Rings and Things

David Ormrod Morley

Balliol College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2020

Contents

List of Notes By David

List of Notes By Mark

1 Introduction

1.1 Background

The notion of describing amorphous materials as random networks dates back to Zachariasen, who in 1932 sketched a simple diagram of a two-dimensional glass [1]. This configuration, reproduced in figure 1.1a, showed a collection of percolating rings with an absence of long-range order. At the time, Zachariasen's image was intended only as schematic to illustrate the analogous effects in three-dimensional glasses. However, some eighty years later, modern synthesis techniques have led to a range of two-dimensional materials including amorphous carbon, silica and germania which can be considered realisations of Zachariasen's glass [2–8]. These advances may yet represent a watershed moment in chemistry, facilitating the development of a wide range of technologically useful materials with applications including catalysis and gas separation [9–11].

It is clear that understanding the structure of amorphous materials is key to this aim. However, due to the relative recentness of these experimental discoveries, much of the existing theory arises from studies of systems on greater length scales. Specifically, in the second half of the 20$^{\text{th}}$ century, much work was carried out on the formation of polycrystals in metals and alloys. By annealing the metal and slicing through the sample, the grains in the polycrystal could be directly imaged; revealing a system of tessellating polygons not dissimilar to an atomic material [12, 13]. Over time it became apparent that the structure of these networks is constrained on a series of different levels.

Firstly the mean ring size (*i.e.* the average number of sides in a polygon) tends to the constant value of six. This is readily explainable via graph theoretic arguments, simply resulting from Euler's formula when each vertex forms part of three edges - as is the case for trivalent atoms or the meeting of three grain boundaries. Intuitively from chemistry we know this to be true: a pristine graphene sheet is a hexagonal net and although a Stone-Wales defect introduces pentagons and heptagons, they occur in pairs to preserve the overall mean ring size [14].

The next level of information is then the explicit distribution of polygon sizes, also known as the ring statistics. With the constraint of a fixed mean, the ring statistics were shown to be relatively well defined, following a lognormal or maximum entropy distribution [15–17]. However, the ring statistics alone are not sufficient to fully describe the network topology. This is because the same set of rings can be arranged in a large number of different ways. Consider again Zachariasen's original configuration. Removing one square achieves a mean ring size of six and allows the constituent rings to be arranged as a periodic tiling. Figures 1.1b-1.1d show three such examples tilings.

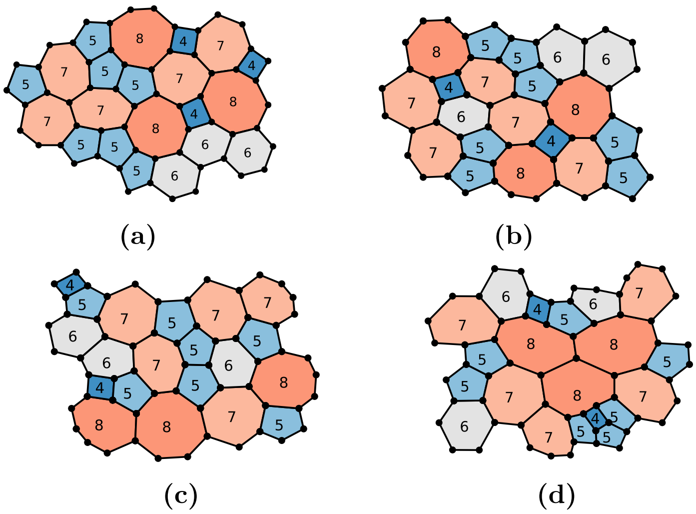


Figure 1.1: Panel (a) shows Zachariasen's glass and panels (b)-(d) three different periodic arrangements based on the glass (with one square removed to satisfy Euler's formula). Moving from panel (b)-(d) there is increased clustering of similar sized rings. The size of the rings are highlighted numerically and by colour.

