian distribution, but with new center and width determined by the center (σ_0) and standard deviation (ν) of the previous distribution and the system's pressure (P):

$$p_{\text{i.g.}}(\sigma) = \exp \left[-(\sigma - \sigma_{0,\text{i.g.}})^2 / 2\nu_{\text{i.g.}}^2 \right] \quad (3.4)$$

where:

$$\begin{aligned} \sigma_{0,\text{i.g.}} &= \frac{\sigma_0}{(1 + 2\nu^2 P / k_B T)} \\ \nu_{\text{i.g.}} &= \frac{\nu}{\sqrt{1 + 2\nu^2 P / k_B T}} \end{aligned} \quad (3.5)$$

These are respectively the center of the new distribution ($\sigma_{0,\text{i.g.}}$) and its standard deviation ($\nu_{\text{i.g.}}$).

Therefore, $p_{\text{i.g.}}$ will be the probability size distribution for an ideal gas configuration of the Gaussian polydisperse hard-disk system in the $NVT - \Delta\mu$ ensemble. We should expect this distribution to describe correctly low packing fraction Gaussian polydisperse hard-disk configurations, for which $\Delta\mu_{\text{excess}}(\sigma)$ will be approximately that of an ideal gas. To test this, we tried comparing the average particle size ($\langle\sigma\rangle$) and the absolute polydispersity ($s = \sqrt{\langle\sigma^2\rangle - \langle\sigma\rangle^2}$) for the $\nu/\sigma_0 = 0.1$ system with those predicted by Eq. 3.5 using the equilibrium pressures of its equation of state. The results can be seen in Fig. 3.12 and show a good agreement for both the average particle size and absolute polydispersity. Furthermore, in Fig. 3.12.b we can see that the fit is almost perfect for very low and very high densities. This is because, as we saw in Fig. 3.11, for both of these regions the system presents low packing fraction.

We can now determine whether this ideal gas particle size distribution correctly predicts the counter-intuitive behaviour of the packing fraction at high pressures. Using Eq. 3.4 we find the packing fraction of a system of 2D hard-disks following an ideal gas expression to be:

$$\eta_{\text{i.g.}} = \rho \frac{\pi \langle\sigma^2\rangle}{4} = \rho \frac{\pi}{4} (\sigma_{0,\text{i.g.}}^2 + \nu_{\text{i.g.}}^2) \quad (3.6)$$

where ρ is the system density. Since for all conditions considered in this chapter the average disk diameter of the system is always significantly larger than the absolute polydispersity, we can simplify Eq. 3.6 by not considering the $\nu_{\text{i.g.}}^2$ contribution. Then, by using the equation of state of an ideal gas ($\rho = P/k_B T$) and by substituting the expression for the average particle size of the ideal gas in Eq. 3.5, we can express the packing fraction of an ideal gas configuration of the Gaussian polydisperse hard-disk system in the $NVT - \Delta\mu$ ensemble as:

$$\eta_{\text{i.g.}}(P) = \frac{\pi \sigma_0^2 P k_B T}{4 (1 + 2\nu^2 P/k_B T)^2} \quad (3.7)$$

and the derivative of the packing fraction with respect pressure is equal to:

$$\frac{d\eta_{\text{i.g.}}}{dP} = \frac{\pi \sigma_0^2}{4 k_B T} \frac{1 - (2\nu^2/k_B T)^2 P^2}{(1 + 2\nu^2 P/k_B T)^4} \quad (3.8)$$

FIGURE 3.13: a) Packing fraction (η) VS density ($\rho\sigma_0^2$) for a system of Gaussian polydisperse hard-disks with $\nu/\sigma_0 = 0.1$ (black) and for an ideal gas configuration of the Gaussian polydisperse hard-disk system under equal pressure (red).

This result tells us that the equation of state will present a maximum in packing fraction for pressure $P_{\max} = \frac{k_B T}{2\nu^2}$ and, therefore, also predicts that the pressure for this maximum will be lower the higher the value of the polydispersity parameter, just as we should expect given our results. This result further confirms that the non monotonic dependence between η and ρ is a general feature for Gaussian-like polydisperse hard-disks, as this behaviour is replicated even for an ideal gas system.

Of course, the equation of state of the $\nu/\sigma_0 = 0.1$ system does not follow the ideal gas expression, due to the interactions between hard-disks. Still, we can again compare our simulation results with those given by Eq. 3.6 using the equilibrium pressures of the system at each point of the equation of state. The results are shown in Fig. 3.13. It can be seen that although the fit is not perfect, both equations of state behave qualitatively the same and they both agree on the location of the maximum in packing fraction.

We thank Dr. Lei Qun-li, who first proposed using an ideal gas expression to determine the excess chemical potential associated to particle exchange during our $NVT - \Delta\mu$ simulations, to approximate the real system probability size distribution.

3.4 Conclusions

By performing large scale computer simulations, we investigate the effect of particle size polydispersity on the phase behaviour of hard-disk systems. We find that with increasing the size polydispersity of hard-disks from zero, the first-order melting transition from hexatic phase to fluid becomes weaker, and completely switches to be continuous at $\nu/\sigma_0 \geq 0.07$ following the celebrated KTHNY scenario. Simultaneously, the density range for stable hexatic phase gets substantially enlarged. Compared with monodisperse hard-disk systems, where hexatic phase is only stable in a very small range of packing fraction of $0.2 \sim 0.3\%$ [37], the packing fraction range for stable hexatic phase in PHDS with $s/\langle\sigma\rangle \simeq 0.07$ is about $2 \sim 3\%$. This suggests new directions in searching for the hexatic phase in 2D polydisperse particle systems. More interestingly, in PHDS with even higher polydispersity, i.e., $0.08 \leq \nu/\sigma_0 \lesssim 0.0835$, we find that at very high density, increasing density of the system can trigger re-entrant transitions from solid to hexatic and hexatic to fluid phases depending on ν . In polydisperse systems, re-entrant transitions, especially the re-entrant melting, were originally predicted theoretically for 3D hard-sphere crystals [122], but proven impossible due to the strong fractionation that was not taken into account in the theory [106]. We find that the absence of strong fractionation in 2D polydisperse systems can indeed induce re-entrant transitions at high density. These show a new difference between phase transitions in polydisperse hard-sphere systems in 2D and 3D, and suggest a new direction in investigating phase transitions in 2D systems by considering the particle size polydispersity, which was overlooked in the past.

Finally, this re-entrant melting occurs along with a counter-intuitive decrease in the system packing fraction with compression, that is determined to be a general feature of the chosen Gaussian distribution for the $NVT - \Delta\mu$ calculations, and not because of the phase transition. In Chapter 4 we will explore the consequences of this finding and present our study of the effect that the probability distribution shape of polydisperse hard-disks has on their phase diagram.

Chapter 4

Effect of particle size distribution in polydisperse hard-disks

4.1 Introduction

As we showed in Chapter 3, particle size polydispersity can change the nature of melting transition for continuously polydisperse hard-disks. Namely, when increasing the degree of polydispersity, the first-order hexatic-fluid transition becomes weaker and eventually disappears at relatively high polydispersity, and simultaneously the density range of stable hexatic phase increases by orders of magnitude. More intriguingly, for systems of relatively highly polydispersed hard-disks, a re-entrant melting transition was observed with increasing the density of system, which was proven impossible in 3D system of polydisperse hard spheres [106, 122].

In the work we presented in Chapter 3, we assumed that the polydisperse hard-disk system is in contact with a dilute reservoir of Gaussian-like particle size distribution. To model the effect of polydispersity, we considered a 2D system of volume V containing N polydisperse hard-disks based on the semigrand canonical ensemble, in which the chemical potential difference between particles of different size is fixed [107–110], and its distribution obeys

$$\frac{\Delta\mu(\sigma)}{k_B T} = -\frac{(\sigma - \sigma_0)^2}{2\nu^2} \quad (4.1)$$

where σ is the particle diameter changing from 0 to ∞ , with k_B and T the Boltzmann constant and temperature of the system, respectively. ν is the polydispersity parameter and the main controlling parameter of the system. ν is the polydispersity parameter, and in the ideal gas limit Eq. 4.1 gives a Gaussian-like particle size distribution centered around σ_0 with the standard deviation ν . This same equation appears in Chapter 3 as Eq. 3.1, but we repeat it in this chapter for convenience.

As we showed in Chapter 3, the re-entrant melting found for the Gaussian polydisperse hard disk system is associated with a counter-intuitive decrease in the system packing fraction under compression. Furthermore, it was shown that this behaviour was related to the behaviour of the Gaussian distribution itself and not a result of the phase transition. Analysing Eq. 4.1, one finds that its most notable artefact is that there is a finite probability of having $\sigma = 0$, which is unrealistic and has pronounced effects at large ν , where the re-entrant melting was observed. This opens the question of whether the reported effects of particle size polydispersity in the phase diagram of hard-disk is general or only exists in the system of artificial Gaussian-like polydisperse hard-disks. To clarify this point, in this chapter we investigate the effect of particle size distribution on the phase behavior of polydisperse hard-disks, and we simulate polydisperse hard-disk systems of two representative particle size distribution, i.e., the lognormal and triangle distributions, in which the probability of having $\sigma = 0$ is exactly zero. The main difference between these two new distributions is that in the triangle distribution, there is finite range of σ having probability zero, while in the lognormal distribution, the particle size probability only vanishes at $\sigma = 0$.

4.2 Lognormal polydisperse hard-disks

As we said in the introduction, the major artefact of Eq. 4.1 is that the probability of having $\sigma = 0$ is a finite positive number and therefore, to avoid this, we need to choose a probability distribution with zero probability of finding a zero sized particle.

FIGURE 4.1: (a) EOS for lognormal polydisperse hard-disk systems with various polydispersity parameter $\nu/\sigma_0 = 0.05$ to 0.15 in the representation of $(P - P^*)\sigma_0^2/k_B T$ vs $(\rho^{-1} - \rho_{hex}^{-1})\sigma_0^{-2}$, where P^* and ρ_{hex} are the pressure and density of the hexatic phase at the fluid-hexatic transition, respectively, and the solid lines are fits of the EOS using 5th order polynomials. (b) $|\langle\Psi_6\rangle|$ as functions of $(\rho^{-1} - \rho_{hex}^{-1})\sigma_0^{-2}$ for systems with $\nu/\sigma_0 = 0.10, 0.13$ and 0.15 .

First, we introduce the lognormal distribution, in which $\log(\sigma/\sigma_0)$ obeys a Gaussian distribution. Therefore, the chemical potential between particles of different size is

$$\frac{\Delta\mu(\sigma)}{k_B T} = -\frac{\log^2(\sigma/\sigma_0)}{2\nu^2} - \log\left(\frac{\sigma}{\sigma_0}\right) \quad (4.2)$$

where ν is the polydispersity parameter, and in the ideal gas limit Eq. 4.2 gives a system of polydisperse particles centered around $\sigma = \sigma_0$ with the standard deviation ν .

We perform $NVT - \Delta\mu$ simulation using the EMMC algorithm with $N = 256^2$ particles based on Eq. 4.2, and the calculated EOS for systems of various polydispersity parameter ν is shown in Fig. 4.1a. One can see that similar to the situation in Gaussian-like polydisperse hard-disks, with increasing ν , the Mayer-Wood loop in EOS becomes smaller implying that first-order transition becomes weaker [111]. For $\nu\sigma_0 \leq 0.1$, the Mayer-Wood loop in EOS completely disappears, while there is clearly a transition from a disordered state to an ordered state indicated by the sharp increase of $|\langle\Psi_6\rangle|$ with increasing density as shown in Fig. 4.1b. By checking the decay of $g(x, 0) - 1$ and $g_6(r)$, we ensure that the ordered phase forming from the fluid is the hexatic phase, which is qualitatively the same as in the system of Gaussian-like polydisperse hard-disks, as shown in Chapter 3.