Whilst they may initially look similar, on closer inspection the three configurations display fundamentally different properties. In figure 1.1b similar sized rings are dispersed throughout the arrangement whilst in 1.1d they are tightly clustered together. Furthermore, given the large number of configurations which may be theoretically possible for any set of ring statistics, only a subset of these may be physically realisable. Empirically, these are found to be the ones in which large rings tend to be surrounded by smaller rings *i.e.* similar to 1.1b. Once again, chemical intuition would support this in the context of atomic materials, as strain is minimised by maintaining bond lengths and angles as close to their equilibrium values as possible. This effect was first noticed in polycrystals and quantified through the Aboav-Weaire law [18, 19]. This law claims that the mean ring size about any given ring can be related to the central ring size by a single fitting parameter. Hence the value of this parameter in some way describes the increased tendency of the small rings to be adjacent to large rings. The Aboav-Weaire parameter therefore provides information on the first-order ring correlations, completing the topological description of the network material.

The novelty and potential usefulness of two-dimensional materials makes them a clear candidate for computational study, in order to complement and supplement experimental endeavours. Taking the example of thin silica films, there have already been multiple complementary computational investigations including both *ab initio* methods and molecular dynamics studies using classical force fields at varying levels of theory [20–27]. In order to perform these simulations, it is necessary to have a starting atomistic configuration. This can be achieved in multiple ways. The most straightforward is to take one of the existing experimental images. These are however limited in size and number and can contain defects or areas which cannot be fully imaged. As a result, computational techniques are often preferable, but generating configurations will the required topological properties (*i.e.* correct ring statistics *and* Aboav-Weaire parameter) has proved surprisingly difficult [28, 29]. Therefore, the first part of this thesis will focus on developing methods to generate configurations of two-dimensional networks in which the topological parameters can be tuned in a controllable manner. These configurations can then be used as a seed for further computational studies, removing the reliance for experimental configurations and opening the door for high-throughput calculations which can be speculative and potentially predictive.

However, the scope of this work extends beyond materials modelling. As previously mentioned, much of the original work in this field focussed on polycrystals of metal oxides with some links to foams and Voronoi polygons [30, 31]. It is now clear that these chemical networks fit into a much wider class of two-dimensional physical networks that are ubiquitous in the natural world, emerging across all physical disciplines and length scales. Traditional examples range from the atomic level of ultra-thin materials, through colloids, foams, epithelial cells and all the way to geological rock formations [32–37]. There are however countless more occurrences, with drying blood, stratocumulus clouds, crocodile scales and geopolitical borders all being the subject of studies [38–41]. More intriguingly, although these systems are incredibly physically diverse, they still have similar structures [42]. This is because they can all be mapped onto the same generic system, which can be equivalently described as a collection of tessellating polygons or percolating rings, and hence they are governed by the same fundamental laws. Understanding the behaviours of two-dimensional networks is therefore key to a wide range of problems in frontier research, not only the directed synthesis of nano-materials but also for example the control of mitotic division [43, 44]; as well as to curiosities such as explaining the arrangement of the stones in Giant's Causeway or cracking in famous artworks [45, 46].

Furthermore, the continuing expansion and maturity of network science as a field has led to significant advances in the description and characterisation of complex networks. This has largely been driven by interest in networks in the more abstract sense of the internet, social media and neural networks [47–49]. To date, the application of these principles in the physical sciences has mostly been confined to topics such as biological signalling pathways. The second half of this thesis will therefore show how robust metrics from network science can be applied to physical two-dimensional networks to better quantify their structure and replace the need for empirical measures such as the Aboav-Weaire law. This also has the effect of tying physical two-dimensional networks into the wider field of network science, showing them to be a unique and interesting addition to the area.

As part of this process, more generic methods will be developed to construct two-dimensional networks across a range of potential models, coordination environments and topologies. This will allow a systematic study into the factors which influence the underlying network properties in two-dimensional systems. These will be compared to two further in-depth studies of network forming structures from the physical sciences. Expand colloid/procrystals bit. The first are Voronoi tessellations formed in colloidal monolayers. which can be simulated via hard particle models [50]. The second are the ring structures in so-called "procrystalline" lattices [51]. Paragraph to wrap up PH here? Yep PH here

1.2 Thesis Structure

2 | Network Theory

The theory underpinning complex networks is discussed, covering the representation of atomic systems as networks and the relationship of the dual network to ring structure. The laws which govern the topological properties of physical networks, namely Euler's law, Lemaître's law and the Aboav-Weaire law are also introduced.

2.1 Network Theory

The scope of what constitutes a complex network is extremely broad, covering everything from the tangible (*e.g.* computational clusters) to the more abstract (*e.g.* social interactions). Yet part of the appeal and power of network science is the ability to quantify and relate these highly disparate systems with the same underlying theory. A network is simply a collection of components termed *nodes* and the connections between them termed *links*, an example of which is given in figure 2.1. There are then two fundamental classes of network based on the nature of the connections. Networks in which the links between nodes are mutual are termed undirected, whereas those in which the links are one-way are termed directed [52]. At the risk of dating this thesis, this is the difference between Facebook (an undirected social network of friends) and Twitter (a directed social network of followers). All the networks considered in this work are undirected and all the theory assumes this property.

2.1.1 Node Degree and Probability Distributions

A key concept in network science is the the node degree, defined as the number of links that each node possesses. A node with k links is then said simply to have

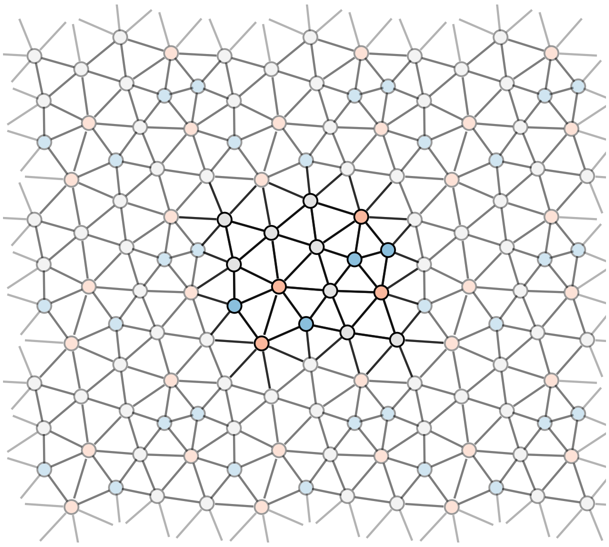


Figure 2.1: Example of a periodic two-dimensional network where nodes are represented by circles and links as lines. Nodes are coloured similarly according to their degree, whilst periodic images are greyed out to highlight the central repeating unit.

degree k, where $k \in \mathbb{N}$. This is illustrated in figure 2.1, which consists of 5- (blue), 6- (grey) and 7- (red) degree nodes. The occurrence and correlations of nodes of given degrees can then be described by a range of probability distributions.

The probability of a randomly selected node having degree k is given by the node degree distribution, denoted p_k. This is a normalised discrete distribution such that

$$\sum_k p_k = 1\,. \tag{2.1}$$

The n^{th} moments of this distribution are then given by:

$$\langle k^n \rangle = \sum_k k^n p_k\,. \tag{2.2}$$

Alternatively, one can also calculate the probability that a randomly selected link has a k-degree node at the end, denoted q_k. This is not the same as the distribution above, as there is greater chance of selecting links which emanate from high degree nodes, in a manner which is proportional to the node degree. As this distribution is normalised, this leads to the relations:

$$\sum_k q_k = 1 \tag{2.3}$$

$$q_k = \frac{kp_k}{\langle k \rangle}\,. \tag{2.4}$$

In addition, one can also evaluate the probability that a randomly chosen link has nodes of degree j,k at either end. This is the node joint degree distribution, denoted e_{jk}. Once again this is normalised and satisfies the following relationships:

$$\sum_{jk} e_{jk} = 1, \tag{2.5}$$

$$\sum_{jk} e_{jk} = q_j \tag{2.6}$$

$$e_{jk} = e_{kj}\ , \tag{2.7}$$

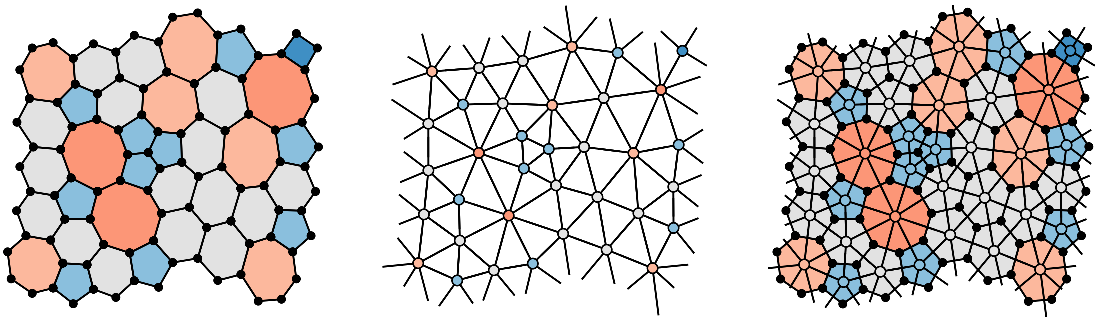
where the final result arises from reciprocal nature of the links in an undirected network. As an example, these three probability distributions are provided for the network in figure 2.1:

$$\mathbf{p} = \frac{1}{16}\begin{array}{cc} \begin{bmatrix} 4 \\ 8 \\ 4 \end{bmatrix} & \begin{array}{c} 5 \\ 6 \\ 7 \end{array} \end{array} \qquad \mathbf{q} = \frac{1}{96}\begin{array}{cc} \begin{bmatrix} 20 \\ 48 \\ 28 \end{bmatrix} & \begin{array}{c} 5 \\ 6 \\ 7 \end{array} \end{array} \qquad \mathbf{e} = \frac{1}{96}\begin{array}{cc} \begin{array}{ccc} 5 & 6 & 7 \end{array} & \\ \begin{bmatrix} 2 & 9 & 9 \\ 9 & 22 & 17 \\ 9 & 17 & 2 \end{bmatrix} & \begin{array}{c} 5 \\ 6 \\ 7 \end{array} \end{array}\ . \tag{2.8}$$

2.1.2 Atomic and Ring Networks

To see how network theory relates to atomic materials, consider the amorphous graphene configuration in figure 2.2a. In this network the nodes represent carbon atoms and the links sp^2 bonds. The node degree in the atomic network for all nodes is then equal to three, being equivalent to the atomic coordination number (which throughout this thesis will be denoted by c). This is problematic, because whilst there is clear disorder in the system, it is not well captured by the atomic network. Due to the fact that the local environment around the atoms is identical, when examining say the node degree distribution any information about the glassy structure is lost. This network is to first order indeterminable from a crystalline hexagonal lattice.

Observing figure 2.2a one can see there is another level of structure in the network, namely that of the ring structure. A ring is strictly any closed path of sequentially linked nodes in a network, but this thesis will use the term in reference only to the primitive rings *i.e.* those which cannot be subdivided into two smaller rings [53]. A



(a) Atomic network. **(b)** Ring network. **(c)** Dual relationship.

Figure 2.2: Panel (a) gives an example of a 3- coordinate periodic atomic network with disordered ring structure. Nodes and links represent atoms and bonds respectively where rings are coloured by size. Panel (b) gives the corresponding ring network where nodes and links represent rings and their adjacencies, where nodes are coloured by degree. Panel (c) shows the dual relationship between the atomic and ring networks, where the node degree in the ring network is equal to the ring size in the atomic network.

ring of size k (or k-ring) is then defined as a ring with k constituent nodes. It is clear that finding and counting the number of rings of each size, often termed calculating the ring statistics, does then quantify the disorder in the system [29]. The ring statistics can be summarised by the normalised probability distribution, p_k.

However, there is a more efficient way of representing and quantifying the ring structure in the system, and that is by constructing the dual network [54]. The dual is generated by placing a node at the centre of each ring and linking the nodes of adjacent (*i.e.* edge-sharing) rings, as can be seen in figure 2.2b. This will be referred to as the ring network. The ring network is a reciprocal lattice in which the node degree, k, is equivalent to the ring size in the atomic network. Similarly, it consists solely of triangles, reflecting the 3-coordinate nature of the underlying atomic network. Hence, the disorder is captured directly in the node properties of the ring network. These characteristics make the ring network preferable for manipulating and analysing the systems in this thesis.

2.2 Topological Laws

There are a number of laws which govern the topological properties of two-dimensional network-forming materials. These laws constrain the ring structure, influencing the network properties in a manner that makes physical networks unique in the field of network science. These laws act on a number of "levels": Euler's law controls the overall mean ring size, Lemaître's law the ring size distribution and the Aboav-Weaire law the ring-ring correlations.