FIGURE 4.2: (a) Phase diagram of the polydisperse hard-disk systems of lognormal distribution in the representation of $\rho\sigma_0^2$ and $s/\langle\sigma\rangle$. The phase boundaries are obtained from $NVT - \Delta\mu$ simulations for systems with $\nu/\sigma_0 = 0.005$ to 0.2 . Inset: the enlarged view of the region of phase diagram at $0.06 \leq s/\langle\sigma\rangle \leq 0.08$. The crosses are the simulation results along the the re-entrant transitions at $\nu/\sigma_0 = 0.13, 0.15, 0.17$, and 0.2 from bottom to top. (b) Low pressure phase diagram of polydisperse hard-disks with lognormal distribution in the representation of η vs $s/\langle\sigma\rangle$, where η is the packing fraction of the system. The state points obtained from simulations at each ν/σ_0 from 0.08 to 0.20 are shown as the symbols. The error bars are smaller than the symbols

FIGURE 4.3: Chemical potential difference $\Delta\mu(\sigma)/k_B T$ for Gaussian-like (Eq. 4.1) and lognormal polydisperse (Eq. 4.2) hard-disks with $\nu/\sigma_0 = 0.1$.

The calculated phase diagram for polydisperse hard-disks systems of lognormal distribution is shown in Fig. 4.2a in the presentation of $\rho\sigma_0^2$ and $s/\langle\sigma\rangle$. One can see that qualitatively, the phase diagram is very similar to the system of Gaussian-like polydisperse hard-disks in Fig. 3.2.b, in which with increasing polydispersity, the first-order fluid-hexatic transition becomes weaker and eventually changes to be a continuous transition, and the density range for stable hexatic phase increases by one order of magnitude. Moreover, as shown in Fig. 4.2a, at $s/\langle\sigma\rangle \gtrsim 0.09$, with increasing ρ , the system can change from fluid to hexatic and to fluid again at very high density, which is seemingly qualitatively the same as in the system of Gaussian-like polydisperse hard-disks (Fig. 3.2.b). However, as shown in Fig. 4.2a, one can see that at fixed large ν/σ_0 , e.g. $\nu/\sigma_0 \gtrsim 0.13$, with increasing the density of the system the polydispersity $s/\langle\sigma\rangle$ decreases, which is in contrast with the situation in the system of Gaussian-like polydisperse hard-disks, where $s/\langle\sigma\rangle$ increases with increasing density at large ν/σ_0 as shown in Fig. 3.2.b. This implies that for the sedimentation of highly polydisperse hard-disks with lognormal distributions, no re-entrant melting of hexatic phase can occur. This can be understood from the different behavior of the chemical potential in the two distributions. In Fig. 4.3, we plot Eq. 4.1 and 4.2 for $\nu/\sigma_0 = 0.1$. One can see that for Gaussian-like polydisperse system, when σ is very small, e.g. $\sigma/\sigma_0 < 10^{-2}$, further decreasing σ does not cost much free energy, which essentially drive the system to decrease the particle size at higher pressure to form a random fluid with increasing the true polydispersity of the system $s/\langle\sigma\rangle$. On the contrary, for the polydisperse hard-disks with lognormal

FIGURE 4.4: EOS for triangle polydisperse hard-disk systems with various polydispersity parameter $\nu/\sigma_0 = 0.25$ to 0.3 in the representation of $(P - P^*)\sigma_0^2/k_B T$ vs $(\rho^{-1} - \rho_{hex}^{-1})\sigma_0^{-2}$, where P^* and ρ_{hex} are the pressure and density of the hexatic phase at the fluid-hexatic transition, respectively, and the solid lines are fits of the EOS using 5th order polynomials.

distributions, the free energy cost for decreasing the the particle size increases dramatically and diverges when approach $\sigma = 0$, this effectively prevents the decrease of particle size as well as the increase in the true polydispersity of the system $s/\langle\sigma\rangle$ at large pressure. Therefore, at very high pressure, polydisperse hard-disks with lognormal distributions are almost “monodisperse” and forming an ordered solid phase.

However, according to the phase diagram in Fig. 4.2a, if one keeps the polydispersity of the system fixed at a certain value above $s/\langle\sigma\rangle \simeq 0.09$, increasing the density, or pressure, of the system, the system indeed can transform from a fluid phase to a hexatic phase and then to a fluid phase again at very high density or pressure.

4.3 Triangular polydisperse hard-disks

In Sec. 4.2, we studied the phase behavior of polydisperse hard-disks with lognormal distributions, where the particle size probability vanishes at $\sigma = 0$. Here we investigate another different type of polydisperse system, i.e., the triangle distribution, in which the particle size probability with a finite range above $\sigma = 0$ is zero.

FIGURE 4.5: Phase diagram for the systems of triangular polydisperse hard-disks in the representations of $\rho\sigma_0^2$ vs $s/\langle\sigma\rangle$ (a) and η vs $s/\langle\sigma\rangle$ for ν/σ_0 from 0.02 to 0.3. The inset in (a) is the enlarged view of the low polydispersity region.

FIGURE 4.6: Polydispersity $s/\langle\sigma\rangle$ as a function of packing fraction η for highly polydisperse hard-disks with triangle distributions at $0.08 \leq \nu/\sigma \leq 0.3$.

In this case, the chemical potential between particles of different size is:

$$\frac{\Delta\mu(\sigma)}{k_B T} = \begin{cases} -\ln\left(\frac{\sqrt{6}\nu - |\sigma - \sigma_0|}{\sqrt{6}\nu}\right) & \text{if } |\sigma - \sigma_0| < \sqrt{6}\nu, \\ -\infty & \text{else,} \end{cases} \quad (4.3)$$

where ν is the polydispersity parameter. This distribution is characterized for having a maximum at $\sigma = \sigma_0$ and linearly decays as it moves away from the maximum until reaching zero probability for a distance $\sqrt{6}\nu$. In the ideal gas limit Eq. 4.3 gives a system of polydisperse particles centered around $\sigma = \sigma_0$ with the standard deviation ν .

First, we perform $NVT - \Delta\mu$ ECMC simulations in a 256^2 particle system to obtain the equation of state for values of ν ranging between 0.01 and $0.30\sigma_0$ as shown in Fig. 4.4. One can see that with increasing ν , the first-order transition from the fluid phase does not become weaker, and even at $\nu/\sigma_0 = 0.3$, there is still a pronounced Mayer-Wood loop in the EOS. By checking the decay of $g(x, 0) - 1$ and $g_6(r)$, we ensure that the ordered phase forming from the fluid is a hexatic phase, and this persistent first-order fluid-hexatic phase transition in triangle polydisperse hard-disks is markedly different from the polydisperse hard-disk systems of Gaussian-like and lognormal polydisperse distributions.

The phase diagrams of polydisperse hard-disks with various triangle distributions are summarized in Fig. 4.5. One can see that the density range for stable hexatic phase does not change significantly with increasing $\nu/\sigma_0 = 0$ to 0.08, and the corresponding packing fraction range η for stable hexatic phase increases from 0.002 at $\nu/\sigma_0 = 0$ to about 0.005 at $\nu/\sigma_0 = 0.3$ (Fig. 4.5b). Moreover, the re-entrant melting transition of hexatic phase does not exist in the range of ν/σ_0 studied from $\nu/\sigma_0 = 0.02$ to 0.3. At high polydispersity, e.g. $\nu/\sigma_0 \gtrsim 0.13$, as shown in Fig. 4.5b, the phase boundary between fluid and hexatic phase converges to $(\eta_{fluid} \simeq 0.72, s_{fluid}/\langle\sigma\rangle \simeq 0.057)$ and $(\eta_{hex} \simeq 0.728, s_{hex}/\langle\sigma\rangle \simeq 0.055)$, and hexatic to solid transition point converges to $(\eta_{solid} \simeq 0.731, s_{solid}/\langle\sigma\rangle \simeq 0.054)$. In Fig. 4.6, we plot the true polydispersity of the system $s/\langle\sigma\rangle$ as a function of packing fraction η for triangle polydisperse hard-disks with various ν/σ_0 , and one can see that for $\nu/\sigma_0 \geq 0.1$, all curves collapse into a single curve, suggesting that at high enough ν/σ_0 , the properties of the system essentially do not depend on the polydispersity parameter ν . To understand this, we define a normalized particle

size

$$\sigma^* = \frac{\sigma}{\sigma_{\min}}, \quad (4.4)$$

where $\sigma_{\min} = \sigma_0 - \sqrt{6}\nu$ is the minimal particle size for the triangle distribution of ν . Then we can re-write Eq. 4.3 into

$$\frac{\Delta\mu(\sigma^*)}{k_B T} = \begin{cases} -\ln\left(\frac{\sqrt{6}\nu - |\sigma_0 - \sigma^* \sigma_{\min}|}{\sqrt{6}\nu}\right) & \text{if } \left|\sigma^* - \frac{\sigma_0}{\sigma_{\min}}\right| < \frac{\sqrt{6}\nu}{\sigma_{\min}}, \\ -\infty & \text{else.} \end{cases} \quad (4.5)$$

At high pressure, when σ is approaching $\sigma_{\min}$, i.e., $\sigma^* \simeq 1$, we have

$$\frac{d\Delta\mu}{d\sigma^*} = \frac{k_B T}{1 - \sigma^*}, \quad (4.6)$$

which implies that at high pressure, the change of $\Delta\mu$ does not depend on polydispersity parameter ν . This suggests for high enough pressures, when the average particle size is close to $\sigma_{\min}$, i.e., $\langle\sigma^*\rangle \simeq 1$, all systems are equivalent, independently of their ν value. Given that in systems of highly polydisperse hard-disks, the distribution of particle size is wider, the pressure or packing fraction required to push $\langle\sigma^*\rangle$ towards 1 is smaller, and this explains the collapse of curves for different ν in Fig. 4.6 at high polydispersity and high packing fraction.

4.4 Conclusions

Following our previous work on the phase transitions in Gaussian-like polydisperse hard-disks detailed in Chapter 3, in this chapter we have investigated the effect of particle size distribution on the phase behaviour of polydisperse hard-disks. Our main results in Chapter 3 were: (i) with increasing the polydispersity of the system, the hexatic-fluid transition becomes weaker and completely switches to a continuous transition following the celebrated KTHNY scenario at high polydispersity; (ii) simultaneously, the stable density range of hexatic phase is enlarged by orders of magnitude; (iii) at high polydispersity parameter ν re-entrant melting transitions of ordered phases, i.e., solid and hexatic phases, are observed with increasing pressure, which suggests that the re-entrant melting transition can be found in sedimentation experiments [42]. However, in the Gaussian-like polydisperse distribution, the probability of having a zero sized particle is finite, which is the artefact of the model. Therefore, this motivated us to investigate the phase behaviour of

polydisperse hard-disks with two other representative types of particle size distributions for which the probability of finding a particle with zero size vanishes to zero, i.e., the lognormal and the triangle distributions.

In systems of polydisperse hard-disks with lognormal distributions, the phase diagram appears qualitatively the same as the Gaussian-like polydisperse hard-disks, and the enhanced stability of hexatic as well as the switch of hexatic-fluid transition to the KTHNY scenario are both found in the system of polydisperse hard-disks with lognormal distribution with increasing polydispersity. Moreover, at a fixed true polydispersity of the system, $s/\langle\sigma\rangle$, with increasing density, re-entrant melting of solid and hexatic phases can be also found in highly polydisperse lognormal systems of hard-disks. However, if at any polydispersity parameter ν , with increasing pressure or density, the true polydispersity of the system monotonically decreases, and no re-entrant melting transition is found. The reason is that at very high pressure, the free energy cost of decreasing the particle size is diverging at zero for lognormal distributions, while it is almost zero for Gaussian-like distributions. This suggests that different from the Gaussian-like polydisperse hard-disks, in the sedimentation of polydisperse hard-disks of lognormal distributions, we would not find any re-entrant melting transitions of ordered phases.

Moreover, in systems of polydisperse hard-disks with triangle distributions, the phase diagram is qualitatively different from those of Gaussian-like and lognormal distributions. For all polydispersity parameter studied, the hexatic-fluid transition in polydisperse hard-disks with triangle distributions is always strongly first order, and we do not observe significant increase of stable density range of hexatic phase with increasing polydispersity of the system. Moreover, we can not reach any system of true polydispersity higher than 0.06. The reason is due to the fact that at high pressure and high polydispersity, the change of chemical potential of different particle size does not depend on the polydispersity of the system. Our results show that the exact particle size distribution plays an important role in the phase behavior of polydisperse hard-disks, and so far we can not reach a universal phase diagram for all polydisperse hard-disks. However, the effect in the stabilization of the hexatic phase and the return to the KTHNY theory melting scenario exposed in Chapter 3 is not affected by distribution shape.

Chapter 5

Adaptive bias Boltzmann generators

Although the whole field of Machine Learning is currently living an explosion in interest and applications, few approaches are as popular as artificial neural networks. Far from a recent development, neural networks were proposed more than half a century ago[124, 125] and then abandoned for several decades. Nowadays, however, they are experiencing a rebirth caused by a combination of their success in image recognition technology[126]; the development of the back propagation algorithm [127, 128] that significantly speeds up their training; and the general increase in computation speed. This way, neural networks have been implemented successfully for a vast range of applications[129, 130].

With respect to applications to physics, neural networks have shown potential in learning order parameters to classify phases[131, 132], learning complex potential energy surfaces[133, 134] and even in the treatment of many-body quantum systems[135]. However, most applications are based on supervised learning, for which one needs a large amount of classified data that the neural network uses to determine the main features that relate them. There are, however some proposed unsupervised learning approaches to sample the Boltzmann distribution of a system[136, 137]. In this fashion, Noé *et al* recently developed a new algorithm called Boltzmann generator[138], that uses trained reversible neural networks to transform configurations of the system expressed into the real space to others expressed in an alternative latent space. This approach, when combined with energy

based training of the involved transformation functions, allowed Noé *et al* to sample the conformational space of several problems in a fraction of the time required for its alternatives. However, although in theory the method can be used with no prior knowledge about the regions of the phase space that need to be sampled, in practice training by example is required correct results.

Considering this, we propose a new method that combines Boltzmann generators with adaptive bias sampling techniques, such as MetaDynamics[139]. The key to the method lies in using the Boltzmann distribution generated by a Boltzmann generator to gradually fill the potential energy surface one wants to sample, instead of using arbitrary repulsive functions based on order parameters. Therefore, in this chapter we present the method and show our results for its application to sample a simple two dimensional four-well potential energy surface.

5.1 Boltzmann generators

The main idea behind Boltzmann generators lies in the training of a set of transformations to connect a complex Boltzmann distribution $p(\mathbf{x})$ expressed in real space, to another simple distribution expressed in an equally sized latent space $p(\mathbf{z})$. This way, each point in latent space is univocally connected to another one in real space, and probability $p(\mathbf{x})$ for a configuration in real space can be related to the probability of its associated latent space configuration $p(\mathbf{z})$.

Since the transformation is univocal, we are limited to use functions that are invertible, so that $(F_{xz}(\mathbf{x}))^{-1} = F_{zx}(\mathbf{z})$. For this reason, Noé *et al* proposed to use invertible neural networks, and more specifically the architecture proposed by Dinh and colleagues[140]. In this case, the first step of the transformation involves dividing the full set of coordinates we want to transform into halves $\mathbf{u}_1$ and $\mathbf{u}_2$. Then, the transformation $F_{uv}(\mathbf{u})$ is expressed as:

$$\begin{aligned}\mathbf{v}_1 &= \mathbf{u}_1 \odot \exp(s_2(\mathbf{u}_2)) + t_2(\mathbf{u}_2) \\ \mathbf{v}_2 &= \mathbf{u}_2 \odot \exp(s_1(\mathbf{v}_1)) + t_1(\mathbf{v}_1)\end{aligned}\tag{5.1}$$

where $\odot$ represents the element wise product. It is easy to see that the output $v = [v_1, v_2]$ can be transformed back to u easily:

$$\begin{aligned}\mathbf{u}_2 &= (\mathbf{v}_2 - t_1(\mathbf{v}_1)) \odot \exp(-s_1(\mathbf{v}_1)) \\ \mathbf{u}_1 &= (\mathbf{v}_1 - t_2(\mathbf{u}_2)) \odot \exp(-s_2(\mathbf{u}_2))\end{aligned}\tag{5.2}$$

so that $(F_{uv})^{-1} = F_{vu}$. This transformation is invertible even if the s_1 , s_2 , t_1 and t_2 functions are not. Therefore, invertible neural networks use neural networks for these functions, so that they can be trained to optimize a certain cost function. This implementation of invertible neural networks is called a RealNVP layer and they are the architecture used for building Boltzmann Generators. Therefore, each invertible neural network determines a layer of the Boltzmann generator $F_{xz,i}$, with a set of trainable parameters θ_i . In general, applications of Boltzmann generators require to concatenate several of these transformations, such that $F_{xz} = [F_{xz,i}]$ and $\theta = [\theta_i]$.

As mentioned, the real potential of using an invertible transformation between the real and the latent space is that we can relate the probabilities in latent space $p(\mathbf{z})$ to those for the real system equilibrium distribution $p(\mathbf{x})$. In order to explain how this is done, let us first define the Jacobian of the transformation:

$$\begin{aligned}\mathbf{J}_{xz}(\mathbf{x}, \theta) &= \left[\frac{\partial F_{xz}(\mathbf{x}, \theta)}{\partial x_i} \right] \\ \mathbf{J}_{zx}(\mathbf{z}, \theta) &= \left[\frac{\partial F_{zx}(\mathbf{z}, \theta)}{\partial z_i} \right]\end{aligned}\tag{5.3}$$

The absolute value of the Jacobian determinant $R_{zx}(\mathbf{z}) = |\det(\mathbf{J}_{zx}(\mathbf{z}, \theta))|$ measures how much a volume element at $\mathbf{z}$ is scaled with the transformation, and similarly for its inverse $R_{xz}(\mathbf{x}) = |\det(\mathbf{J}_{xz}(\mathbf{x}, \theta))|$. Knowing this, one can relate the probability distributions in real and latent space as:

$$\begin{aligned}p(\mathbf{x}) &= p(\mathbf{z})R_{zx}(\mathbf{z}) = p(F_{xz}(\mathbf{x}))R_{xz}^{-1}(\mathbf{x}) \\ p(\mathbf{z}) &= p(\mathbf{x})R_{xz}(\mathbf{x}) = p(F_{zx}(\mathbf{z}))R_{zx}^{-1}(\mathbf{z})\end{aligned}\tag{5.4}$$

And since the Jacobian of invertible neural networks is triangular, the calculation of this determinant is elementary.

Therefore, the method uses a simple probability distribution in latent space (Gaussian, logarithmic-normal, etc) and trains a Boltzmann generator so the resulting $p(\mathbf{x})$ probability distribution samples appropriately all relevant regions of the system's equilibrium distribution. In Machine Learning, training is usually performed by minimizing a certain cost function for a sample of problems. The other achievement of Noé *et al* is proposing the Kullback-Leibler divergence between the real Boltzmann distribution of the system and that calculated by the Boltzmann generator as this cost function. This KL divergence can be calculated for configurations expressed in the latent space $\mathbf{z}$ as:

$$J_{KL} = \mathbb{E}_{\mathbf{z}} [u(F_{zx}(\mathbf{z})) - \log R_{zx}(\mathbf{z})] \quad (5.5)$$

where $u(F_{zx}(\mathbf{z}))$ is the energy for the corresponding configuration in Cartesian space. The KL divergence is equivalent to the free energy of a configuration, containing both a potential and an entropic term. The $\mathbb{E} [u(F_{zx}(\mathbf{z}))]$ term encompasses the potential energy for the configuration and the $\mathbb{E} [R_{zx}(\mathbf{z})]$ term is equal to the entropic contribution at the chosen temperature plus a constant.

Given that this cost function requires no previous knowledge about a configuration to evaluate it, it can be applied for non supervised training. This way, the training process using energy training for the Boltzmann generator works as follows. The method first initializes the parameters for the transformation functions. After this, a training set is created by sampling the latent space, and the transformation function is actualized according to the gradient of the KL divergence. This training process continues until convergence.

Once we have found the optimal transformation function, the resulting Boltzmann distribution calculated by the Boltzmann generator usually requires some modification in order to better adjust to the real system's. In the case of Noé *et al*. they choose the simplest approach by assigning the statistical weight $w(\mathbf{x}) = e^{-u(\mathbf{x})} p_X(\mathbf{x})$ to every $\mathbf{x}$ configuration and consequently calculating the desired statistics.

However, although the whole process described here requires no previous knowledge about the equilibrium configuration one wants to sample, in practice a pure optimization of the KL divergence leads to the Boltzmann generator being trapped on the closest potential energy minima. Therefore, the actual implementation of the method proposed by Noé *et al*. involved a combined training by energy and

example, with the Boltzmann generator trying to maximize the probability $p_X(\mathbf{x})$ for some given reference configurations. Consequently, the method has yet only been applied for problems in which the target distribution is roughly known, such as the equilibrium between two configurations of a protein.

5.2 Method

We propose a new method to improve the applicability of Boltzmann generators for exploring potential energy surfaces and correctly sampling all relevant regions of the Boltzmann distribution, based on the adaptive bias strategy that characterizes other well known methods such as Metadynamics. In Metadynamics, the selected sampling method (local Monte Carlo, molecular dynamics, etc) is repeated iteratively while the system’s potential energy surface is gradually deformed, so future propagation avoids sampling already visited configurations. This way, as the process progresses, the different energy wells are filled by an artificial repulsive potential, allowing to eventually visit the full dimensional space.

However, a general problem with Metadynamics is the building of this artificial repulsive potential. Since one needs to evaluate how *close* is a given configuration to those already visited, the process necessarily implies the definition of a certain transition coordinate, the most common being the module of the Cartesian vector difference. This, in turn, can lead to the system becoming trapped by the repulsive potential, specially for systems with many degrees of freedom[141].

However, Boltzmann generators provide us with a non arbitrary method to build this repulsive potential using the obtained $p_X(\mathbf{x})$ distribution. In this way, for every iteration of the procedure we obtain a distribution $p_k(\mathbf{x})$ that samples preferentially those regions of the $\mathbf{x}$ space not already visited. This probability distribution can then be used to actualize the repulsive potential, so the process can continue.

We found two different approaches to implement the method, that differentiate between each other in how the information about the configurations already visited in previous iterations is stored. The easiest solution is to use a different Boltzmann generator associated to each $p_k(\mathbf{x})$ obtained. However, the high memory cost of keeping and the high computational cost of evaluating such a large amount of highly complex neural networks makes this method not practical. Fortunately, there is an

alternative: using a single Boltzmann generator to store the full information about all previous $p_k(\mathbf{x})$ generated. This is the approach we decide to implement.