2.2.1 Euler's Law

Euler's law constrains the mean ring size, $\langle k \rangle$, in an atomic network or equivalently the mean node degree of the ring network. The atomic networks studied in this work are all two-dimensional, connected (there is a path between any two nodes) and planar (they have no overlapping links) and so are subject to Euler's formula which states:

$$N + V - E = \chi, \tag{2.9}$$

where N, V, E are the number of rings, vertices and edges in the network and χ in an integer termed the Euler characteristic, which is dependent on the global topology of the system. Each vertex represents an atom and the number of edges emanating from each vertex is then the coordination number.

For generality consider an atomic network with atoms of assorted coordination numbers, c. If the proportion of each coordination type is x_c, then the mean coordination number is given by $\langle c \rangle = \sum\limits_c c x_c$. This allows the number of edges to be written in terms of the number of vertices as $E = \frac{V}{2}\langle c \rangle$. In turn the mean ring size is simply the total number of vertices per ring, allowing for multiple counting, such that $\langle k \rangle = \frac{V}{N}\langle c \rangle$. Substituting these two expressions into equation (2.9) leads to the expression:

$$\langle k \rangle = \frac{2\langle c \rangle \left(1 - \chi/N\right)}{\langle c \rangle - 2}. \tag{2.10}$$

Hence the average node degree in the ring network (equivalent to the mean ring size of the physical network), is simply related to the average degree of the physical network (*i.e.* local coordination environment), the topology of the system and the number of rings.

Although equation (2.10) may appear simple, it is a very powerful constraint. To demonstrate this consider a two-dimensional lattice with two possible coordination environments $c = 3, 4$. The planar case with periodic boundary conditions (mimicking an infinite planar lattice) maps onto the torus with $\chi = 0$, and so:

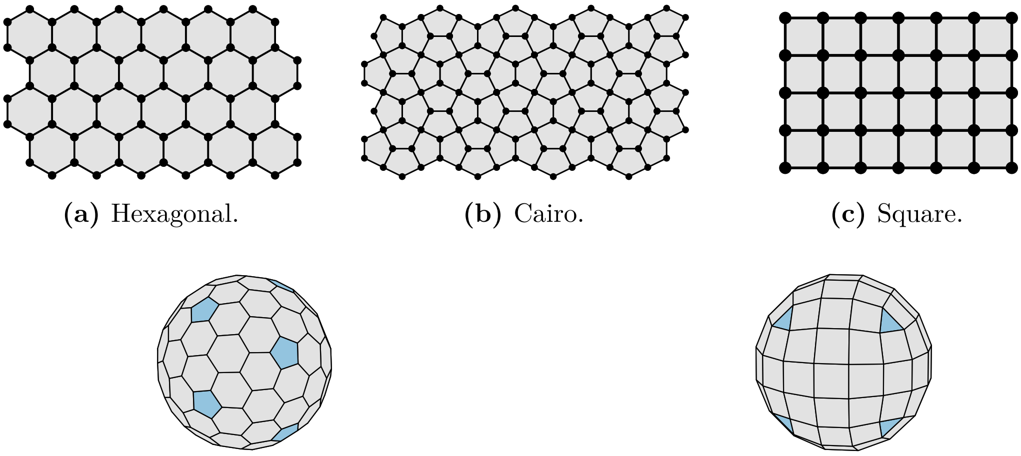
$$\langle k \rangle = \begin{cases} 6, & x_3 = 1 \\ 4, & x_4 = 1 \\ 5, & x_3 = 2/3, x_4 = 1/3 \end{cases} . \tag{2.11}$$

To reiterate in plain terms, this means that if there is a material consisting of atoms all forming exactly three bonds (as for amorphous carbon), the mean ring size *must* be equal to six. Similarly if all atoms form four bonds the mean ring size is four, and if there is a two-thirds to one-third mixture of coordination environments the mean ring size is five. The simplest illustrations of these are the hexagonal, square and cairo regular tilings, shown in figure 2.3, but this law holds equally well for amorphous configurations. For aperiodic systems strictly $\chi = 1$, but as $N \to \infty$, the proportion of vertices with unsatisfied coordination on the sample perimeter become negligible overall as does the term in χ. Therefore in reality these relationships hold, and remain as applicable to amorphous graphene as the basalt columns in Fingal's Cave, and the Penrose tiling [37, 55].