To describe our method, we begin with a system described in the dimensional space $\mathbf{x}$ for which we know its potential energy operator $u(\mathbf{x})$. Then, the process progresses as follows:

- For the potential $u_k(\mathbf{x})$ we train a Boltzmann generator to find the probability distribution $p_k(\mathbf{x})$ that minimizes the KL divergence with respect the Boltzmann distribution determined by $u_k(\mathbf{x})$:

$$J_{KL}(p_k|u_k) = \int p_k(\mathbf{x}) [\log p_k(\mathbf{x}) + u_k(\mathbf{x})] d\mathbf{x} \quad (5.6)$$

- Train a Boltzmann generator to find the probability distribution $q_k(\mathbf{x})$ that minimizes the KL divergence:

$$J_{KL}[(q+p)_k|q_k] = \int (q+p)_k [\log [(q+p)_k(\mathbf{x})] - \log q_k(\mathbf{x})] d\mathbf{x} \quad (5.7)$$

where $(q+p)_k = \frac{k}{k+1}q_{k-1} + \frac{1}{k+1}p_k$ is the norm conservative linear combination between the distribution containing all visited configurations up to the $(k-1)$ th iteration for q_{k-1} and p_k .

- Train a neural network to learn the next iteration repulsive potential $\hat{u}_{k+1}(\mathbf{x})$ by training from examples sampled from the $q_k(\mathbf{x})$ distribution.

$$\hat{u}_{k+1}(\mathbf{x}) = \hat{u}_k(\mathbf{x}) + F_{\text{rep}}(p_k(\mathbf{x})) \quad (5.8)$$

where $F_{\text{rep}}(p_k(\mathbf{x}))$ is an appropriate repulsive potential designed to destabilize configurations visited during the k th iteration.

- Repeat the process from point one using the potential $u_{k+1}(\mathbf{x})$ defined as:

$$u_{k+1}(\mathbf{x}) = u(\mathbf{x}) + \gamma \frac{q_k(\mathbf{x})}{q_k(\mathbf{x}) + q_0} \hat{u}_k(\mathbf{x}) \quad (5.9)$$

where γ and ϵ are constants that determine both the intensity and the reach of the repulsion potential.

In our case, we choose as our repulsive function $F_{\text{rep}}(p)$:

$$F_{\text{rep}}(p) = \log(p + \epsilon) - \log(\epsilon) \quad (5.10)$$

There are many advantages to this implementation of our method. First, it only implies maintaining two Boltzmann generators (one to sample p_k and another for q_k) and another neural network to store the artificial repulsion potential. Furthermore, the form of Eq. 5.9 guarantees that the repulsive potential is always zero for non visited regions of space, for which the neural network $\hat{u}_k$ cannot be trained.

5.3 Results

We extend the original code by Noé *et al.* to implement the adaptive bias strategy. From the original code we take the RealNVP Boltzmann generator structure, which is applied for learning the different $p_k(\mathbf{x})$ and $q_k(\mathbf{x})$ probability distributions generated during the simulation. Our work consists of programming the new KL divergence functions represented by functions 5.6 and 5.7, designing a Neural Network to learn the cumulative repulsive potential $\hat{u}_k(\mathbf{x})$ and programming the biased potential energy operator in Eq. 5.9. We use the Python implementation of the Keras package to build the different Neural Networks and Tensorflow as the back-end engine.

To test the applicability of our method we study a 2D multiple well potential energy surface represented in Fig. 5.1. The potential presents four energy wells at $\mathbf{x} = (-1, -1), (1, -1), (-1, 1)$ and $(1, 1)$ with energies $-20, -16, -16$ and $-14k_B T$ respectively.

Our previous experience with the Boltzmann generators method showed that the pure energetic training approach cannot overcome energy barriers over $10k_B T$. This is confirmed for our four-well model. For a 10 layer RealNVP Boltzmann generator trained to minimize the KL divergence, the resulting $p(\mathbf{x})$ distribution collapses always to the lowest energy well at $(-1, -1)$. This can be seen in Fig. 5.2, that shows a typical free energy surface obtained through this method, with only one well sampled.

FIGURE 5.1: Potential Energy Surface for the 2D four well system used to test the adaptive bias Boltzmann generators algorithm

We apply our algorithm to the given problem, using values for the control parameters confirmed to give a stable and reproducible adaptive bias process for other simple one dimensional problems. To learn the cumulative repulsive potential we use a neural network with two hidden layers, each consisting of 100 neurons and for all Boltzmann generators we use a ten RealNVP layer structure.

We find that the method typically samples all four wells in the potential in just four iterations, leading to an almost flat biased potential $u_k(\mathbf{x})$. Fig. 5.7 and 5.8 show the successive repulsive potentials and the effective potential energy surface of the system after each iteration. As it can be seen, just applying a single Boltzmann generator based repulsive potential is enough to level the energy field and to allow the system to surpass the $10k_B T$ energy barriers between the wells.

For as many times as we repeat the simulation, we never find that the method takes more than five iterations to sample all wells. We therefore conclude that each iteration of the method is enough to cross barriers of about $10k_B T$ in energy.

FIGURE 5.2: Energy sampling of the four-well model using a single Boltzmann Generator and energy training. Non colored regions are not sampled by the method.

Finally, we sample 1000 configurations from the final $q_k(\mathbf{x})$ distribution and use them to train a new Boltzmann generator using an example and energy combined optimization function, in order to obtain the final predicted distribution of our method. In this aspect our algorithm just applies directly the Boltzmann generator method, and therefore any of the strategies proposed in the original paper to further optimize performance (such as the use of a different Boltzmann generators for each energy well), could be applied in a straight-forward manner.

In Fig. 5.5 we can see a summary of the results for the global sampling. Subfigures c) and d) show how the non linear transformation links different circular fractions of the two dimensional Gaussian distribution to each of the four energy wells, with the most central (and therefore most probable) region being linked to the deepest energy well (red points in the figure).

Fig. 5.5 also shows the free energy surface obtained from sampling the final Boltzmann generator. As expected, the four wells are sampled and the normalized energy differences between the four minimum points agree with the exact one.

FIGURE 5.3: Successive biased potentials $u_k(\mathbf{x})$ used to sample $p_{k+1}(\mathbf{x})$. It can be seen that as the algorithm fills the different energy wells the potential becomes flat

Another interesting feature in the figure is that several energy wells are connected by energy paths that roughly correspond to the minimum energy path. This effect was already pointed out in the original paper by Noé *et al.* and is a direct effect of the method used. The Boltzmann generator finds a simple linear transformation that moves through all the different wells as one samples the latent space. Therefore, the function also samples intermediate states connecting the different low energy states. In the future, this could be applied for finding the activated complex configuration for a reactions for which the initial and final state are known (saddle point optimization).

FIGURE 5.4: Successive repulsive potentials $\hat{u}_k(\mathbf{x})$ obtained

5.4 Conclusions

We have proposed and implemented a new application of the Boltzmann generators algorithm designed by Noé *et al.* that is able to sample complex potential energy surfaces by using an adaptive bias potential approach similar to that of MetaDynamics. The novelty of the method lies in that Boltzmann Generators allow us to use the real probability distribution function of the system to build the artificial repulsive potential required. This way, the method works iteratively and only requires the use of two Boltzmann generators, one that is trained at each step to sample the current deformed potential energy surface and another stores the overall information collected about the real system Boltzmann distribution through all iterations. Even more importantly, the method requires no prior knowledge about the system in order to be implemented and is therefore completely general.

We also show the functioning of the method with a simple toy model: a two dimensional potential energy surface with four wells, with energy barriers of about

FIGURE 5.5: Results for the final ML+KL optimization of a Boltzmann generator using the information collected during the sampling process: a) Non normalized ($\log(p_X(\mathbf{x}))$) free energy surface, b) Normalized ($u(\mathbf{x}) - \log(p_X(\mathbf{x}))$) free energy surface, c) and d) a collection of equilibrium configurations in Cartesian space taken from MC simulations and their respective representation in the latent space

$10k_B T$. While the original method by Noé *et al.* requires learning by example to sample all four energy wells, our method allows us to find all relevant configurations with no prior information. Even more, the final Boltzmann Generator also samples small regions of the energy wall, connecting the four wells, even though these regions were not sampled during the simulation. This effect was already pointed out by Noé *et al.*, and is feature of the energy training.

The method is ideal to sample potential energy surfaces with several minimums for which common Metadynamics cannot be applied. We plan to apply our method to study the free energy landscape of sticky hard sphere clusters, a system characterized by its high amount of stable configurations (that differ between each other in geometry but not energy) and that so far has been approached by building all possible geometries through analytical and numerical methods [142–145]. This system is completely athermal and the probabilities for each assembly with maximum

amount of bonds are purely determined by their dynamic entropy. So far, the complexity of the potential has not allowed to determine this entropical stabilization through thermodynamic integration[146]. Therefore, we will be able to compare our simulation results with those obtained through analytic approximate methods for the first time.

Chapter 6

Conclusions

6.1 Summary

Thorough this thesis we focused on the study of colloidal systems for which entropy can be applied as an effective method to control their phase diagram, allowing new crystal phases to emerge or stabilizing those already present. Our contributions can be divided in three parts, corresponding to Chapter 2; Chapter 3 and 4; and Chapter 5, respectively.

In Chapter 2 we focused on the self-assembly of mobile DNA coated colloids, a new recent implementation of DNA coated colloids that can rearrange themselves more easily as they grow, this way being less likely to be kinetically arrested[57]. By designing a new computational approach to calculate the complex multi-body interaction of the system[1], Angioletti-Uberti *et al.* were able to prove that mDNACC can be programmed to control the "valency" of their self-assembly structures without the need of patches, by tuning non specific repulsions between particles. However, the exact principles that control the self-assembly of mDNACC are not yet fully understood.

In our work, we show that mDNACC colloids can self-assemble to form open or *floppy* crystals, characterized for presenting fewer particle constraints than degrees of freedom and therefore not being rigid. We confirm the results by Angioletti-Uberti *et al.* and show that a 50:50 mixtures of colloids A and B with complementary DNA linkers can stabilize square and CsCl crystals for 2D and 3D respectively,

without the need of repulsion between particles. We also prove, using thermodynamic integration to the reference Einstein crystal, that this stabilization is purely entropic in nature, due to the higher mobility of floppy lattices.

Finally, the original method was modified to be able to calculate energy in the zero temperature limit, when all linkers are attached. For this case we also confirm stable floppy lattices at low temperatures and comparing our results with a simple sticky-sphere model, we show that this stabilization is also due to the higher vibrational entropy of the floppy crystals. Therefore, our results offer a new strategy for the self assembly of floppy crystals stable for any temperature lower than a limit. All our results has been published in [17].

Chapters 3 and 4 are dedicated to the study of melting of polydisperse hard-disks. Despite their apparent simplicity, the melting mechanism of hard-disks has been the subject of much discussion, specially since the publication of the KTHNY theory that proposes a microscopic mechanism of 2D melting based in topological changes. This discussion appears now to be over, after Bernard and Krauth[147] proved that the process occurs through a continuous solid-hexatic transition, closely followed by a first order hexatic-fluid transition. Their results also suggest a small likelihood of observing the hexatic phase experimentally, since their calculated single phase packing fraction stability range was only $\eta = [0.718 - 0.720]$.

This last result contrasts with the much wider range found experimentally analyzing the sedimentation of 2D spherical colloids on a tilted plane[42]. This discrepancy led us to study whether the variation in the experimental particle sizes could explain this disagreement. By studying the phase diagram of Gaussian distributed polydisperse hard-disks in the $NVT\Delta\mu$ ensemble, we first found that polydispersity can induce a return to the KTHNY melting mechanism, with two consecutive continuous transitions, for system's average polydispersities over 7% ($s/\langle\sigma\rangle > 0.07$). We also found that polydispersity can indeed increase the packing fraction stability range of the hexatic phase for over one order of magnitude, and confirm that this favoring occurs due to the increase of dislocations for high polydispersity systems, that destabilize the solid phase. Finally, our most surprising result is the presence of re-entrant melting, that is, for systems with polydispersity parameters over 8% ($\nu/\sigma_0 > 0.08$), compressing the solid leads to the formation of a new fluid phase at high pressures, with the density vs average polydispersity phase diagram presenting a low and high pressure branch for the solid-hexatic and

hexatic-fluid transitions. This new phase is confirmed to be a fluid by analyzing its time correlation function. Therefore, for the polydisperse hard-disks system, we confirm that introducing disorder in the form of variation of particle size leads to very significant effects in the physics of the melting. All our results were published in [18].

However, a logical critique against our results is their generality. In real life experimental implementations it is not possible to synthesize a solution of colloids that follow a perfect Gaussian-like distribution with respect to their size. Therefore, in Chapter 4 we studied the effect that the size probability distribution on the phase diagram of polydisperse hard-disks. We do this by obtaining the phase diagrams for a lognormal distributed system and a triangular one. For the lognormal system we confirm that polydispersity can induce a return to the KTHNY melting mechanism and a significant increase in the hexatic phase stability range. However, with respect to the re-entrant melting, we find that the lognormal system cannot present it, since unlike for the Gaussian distribution, compression of the system leads to a monotonical decrease of average polydispersity. We do, however, find that increasing the polydispersity parameter of the exchange potential can lead to a counter-intuitive decrease in the average polydispersity at transition.

Finally, we find that for the triangular distribution, polydispersity cannot induce significant effects in the melting of hard-disks, as all high valued polydispersity parameter equations of state collapse to a common isotherm at the transition to the solid. Because of this, the formed hexatic and solid phase always present average polydispersities under 6%, too low for the observed effects in the case of the Gaussian. We explain the occurring of this collapse by showing that at high pressures the triangular distribution behaves identically independently of the polydispersity parameter.

From our results from this chapter we can conclude that polydispersity is indeed a reliable tool for stabilizing the hexatic phase of the hard-disk system, with the effect of the distribution shape being minimal. However, probability distribution shape significantly affects the observed re-entrant melting for the Gaussian like distribution and it is not possible to describe a general phase diagram.

Finally, in Chapter 5, we first introduced Boltzmann generators[138], a new application of deep learning to the sampling of free energies. The key idea of Boltzmann

generators is to link the real system's probability distribution in the Cartesian space $p(\mathbf{x})$ to a simple Gaussian distribution in latent space $p(\mathbf{z})$ through a trainable invertible transformation. This allow us to obtain the probability of a certain configuration in real space from that of its equivalent configuration in latent space.

Then, we proposed a new method to sample the potential energy surface of complex multidimensional systems: adaptive bias Boltzmann generators. Our method follows the same strategy as other adaptive bias methods, such as Meta-Dynamics or the Wang-Landau algorithm. It helps the simulation to surpass energy barriers by filling the potential wells using an artificial repulsive potential that favours the exploration of non visited configurations. However, unlike the previously mentioned methods, adaptive bias Boltzmann generators do not require the definition of an arbitrary order parameter to determine how *close* is a given configuration to those already visited. Instead, it uses the probability distribution $p(\mathbf{x})$ given by Boltzmann Generators to build this repulsive potential. As all adaptive bias methods, it works iteratively, with the repulsive potential evolving as new information is obtained.

Finally, after describing the new method, we show its functioning by applying it to sample a simple four-well two dimensional potential. We believe this method will be able in the future to sample the free energy of sticky sphere clusters, a complex problem in which stability is determined purely by mobility and that so far has only been treated by calculating entropy using simple analytical methods based on approximations.

6.2 Final remarks

6.2.1 Polydisperse hard-disks

Our findings for the polydisperse hard-disk phase diagram open a new door for experimental scientist to observe the hexatic phase in hard sphere like colloids that was apparently closed by previous publications. Indeed, the significant packing fraction stabilization observed and its robustness against changes in the used probability size distribution suggest that our results are general and suited to be

confirmed experimentally. However, our results are not likely to explain the discrepancies found for the experiment by Thorneywork *et al* that were commented on Chapter 1, as in their case particle polydispersity was below 2%, significantly below the 7% limit found in our simulations required for a significant effect on the hard-disk phase diagram to be observed. Other experimental conditions that may have caused this stabilization of the hexatic phase include particle shape, pinning of particles to the container, out of plane fluctuations and vacancies.

With respect of further directions of research, we believe that the most interesting question left unanswered in our research is the effect that the re-entrant melting observed for Gaussian distributed polydisperse hard-disks would have in the sedimentation of polydisperse hard spheres. One would expect that the pressure gradient naturally caused by the incremental weight of the different layers of hard spheres should induce an initial formation of the quasi-solid, followed by a high pressure re-entrant liquid for even deeper layers. This would effectively conform a solid sandwiched between two stable liquid phases.

6.2.2 Adaptive bias Boltzmann generators

For this section, our work is yet preliminary and it is therefore the one that offers more possible paths of future research. The next obvious research direction is, as commented in Chapter 5, the application of the method to the study of a realistic problem. As we explained, the method is ideal for the sampling of the potential energy surface with many energy minimums and for which it is not easy to describe an order parameter that permits the use of Metadynamics. Currently we are working to implement the method to the study of the conformations of sticky particles clusters of up to ten particles. This would mean the first study of the free energy surface of this system using a computational method.

Once we implement the method for a realistic problem, it would be interesting comparing its performance against other adaptive bias potential methods, such as local elevation[148], conformational flooding[149], metadynamics[139], the Wang-Landau algorithm[149] and the adaptive biasing force algorithm[150]. It is possible that adaptive bias Boltzmann generators outperforms its alternatives even in systems for which an useful transition coordinate can be defined, since our method can surpass a $10k_B T$ energy barrier in a single iteration.

Also, as we briefly commented in Chapter 5, our method allows the use of any arbitrary Boltzmann generator’s technique to compute free energies, as its function is mostly to probe all interesting regions of a complex potential energy surface. In this work, we generated our final probability distribution for the 2D four well system using a single Boltzmann Generator that simultaneously captures multiple metastable states and maps them to the same latent space Z . However, Noé and colleagues have already proposed other approaches more suitable for complex systems, when it could be difficult to train a single Boltzmann generator to sample all relevant configurations. This is done by using several Boltzmann generators and relating them to a common reference for which free energy is known.

In the case of Boltzmann generators, finding a the free energy difference between the configurations sampled by two different latent is simple, thanks to the energy loss function J_{KL} that gives us, as we showed in Chapter 5, an account of the free energy difference of transforming a configuration in latent space to its correspondent in the Cartesian. Therefore, we can obtain an average of the cost function for the two transformations ($\langle J_{KL}^1 \rangle$ and $\langle J_{KL}^2 \rangle$) by sampling from the two Gaussian prior distributions as $\Delta A_{12} = \langle J_{KL}^2 \rangle - \langle J_{KL}^1 \rangle$.

This approach could not only be useful for the final free energy sampling, but also during the iterative adaptive bias process. Namely, the probability distribution $q_k(\mathbf{x})$ that samples all already visited regions of the Cartesian space could be implemented using as many Boltzmann Generators as needed with the probability for a certain configuration being taken as the maximum probability given by all Boltzmann generators $q_k(\mathbf{x}) = \max(q_{k,1}(\mathbf{x}), \dots, q_{k,m}(\mathbf{x}))$, and the sampling divided accordingly to their Boltzmann weights. This approach also opens the possibility of parallelization, with several independent adaptive bias Boltzmann generators being run simultaneously and exchanging information between iterations.

These are only some logical future steps in order to make our method as general as possible. We believe this is a promising algorithm that could help finding new metastable states in systems where other adaptive bias potentials methods have failed. Therefore, this could be a promising research project.

Appendix A

Additional figure for Chapter 3

FIGURE A.1: Packing fraction η as functions of density $\rho\sigma_0^2$ for PHDS with $\nu/\sigma_0 = 0.082$ (black), 0.083 (red), and 0.0835 (green). Errors are smaller than the symbols.

Bibliography

- [1] Stefano Angioletti-Uberti, Bortolo M. Mognetti, and Daan Frenkel. *Phys. Chem. Chem. Phys.*, 18(9):6373–6393, MAR 7 2016. [vii](#), [24](#), [28](#), [83](#)
- [2] William G. Hoover and Francis H. Ree. Melting transition and communal entropy for hard spheres. *The Journal of Chemical Physics*, 49(8):3609–3617, 1968. doi: 10.1063/1.1670641. URL <https://doi.org/10.1063/1.1670641>. [2](#)
- [3] G. Ernst, C. Broholm, G. R. Kowach, and A. P. Ramirez. Phonon density of states and negative thermal expansion in zrw2o8. *Nature*, 396(6707):147–149, Nov 1998. ISSN 1476-4687. doi: 10.1038/24115. URL <https://doi.org/10.1038/24115>. [2](#)
- [4] G. N. Greaves, A. L. Greer, R. S. Lakes, and T. Rouxel. Poisson’s ratio and modern materials. *Nature Materials*, 10(11):823–837, Nov 2011. ISSN 1476-4660. doi: 10.1038/nmat3134. URL <https://doi.org/10.1038/nmat3134>. [2](#)
- [5] Joseph N. Grima, Ruben Gatt, Victor Zammit, Jennifer J. Williams, Kenneth E. Evans, Andrew Alderson, and Richard I. Walton. Natrolite: A zeolite with negative poissons ratios. *Journal of Applied Physics*, 101(8):086102, 2007. doi: 10.1063/1.2718879. URL <https://doi.org/10.1063/1.2718879>.
- [6] K. Sun, A. Souslov, X. Mao, and T. C. Lubensky. Surface phonons, elastic response, and conformal invariance in twisted kagome lattices. *PNAS*, 109(31):12369–12374, 2012. ISSN 0027-8424. doi: 10.1073/pnas.1119941109. URL <https://doi.org/10.1073/pnas.1119941109>. [2](#)
- [7] J. D. Joannopoulos, Pierre R. Villeneuve, and Shanhui Fan. Photonic crystals: putting a new twist on light. *Nature*, 386(6621):143–149, Mar 1997. ISSN 1476-4687. doi: 10.1038/386143a0. URL <https://doi.org/10.1038/386143a0>. [2](#)
- [8] Alexander Moroz. Metallo-dielectric diamond and zinc-blende photonic crystals. *Phys. Rev. B*, 66:115109, Sep 2002. doi: 10.1103/PhysRevB.66.115109. URL <https://link.aps.org/doi/10.1103/PhysRevB.66.115109>. [2](#)