This analysis also extends to spherical topology where $\chi = 2$, and so:

$$\langle k \rangle = \begin{cases} \frac{6N-12}{N}, & x_3 = 1 \\ \frac{4N-8}{N}, & x_4 = 1. \end{cases} \tag{2.12}$$

These relationships are the origin of the 12 pentagon rule for 3-coordinate fullerenes (the "football problem"), or equivalently an "8 triangle rule" in the 4-coordinate case, as this is the only way to satisfy these equations if the allowed ring sizes are limited to $k = 5, 6$ and $k = 3, 4$ respectively (as in figures 2.3d, 2.3e) [56]. Much of the richness in the behaviour of two-dimensional physical networks stems from this fundamental constraint on the network average degree.



(a) Hexagonal. **(b)** Cairo. **(c)** Square.

(d) 3-coordinate fullerene. **(e)** 4-coordinate fullerene.

Figure 2.3: Panels (a)-(c) give regular planar tilings of 6-, 5- and 4- rings, where the ring size is related to the underlying atomic coordination. Panels (d) and (e) show the 3- and 4- coordinate tilings in spherical topology, where the mean ring size is reduced due to the change in the Euler characteristic.

2.2.2 Lemaître's Law

Knowing that the mean node degree is fixed by Euler's law, the next level of available information is the form of the underlying degree distribution, p_k. Interestingly, the degree distributions found in physical ring networks seem relatively well defined. For instance, it has been noted in models and realisations of two-dimensional silica glass that the ring statistics looked to follow a lognormal distribution [11, 15]. Lemaître *et al.* demonstrated that the distribution in 3-coordinate networks systems can be well described by a maximum entropy distribution [57]. Lemaître's maximum entropy method is summarised here, trivially extended to arbitrary coordination.

The entropy of a probability distribution is defined as

$$\mathcal{S} = -\sum_k p_k \log p_k. \tag{2.13}$$

In addition, the degree distribution has the following constraints:

$$\sum_k p_k = 1, \tag{2.14}$$

$$\sum_k k p_k = \langle k \rangle, \tag{2.15}$$

$$\sum_k \frac{p_k}{k} = \text{constant}, \tag{2.16}$$

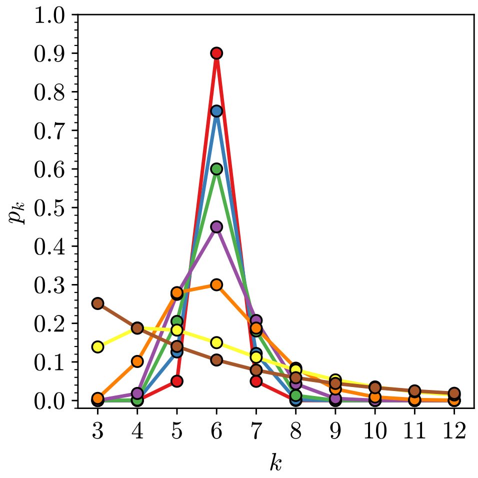
where the first two constraints correspond to the normalisation condition and the fixed mean ring size, and the final constraint will be discussed below. The entropy can then be maximised using Lagrange's method of undetermined multipliers to yield the result:

$$p_k = \frac{e^{-\lambda_1 k - \lambda_2 / k}}{\sum\limits_k e^{-\lambda_1 k - \lambda_2 / k}}, \tag{2.17}$$

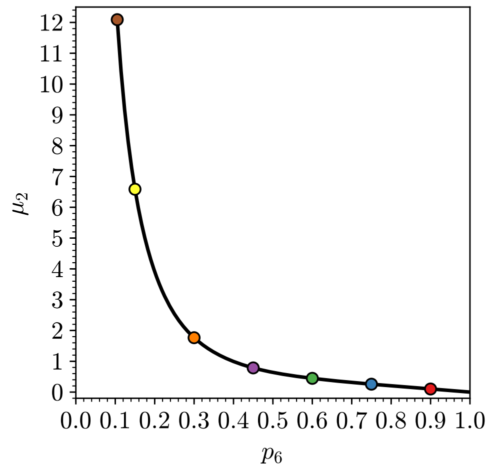
which can be solved numerically by substitution into equations (2.15),(2.16). By allowing the chosen constant to vary, a family of maximum entropy curves can be generated, as in figure 2.4a. The resulting distributions can be summarised by relating the variance, $\mu_2 = \langle k^2 \rangle - \langle k \rangle^2$, to a single chosen node degree probability, leading to the plot known as Lemaître's law, given in figure 2.4b. It is usually framed in the context of the proportion of hexagons in a system, p_6, for the precise reason that most networks have $\langle k \rangle = 6$ and p_6 as the largest contribution. Many experimental and theoretical studies have shown good agreement to this law [58–60].