- [9] Henry Cohn and Abhinav Kumar. Algorithmic design of self-assembling structures. *Proceedings of the National Academy of Sciences*, 106(24):9570–9575, 2009. [2](#)
- [10] Erik Edlund, Oskar Lindgren, and M Nilsson Jacobi. Novel self-assembled morphologies from isotropic interactions. *Physical review letters*, 107(8):085501, 2011.
- [11] Salvatore Torquato. Inverse optimization techniques for targeted self-assembly. *Soft Matter*, 5(6):1157–1173, 2009.
- [12] Krassimir P Velikov, Christina G Christova, Roel PA Dullens, and Alfons van Blaaderen. Layer-by-layer growth of binary colloidal crystals. *Science*, 296(5565):106–109, 2002.
- [13] Sharon C Glotzer and Michael J Solomon. Anisotropy of building blocks and their assembly into complex structures. *Nature materials*, 6(8):557–562, 2007. [2](#)
- [14] Xiaoming Mao, Qian Chen, and Steve Granick. Entropy favours open colloidal lattices. *Nature Materials*, 12(3):217–222, MAR 2013. ISSN 1476-1122. [2](#), [34](#), [36](#)
- [15] Xiaoming Mao. Entropic effects in the self-assembly of open lattices from patchy particles. *Phys. Rev. E*, 87(6):062319, JUN 24 2013. ISSN 1539-3755. [34](#), [36](#)
- [16] D Zeb Rocklin and Xiaoming Mao. Self-assembly of three-dimensional open structures using patchy colloidal particles. *Soft Matter*, 10(38):7569–7576, 2014. [2](#)
- [17] Hao Hu, Pablo Sampedro Ruiz, and Ran Ni. Entropy stabilizes floppy crystals of mobile dna-coated colloids. *Physical review letters*, 120(4):048003, 2018. [3](#), [32](#), [84](#)
- [18] Pablo Sampedro Ruiz, Qun-li Lei, and Ran Ni. Melting and re-entrant melting of polydisperse hard disks. *Communications Physics*, 2(1):1–6, 2019. [3](#), [46](#), [85](#)
- [19] BI Halperin and David R Nelson. Theory of two-dimensional melting. *Physical Review Letters*, 41(2):121, 1978. [5](#)
- [20] David R Nelson and BI Halperin. Dislocation-mediated melting in two dimensions. *Physical Review B*, 19(5):2457, 1979. [5](#)
- [21] Paul M Chaikin, Tom C Lubensky, and Thomas A Witten. *Principles of condensed matter physics*, volume 10. Cambridge university press Cambridge, 1995. [9](#), [11](#)
- [22] S. T. Chui. Grain-boundary theory of melting in two dimensions. *Phys. Rev. Lett.*, 48:933–935, 1982. [14](#), [37](#)

- [23] Eytan Domany, Michael Schick, and Robert H Swendsen. First-order transition in an xy model with nearest-neighbor interactions. *Physical review letters*, 52(17):1535, 1984. [14](#)
- [24] Sebastian C. Kapfer and Werner Krauth. Two-dimensional melting: From liquid-hexatic coexistence to continuous transitions. *Phys. Rev. Lett.*, 114: 035702, 2015. [15](#), [37](#), [42](#)
- [25] Peter Bladon and Daan Frenkel. Dislocation unbinding in dense two-dimensional crystals. *Physical review letters*, 74(13):2519, 1995. [15](#)
- [26] Joshua A. Anderson, James Antonaglia, Jaime A. Millan, Michael Engel, and Sharon C. Glotzer. Shape and symmetry determine two-dimensional melting transitions of hard regular polygons. *Phys. Rev. X*, 7:021001, 2017. [15](#), [37](#), [46](#)
- [27] C. A. Murray and D. H. Van Winkle. Experimental observation of two-stage melting in a classical two-dimensional screened coulomb system. *Phys. Rev. Lett.*, 58:1200–1203, 1987. [15](#), [37](#)
- [28] Gert Lube, Herbert E Huppert, R Stephen J Sparks, and Armin Freundt. Collapses of two-dimensional granular columns. *Physical Review E*, 72(4): 041301, 2005. [15](#)
- [29] Sven Deutschländer, Tobias Horn, Hartmut Löwen, Georg Maret, and Peter Keim. Two-dimensional melting under quenched disorder. *Phys. Rev. Lett.*, 111:098301, 2013. [15](#), [38](#)
- [30] Weikai Qi and Marjolein Dijkstra. Destabilisation of the hexatic phase in systems of hard disks by quenched disorder due to pinning on a lattice. *Soft Matter*, 11:2852, 2015. [15](#), [38](#)
- [31] BJ Alder and TE Wainwright. Phase transition in elastic disks. *Physical Review*, 127(2):359, 1962. [15](#)
- [32] Kevin J Naidoo and Jurgen Schnitker. Melting of two-dimensional colloidal crystals: A simulation study of the yukawa system. *The Journal of chemical physics*, 100(4):3114–3121, 1994. [15](#)
- [33] KW Wojciechowski. Two-dimensional isotropic system with a negative poisson ratio. *Physics Letters A*, 137(1-2):60–64, 1989. [15](#)
- [34] Julio F Fernández, Juan J Alonso, and Jolanta Stankiewicz. One-stage continuous melting transition in two dimensions. *Physical review letters*, 75(19): 3477, 1995. [15](#)
- [35] Andreas Jaster. Computer simulations of the two-dimensional melting transition using hard disks. *Physical Review E*, 59(3):2594, 1999. [15](#)

- [36] Martin A Bates and Daan Frenkel. Influence of vacancies on the melting transition of hard disks in two dimensions. *Physical Review E*, 61(5):5223, 2000. [15](#)
- [37] Etienne P. Bernard and Werner Krauth. Two-step melting in two dimensions: First-order liquid-hexatic transition. *Phys. Rev. Lett.*, 107:155704, Oct 2011. doi: 10.1103/PhysRevLett.107.155704. URL <https://link.aps.org/doi/10.1103/PhysRevLett.107.155704>. [15](#), [37](#), [38](#), [42](#), [44](#), [57](#)
- [38] Michael Engel, Joshua A. Anderson, Sharon C. Glotzer, Masaharu Isobe, Etienne P. Bernard, and Werner Krauth. Hard-disk equation of state: First-order liquid-hexatic transition in two dimensions with three simulation methods. *Phys. Rev. E*, 87:042134, Apr 2013. doi: 10.1103/PhysRevE.87.042134. URL <https://link.aps.org/doi/10.1103/PhysRevE.87.042134>. [15](#), [20](#), [37](#), [38](#), [42](#)
- [39] John Russo and Nigel B. Wilding. Disappearance of the hexatic phase in a binary mixture of hard disks. *Phys. Rev. Lett.*, 119:115702, 2017. [16](#), [38](#)
- [40] Andrew H Marcus and Stuart A Rice. Observations of first-order liquid-to-hexatic and hexatic-to-solid phase transitions in a confined colloid suspension. *Physical review letters*, 77(12):2577, 1996. [16](#)
- [41] Pallop Karnchanaphanurach, Binhua Lin, and Stuart A. Rice. Melting transition in a quasi-two-dimensional colloid suspension: Influence of the colloid-colloid interaction. *Phys. Rev. E*, 61:4036–4044, 2000. [16](#), [37](#)
- [42] Alice L. Thorneywork, Joshua L. Abbott, Dirk G. A. L. Aarts, and Roel P. A. Dullens. Two-dimensional melting of colloidal hard spheres. *Phys. Rev. Lett.*, 118:158001, Apr 2017. doi: 10.1103/PhysRevLett.118.158001. URL <https://link.aps.org/doi/10.1103/PhysRevLett.118.158001>. [16](#), [38](#), [40](#), [67](#), [84](#)
- [43] Ben Widom. Some topics in the theory of fluids. *The Journal of Chemical Physics*, 39(11):2808–2812, 1963. [18](#)
- [44] Evert Jan Meijer. *Computer Simulations of Molecular Solids and Colloidal Dispersions*. PhD thesis, Rijksuniversiteit Utrecht, The Netherlands, 1993. An optional note. [18](#)
- [45] Daan Frenkel and Anthony JC Ladd. New monte carlo method to compute the free energy of arbitrary solids. application to the fcc and hcp phases of hard spheres. *The Journal of chemical physics*, 81(7):3188–3193, 1984. [19](#)
- [46] M. Michel, S.C. Karpfer, and W. Krauth. Generalized event-chain monte carlo: Constructing rejection-free global-balance algorithms from infinitesimal steps. *J. Chem. Phys.*, (140):054116, 2013. [20](#)

- [47] Etienne P. Bernard, Werner Krauth, and David B. Wilson. Event-chain monte carlo algorithms for hard-sphere systems. *Phys. Rev. E*, 80:056704, Nov 2009. doi: 10.1103/PhysRevE.80.056704. URL <https://link.aps.org/doi/10.1103/PhysRevE.80.056704>. 20
- [48] Persi Diaconis, Susan Holmes, and Radford M Neal. Analysis of a nonreversible markov chain sampler. *Annals of Applied Probability*, pages 726–752, 2000. 21
- [49] J.D. Watson and F.H.C.A. Crick. *Nature*, 171:737, 1953. 23
- [50] CA Mirkin, RL Letsinger, RC Mucic, and JJ Storhoff. A dna-based method for rationally assembling nanoparticles into macroscopic materials. *Nature*, 382(6592):607–609, AUG 15 1996. ISSN 0028-0836. 23
- [51] AP Alivisatos, KP Johnsson, XG Peng, TE Wilson, CJ Loweth, MP Bruchez, and PG Schultz. Organization of ‘nanocrystal molecules’ using dna. *Nature*, 382(6592):609–611, AUG 15 1996. ISSN 0028-0836. 23
- [52] Robert J. Macfarlane, Byeongdu Lee, Matthew R. Jones, Nadine Harris, George C. Schatz, and Chad A. Mirkin. Nanoparticle superlattice engineering with dna. *Science*, 334(6053):204–208, OCT 14 2011. ISSN 0036-8075. 23
- [53] Dmytro Nykypanchuk, Mathew M. Maye, Daniel van der Lelie, and Oleg Gang. Dna-guided crystallization of colloidal nanoparticles. *Nature*, 451(7178):549–552, JAN 31 2008. ISSN 0028-0836.
- [54] Marie T. Casey, Raynaldo T. Scarlett, W. Benjamin Rogers, Ian Jenkins, Talid Sinno, and John C. Crocker. Driving diffusionless transformations in colloidal crystals using dna handshaking. *Nature Communications*, 3:1209, NOV 2012. 28
- [55] Sung Yong Park, Abigail K. R. Lytton-Jean, Byeongdu Lee, Steven Weigand, George C. Schatz, and Chad A. Mirkin. Dna-programmable nanoparticle crystallization. *Nature*, 451(7178):553–556, JAN 31 2008. ISSN 0028-0836.
- [56] Evelyn Auyeung, Ting I. N. G. Li, Andrew J. Senesi, Abrin L. Schmucker, Bridget C. Pals, Monica Olvera de la Cruz, and Chad A. Mirkin. Dna-mediated nanoparticle crystallization into wulff polyhedra. *Nature*, 505(7481):73–77, JAN 2 2014. ISSN 0028-0836.
- [57] Yu Wang, Yufeng Wang, Xiaolong Zheng, Etienne Ducrot, Jeremy S. Yodh, Marcus Weck, and David J. Pine. Crystallization of dna-coated colloids. *Nature Communications*, 6:7253, JUN 2015. 23, 28, 83
- [58] Yufeng Wang, Yu Wang, Xiaolong Zheng, Etienne Ducrot, Myung-Goo Lee, Gi-Ra Yi, Marcus Weck, and David J. Pine. Synthetic strategies toward dna-coated colloids that crystallize. *J. Am. Chem. Soc.*, 137(33):10760–10766, AUG 26 2015. ISSN 0002-7863. 23

- [59] Yufeng Wang, Yu Wang, Dana R. Breed, Vinothan N. Manoharan, Lang Feng, Andrew D. Hollingsworth, Marcus Weck, and David J. Pine. Colloids with valence and specific directional bonding. *Nature*, 491(7422):51–U61, NOV 1 2012. ISSN 0028-0836. [23](#), [24](#)
- [60] Ian C. Jenkins, Marie T. Casey, James T. McGinley, John C. Crocker, and Talid Sinno. Hydrodynamics selects the pathway for displacive transformations in dna-linked colloidal crystallites. *Proc. Natl. Acad. Sci USA*, 111(13):4803–4808, APR 1 2014. ISSN 0027-8424. [23](#), [34](#)
- [61] Minseok Song, Yajin Ding, Hasan Zerbe, Mark A. Snyder, and Jeetain Mittal. Binary superlattice design by controlling dna-mediated interactions. *arXiv:1703.03465*, 2017. [23](#), [25](#), [28](#)
- [62] Stef A. J. van der Meulen and Mirjam E. Leunissen. Solid colloids with surface-mobile dna linkers. *J. Am. Chem. Soc.*, 135(40):15129–15134, OCT 9 2013. ISSN 0002-7863. [24](#), [25](#)
- [63] Stef A. J. van der Meulen, Gesa Helms, and Marileen Dogterom. Solid colloids with surface-mobile linkers. *J. Phys.: Condens. Matter Phys.*, 27:233101, JUN 17 2015. [24](#)
- [64] Lang Feng, Remi Dreyfus, Ruojie Sha, Nadrian C. Seeman, and Paul M. Chaikin. Dna patchy particles. *Adv. Mater.*, 25(20):2779–2783, MAY 28 2013. ISSN 0935-9648. [24](#)
- [65] Stefano Angioletti-Uberti, Patrick Varilly, Bortolo M. Mognetti, and Daan Frenkel. Mobile linkers on dna-coated colloids: Valency without patches. *Phys. Rev. Lett.*, 113(12):128303, SEP 17 2014. [24](#), [25](#), [28](#), [36](#)
- [66] Patrick Varilly, Stefano Angioletti-Uberti, Bortolo M. Mognetti, and Daan Frenkel. A general theory of dna-mediated and other valence-limited colloidal interactions. *J. Chem. Phys.*, 137(9):094108, SEP 7 2012.
- [67] Stefano Angioletti-Uberti, Patrick Varilly, Bortolo M. Mognetti, Alexei V. Tkachenko, and Daan Frenkel. Communication: A simple analytical formula for the free energy of ligand-receptor-mediated interactions. *J. Chem. Phys.*, 138(2):021102, JAN 14 2013. ISSN 0021-9606. [25](#), [28](#)
- [68] Daan Frenkel and Ben Smit. *Understanding Molecular Simulation: from Algorithms to applications, 2nd edition, Elsevier (Singapore) Pte Ltd. (2010)*. [25](#)
- [69] Robert Aveyard, John H. Clint, Dieter Nees, and Vesselin N. Paunov. Compression and structure of monolayers of charged latex particles at air/water and octane/water interfaces. *Langmuir*, 16(4):1969–1979, 2000. doi: 10.1021/la990887g. [25](#)

- [70] Behnaz Bozorgui and Daan Frenkel. Liquid-vapor transition driven by bond disorder. *Phys. Rev. Lett.*, 101:045701, Jul 2008. doi: 10.1103/PhysRevLett.101.045701. URL <https://link.aps.org/doi/10.1103/PhysRevLett.101.045701>. 27
- [71] Francisco J. Martinez-Veracoechea, Bianca M. Mladek, Alexei V. Tkachenko, and Daan Frenkel. Design rule for colloidal crystals of dna-functionalized particles. *Phys. Rev. Lett.*, 107(4):045902, JUL 20 2011. ISSN 0031-9007. 27, 28
- [72] A.-P. Hynninen, M. E. Leunissen, A. van Blaaderen, and M. Dijkstra. Cuau structure in the restricted primitive model and oppositely charged colloids. *Phys. Rev. Lett.*, 96:018303, Jan 2006. doi: 10.1103/PhysRevLett.96.018303. URL <https://link.aps.org/doi/10.1103/PhysRevLett.96.018303>. 27
- [73] Ryan V Thaner, Youngeun Kim, Ting I N G Li, R J Macfarlane, S T Nguyen, Monica Olvera de la Cruz, and C A Mirkin. Entropy-driven crystallization behavior in dna-mediated nanoparticle assembly. *Nano Lett.*, 15:5545–5551, 2015. 27
- [74] W. Benjamin Rogers, William M. Shih, and Vinodhan N. Manoharan. Using dna to program the self-assembly of colloidal nanoparticles and microparticles. *Nature Reviews Materials*, 1(3):16008, MAR 2016. 27
- [75] Bianca M. Mladek, Julia Fornleitner, Francisco J. Martinez-Veracoechea, Alexandre Dawid, and Daan Frenkel. Quantitative prediction of the phase diagram of dna-functionalized nanosized colloids. *Phys. Rev. Lett.*, 108(26):268301, JUN 25 2012. ISSN 0031-9007. 28, 36
- [76] Ke Chen, Wouter G. Ellenbroek, Zexin Zhang, Daniel T. N. Chen, Peter J. Yunker, Silke Henkes, Carolina Brito, Olivier Dauchot, Wim van Saarloos, Andrea J. Liu, and A. G. Yodh. Low-frequency vibrations of soft colloidal glasses. *Phys. Rev. Lett.*, 105(2):025501, JUL 9 2010. ISSN 0031-9007. 34
- [77] M. Dennison, M. Sheinman, C. Storm, and F. C. MacKintosh. Fluctuation-stabilized marginal networks and anomalous entropic elasticity. *Phys. Rev. Lett.*, 111(9):095503, AUG 27 2013. ISSN 0031-9007. 34
- [78] Qian Chen, Sung Chul Bae, and Steve Granick. Directed self-assembly of a colloidal kagome lattice. *Nature*, 469(7330):381–384, JAN 20 2011. ISSN 0028-0836. 36
- [79] S. F. Shimobayashi, B. M. Mognetti, L. Parolini, D. Orsi, P. Cicuta, and L. Di Michele. Direct measurement of dna-mediated adhesion between lipid bilayers. *Phys. Chem. Chem. Phys.*, 17(24):15615–15628, 2015. ISSN 1463-9076. 36
- [80] L. Parolini, B.M. Mognetti, J. Kotar, E. Eiser, P. Cicuta, and L. Di Michele. Volume and porosity thermal regulation in lipid mesophases by coupling mobile ligands to soft membranes. *Nat. Commun.*, 6(5948), 2015.

- [81] L. Parolini, J. Kotar, L. Di Michele, and B.M. Mognetti. Controlling self-assembly kinetics of dna-functionalized liposomes using toehold exchange mechanism. *ACS Nano*, 10(2392), 2016.
- [82] S.J. Bachmann, J. Kotar, L. Parolini, A. Saric, P. Cicuta, L. Di Michele, and B.M. Mognetti. Melting transition in lipid vesicles functionalised by mobile dna linkers. *Soft Matter*, 12(7804), 2016. [36](#)
- [83] Frank Smallenburg, Ludwik Leibler, and Francesco Sciortino. Patchy particle model for vitrimers. *Phys. Rev. Lett.*, 111:188002, Oct 2013. doi: 10.1103/PhysRevLett.111.188002. URL <https://link.aps.org/doi/10.1103/PhysRevLett.111.188002>. [36](#)
- [84] Flavio Romano and Francesco Sciortino. Switching bonds in a dna gel: An all-dna vitrimer. *Phys. Rev. Lett.*, 114:078104, Feb 2015. doi: 10.1103/PhysRevLett.114.078104. URL <https://link.aps.org/doi/10.1103/PhysRevLett.114.078104>. [36](#)
- [85] Katherine J. Strandburg. Two-dimensional melting. *Rev. Mod. Phys.*, 60: 161–207, Jan 1988. doi: 10.1103/RevModPhys.60.161. URL <https://link.aps.org/doi/10.1103/RevModPhys.60.161>. [37](#)
- [86] J. G. Dash. History of the search for continuous melting. *Rev. Mod. Phys.*, 71:1737–1743, Oct 1999. doi: 10.1103/RevModPhys.71.1737. URL <https://link.aps.org/doi/10.1103/RevModPhys.71.1737>.
- [87] U. Gasser. Crystallization in three- and two-dimensional colloidal suspensions. *Journal of Physics: Condensed Matter*, 21(20):203101, 2009. [37](#)
- [88] J.M. Kosterlitz and D.J. Thouless. Ordering, metastability and phase transitions in two-dimensional systems. *Journal of Physics C: Solid State Physics*, 6(7):1181, 1973. [37](#), [42](#)
- [89] B. I. Halperin and David R. Nelson. Theory of two-dimensional melting. *Phys. Rev. Lett.*, 41:121–124, 1978.
- [90] David R. Nelson and B. I. Halperin. Dislocation-mediated melting in two dimensions. *Phys. Rev. B*, 19:2457–2484, 1979.
- [91] A. P. Young. Melting and the vector coulomb gas in two dimensions. *Phys. Rev. B*, 19:1855–1866, 1979. [42](#)
- [92] V. Berezinskii. Destruction of long range order in one-dimensional and two-dimensional systems having a continuous symmetry group. i. classical systems. *Sov. Phys. JETP*, 32:493, 1971. [37](#)
- [93] Kurt Binder, Surajit Sengupta, and Peter Nielaba. The liquid-solid transition of hard discs: first-order transition or kosterlitz-thouless-halperin-nelson-young scenario? *Journal of Physics: Condensed Matter*, 14(9):2323, 2002. [37](#)

- [94] Y. Saito. Melting of dislocation vector systems in two dimensions. *Phys. Rev. Lett.*, 48:1114–1117, 1982. [37](#)
- [95] K. Zahn, R. Lenke, and G. Maret. Two-stage melting of paramagnetic colloidal crystals in two dimensions. *Phys. Rev. Lett.*, 82:2721–2724, 1999. [37](#)
- [96] Y. Han, N. Y. Ha, A. M. Alsayed, and A. G. Yodh. Melting of two-dimensional tunable-diameter colloidal crystals. *Phys. Rev. E*, 77:041406, 2008.
- [97] Stuart A. Rice. Structure in confined colloid suspensions. *Chemical Physics Letters*, 479(1):1 – 13, 2009.
- [98] Andrew H. Marcus and Stuart A. Rice. Observations of first-order liquid-to-hexatic and hexatic-to-solid phase transitions in a confined colloid suspension. *Phys. Rev. Lett.*, 77:2577–2580, 1996.
- [99] H. H. von Grünberg, P. Keim, K. Zahn, and G. Maret. Elastic behavior of a two-dimensional crystal near melting. *Phys. Rev. Lett.*, 93:255703, 2004.
- [100] P. Keim, G. Maret, and H. H. von Grünberg. Frank’s constant in the hexatic phase. *Phys. Rev. E*, 75:031402, 2007.
- [101] Xinliang Xu and Stuart A. Rice. Liquid-to-hexatic phase transition in a quasi-two-dimensional colloid system. *Phys. Rev. E*, 78:011602, 2008. [37](#)
- [102] Yan-Wei Li and Massimo Pica Ciamarra. Role of cell deformability in the two-dimensional melting of biological tissues. *Phys. Rev. Materials*, 2:045602, 2018. [37](#)
- [103] T. Kawasaki, T. Araki, and H. Tanaka. Correlation between dynamic heterogeneity and medium-range order in two-dimensional glass-forming liquids. *Phys. Rev. Lett.*, 99:215701, 2007. [38](#)
- [104] T. Kawasaki and H. Tanaka. Structural signature of slow dynamics and dynamic heterogeneity in two-dimensional colloidal liquids: Glassy structural order. *J Phys. Condens. Matter.*, 23:194121, 2011.
- [105] J. Russo and H. Tanaka. Assessing the role of static length scales behind glassy dynamics in polydisperse hard disks. *Proc. Natl Acad. Sci. USA*, 112:6920, 2015. [38](#)
- [106] Moreno Fasolo and Peter Sollich. Equilibrium phase behavior of polydisperse hard spheres. *Phys. Rev. Lett.*, 91:068301, Aug 2003. doi: 10.1103/PhysRevLett.91.068301. URL <https://link.aps.org/doi/10.1103/PhysRevLett.91.068301>. [38](#), [51](#), [57](#), [59](#)
- [107] David A. Kofke and Eduardo D. Glandt. Monte carlo simulation of multi-component equilibria in a semigrand canonical ensemble. *Molecular Physics*, 64(6):1105–1131, 1988. [38](#), [59](#)

- [108] Peter G. Bolhuis and David A. Kofke. Monte carlo study of freezing of polydisperse hard spheres. *Phys. Rev. E*, 54:634–643, Jul 1996. doi: 10.1103/PhysRevE.54.634. URL <https://link.aps.org/doi/10.1103/PhysRevE.54.634>.
- [109] David A. Kofke and Peter G. Bolhuis. Freezing of polydisperse hard spheres. *Phys. Rev. E*, 59:618–622, Jan 1999. doi: 10.1103/PhysRevE.59.618. URL <https://link.aps.org/doi/10.1103/PhysRevE.59.618>.
- [110] Sander Pronk and Daan Frenkel. Melting of polydisperse hard disks. *Phys. Rev. E*, 69:066123, Jun 2004. doi: 10.1103/PhysRevE.69.066123. URL <https://link.aps.org/doi/10.1103/PhysRevE.69.066123>. [38](#), [42](#), [44](#), [45](#), [59](#)
- [111] Joseph E. Mayer and Wm. W. Wood. Interfacial tension effects in finite, periodic, twodimensional systems. *The Journal of Chemical Physics*, 42(12): 4268–4274, 1965. [42](#), [61](#)
- [112] Nadezhda Gribova, Axel Arnold, Tanja Schilling, and Christian Holm. How close to two dimensions does a lennard-jones system need to be to produce a hexatic phase? *The Journal of Chemical Physics*, 135(5):054514, 2011. [44](#)
- [113] Arthur E. Ferdinand and Michael E. Fisher. Bounded and inhomogeneous ising models. i. specific-heat anomaly of a finite lattice. *Phys. Rev.*, 185: 832–846, 1969. [44](#)
- [114] Yan-Wei Li and Massimo Pica Ciamarra. Accurate determination of the translational correlation function of two-dimensional solids. *Physical Review E*, 100(6):062606, 2019. [45](#)
- [115] Amir Hajibabaei and Kwang S Kim. First-order and continuous melting transitions in two-dimensional lennard-jones systems and repulsive disks. *Physical Review E*, 99(2):022145, 2019. [46](#)
- [116] Weikai Qi, A.P. Gantapara, and Marjolein Dijkstra. Two-stage melting induced by dislocations and grain boundaries in monolayers of hard spheres. *Soft Matter*, 10:5449, 2014. [49](#)
- [117] Ludovic Berthier, Daniele Coslovich, Andrea Ninarello, and Misaki Ozawa. Equilibrium sampling of hard spheres up to the jamming density and beyond. *Phys. Rev. Lett.*, 116:238002, 2016. [49](#)
- [118] Andrea Ninarello, Ludovic Berthier, and Daniele Coslovich. Models and algorithms for the next generation of glass transition studies. *Phys. Rev. X*, 7:021039, 2017.
- [119] Harukuni Ikeda, Francesco Zamponi, and Atsushi Ikeda. Mean field theory of the swap monte carlo algorithm. *The Journal of Chemical Physics*, 147 (23):234506, 2017. [49](#)

- [120] Mengjie Zu, Jun Liu, Hua Tong, and Ning Xu. Density affects the nature of the hexatic-liquid transition in two-dimensional melting of soft-core systems. *Phys. Rev. Lett.*, 117:085702, Aug 2016. doi: 10.1103/PhysRevLett.117.085702. URL <https://link.aps.org/doi/10.1103/PhysRevLett.117.085702>. 51
- [121] D.E. Dudalov, Yu.D. Fomin, E.N. Tsiok, and V.N. Ryzhov. Anomalous melting scenario of the two-dimensional core-softened system. *arXiv:1311.7534v4*, 2014. 51
- [122] Paul Bartlett and Patrick B. Warren. Reentrant melting in polydispersed hard spheres. *Phys. Rev. Lett.*, 82:1979–1982, Mar 1999. doi: 10.1103/PhysRevLett.82.1979. URL <https://link.aps.org/doi/10.1103/PhysRevLett.82.1979>. 51, 57, 59
- [123] D. C. Rapaport. *The Art of Molecular Dynamics Simulation*. Cambridge University Press, 2004. 51
- [124] Alekse Grigorevich Ivakhnenko and Valentin Grigorevich Lapa. *Cybernetic predicting devices*. 1965. 69
- [125] A G Ivakhnenko; V G Lapa; Robert N McDonough. *Cybernetic and forecasting devices*. 1967. 69
- [126] G. E. Hinton and R. R. Salakhutdinov. Reducing the dimensionality of data with neural networks. *Science*, 313(5786):504–507, 2006. ISSN 0036-8075. doi: 10.1126/science.1127647. URL <https://science.sciencemag.org/content/313/5786/504>. 69
- [127] Seppo Linnainmaa. *The representation of the cumulative rounding error of an algorithm as a Taylor expansion of the local rounding errors*. PhD thesis, 1981. 69
- [128] P. J. Werbos. Applications of advances in nonlinear sensitivity analysis. pages 762–770, 1981. 69
- [129] Javier Gonzalez-Dominguez, Ignacio Lopez-Moreno, and Hasim Sak. Automatic language identification using long short-term memory recurrent neural networks. In *Interspeech*, 2014. 69
- [130] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. *Commun. ACM*, 60(6): 8490, May 2017. ISSN 0001-0782. doi: 10.1145/3065386. URL <https://doi.org/10.1145/3065386>. 69
- [131] Juan Carrasquilla and Roger G. Melko. Machine learning phases of matter. *Nature Physics*, 13(5):431–434, May 2017. ISSN 1745-2481. doi: 10.1038/nphys4035. URL <https://doi.org/10.1038/nphys4035>. 69

- [132] Alan Morningstar and Roger G. Melko. Deep learning the ising model near criticality, 2017. 69
- [133] Jörg Behler and Michele Parrinello. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.*, 98:146401, Apr 2007. doi: 10.1103/PhysRevLett.98.146401. URL <https://link.aps.org/doi/10.1103/PhysRevLett.98.146401>. 69
- [134] Jrg Behler. Perspective: Machine learning potentials for atomistic simulations. *The Journal of Chemical Physics*, 145(17):170901, 2016. doi: 10.1063/1.4966192. URL <https://doi.org/10.1063/1.4966192>. 69
- [135] Giuseppe Carleo and Matthias Troyer. Solving the quantum many-body problem with artificial neural networks. *Science*, 355(6325):602–606, 2017. ISSN 0036-8075. doi: 10.1126/science.aag2302. URL <https://science.sciencemag.org/content/355/6325/602>. 69
- [136] J. Tubiana and R. Monasson. Emergence of compositional representations in restricted boltzmann machines. *Phys. Rev. Lett.*, 118:138301, Mar 2017. doi: 10.1103/PhysRevLett.118.138301. URL <https://link.aps.org/doi/10.1103/PhysRevLett.118.138301>. 69
- [137] S. Cocco, R. Monasson, L. Posani, S. Rosay, and J. Tubiana. Statistical physics and representations in real and artificial neural networks. *Physica A: Statistical Mechanics and its Applications*, 504:4576, Aug 2018. ISSN 0378-4371. doi: 10.1016/j.physa.2017.11.153. URL <http://dx.doi.org/10.1016/j.physa.2017.11.153>. 69
- [138] Frank Noé, Simon Olsson, Jonas Köhler, and Hao Wu. Boltzmann generators: Sampling equilibrium states of many-body systems with deep learning. *Science*, 365(6457), 2019. ISSN 0036-8075. doi: 10.1126/science.aaw1147. URL <https://science.sciencemag.org/content/365/6457/eaaw1147>. 69, 85
- [139] Alessandro Laio and Michele Parrinello. Escaping free-energy minima. *Proceedings of the National Academy of Sciences of the United States of America*, 99(20):12562–12566, Oct 2002. ISSN 0027-8424. doi: 10.1073/pnas.202427399. URL <https://pubmed.ncbi.nlm.nih.gov/12271136>. 12271136[pmid]. 70, 87
- [140] Laurent Dinh, Jascha Sohl-Dickstein, and Samy Bengio. Density estimation using real nvp, 2016. 70
- [141] Jeffrey Comer, James C. Gumbart, Jrme Hnin, Tony Lelivre, Andrew Pohorille, and Christophe Chipot. The adaptive biasing force method: Everything you always wanted to know but were afraid to ask. *The Journal of Physical Chemistry B*, 119(3):1129–1151, 2015. doi: 10.1021/jp506633n. URL <https://doi.org/10.1021/jp506633n>. PMID: 25247823. 73
- [142] V. Manoharan N. Arkus and M. Brenner. Deriving finite sphere packings. *SIAM Journal on Discrete Mathematics*, 25(4):18601901, 2011. 80

-
- [143] Jared Harwayne-Gidansky Robert S Hoy and Corey S O'Hern. Structure of finite sphere packings via exact enumeration: Implications for colloidal crystal nucleation. *Physical Review E*, 85(5):18601901, May 2012.
- [144] Robert S Hoy. Structure and dynamics of model colloidal clusters with short-range attractions. *Physical Review E*, 91(1):0123037, 2015.
- [145] M. Holmes-Cerfon. Enumerating rigid sphere packings. *SIAM Review*, 58(2):229244, May 2016. [80](#)
- [146] Yoav Kallus and Miranda Holmes-Cerfon. Free energy of singular sticky-sphere clusters. *Phys. Rev. E*, 95:022130, Feb 2017. doi: 10.1103/PhysRevE.95.022130. URL <https://link.aps.org/doi/10.1103/PhysRevE.95.022130>. [81](#)
- [147] Etienne P. Bernard and Werner Krauth. Two-step melting in two dimensions: First-order liquid-hexatic transition. *Phys. Rev. Lett.*, 107:155704, Oct 2011. doi: 10.1103/PhysRevLett.107.155704. URL <https://link.aps.org/doi/10.1103/PhysRevLett.107.155704>. [84](#)
- [148] Thomas Huber, Andrew E Torda, and Wilfred F Van Gunsteren. Local elevation: a method for improving the searching properties of molecular dynamics simulation. *Journal of computer-aided molecular design*, 8(6):695–708, 1994. [87](#)
- [149] Helmut Grubmüller. Predicting slow structural transitions in macromolecular systems: Conformational flooding. *Physical Review E*, 52(3):2893, 1995. [87](#)
- [150] Eric Darve and Andrew Pohorille. Calculating free energies using average force. *The Journal of chemical physics*, 115(20):9169–9183, 2001. [87](#)