Simple extensions of the classic law are however possible, by modifying the mean degree or the permitted degree range. For instance, k is usually taken in the interval $k \geq 3$ (as the triangle, $k = 3$, is the smallest polygon), but there can be manifestations of physical systems where only certain degrees are accessible [61]. Additional examples of such systems will be procrystalline lattices explored in chapter 8. The resulting Lemaître curves for a selection of these modifications are given in figure 2.4c. A discussion of these will be recur throughout this thesis, but one can see that the application of the allowable ring size constraints leads to marked differences in the maximum entropy solutions. The maximum value of these curves can be simply determined by removing constraint (2.16), equivalent to setting $\lambda_2 = 0$ in equation (2.17).

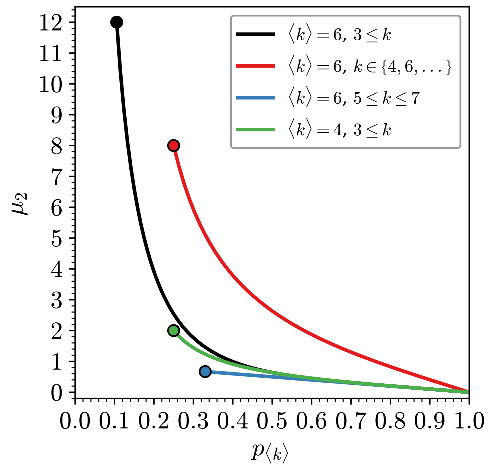
The only somewhat puzzling aspect of this successful theory is the choice of constraint (2.16). It was originally rationalised on the basis that the areas of rings of a given size, A_k, can be well fit by an expression $A_k = ak + b + c/k$, where a, b and c are constants. As noted at the time, this is by no means true for all systems and in fact is contrary to the widely known Lewis law, which states that A_k is linear in k



(a) Maximum entropy distributions.



(b) Lemaître's law.



(c) Extensions to Lemaître's law.

Figure 2.4: Illustration of Lemaître's maximum entropy method. Panel (a) gives examples of explicit maximum entropy distributions with different values of p_6. Panel (b) shows how these distributions can be summarised in a plot of p_6 *vs.* μ_2 (Lemaître's law). Panel (c) provides extensions to the law by modifying the underlying constraints of the mean ring size and allowable k-range.

for many observable networks [62–64]. Despite this, the universality of the Lemaître law suggests that there must be a physical basis to (2.16), and in the section 6.2.2 it will be demonstrated that it can be regenerated by considering ring adjacencies.

2.2.3 Aboav-Weaire Law

The ring statistics given by Lemaître's law are an important measure for physical networks, but they do not provide a complete characterisation of the ring structure, as they say nothing about the ring adjacencies. This is important because whilst with the same ring statistics it is theoretically possible to organise the rings in many different arrangements, it is well known experimentally that only a subsection of these are observed. The vast majority of physical systems have a preference for small rings ($k < \langle k \rangle$) be adjacent to large rings ($k > \langle k \rangle$). This effect was first noted in the grains of polycrystals by Aboav [18]. Aboav quantified these ring correlations by measuring the mean ring size about a k-ring, denoted m_k, and found empirically that $m_k \approx 5 + 8/k$.

In an attempt to explain this observation, Weaire came across the following relation

$$\sum_k k m_k p_k = \sum_k k^2 p_k = \mu_2 + \langle k \rangle^2 \,, \tag{2.18}$$

known as Weaire's sum rule [19]. From this he suggested the modification of $m_k = 5 + (6 + \mu_2)/k$ which satisfied this rule. Aboav's original equation then became a special case when $\mu_2 = 2$, which is close to the expected value for a random collection of Voronoi polygons (see section **??**). Aboav then proposed that if a generic form of $m_k = A + B/k$ was used in conjunction with Weaire's sum rule then

$$m_k = A + \frac{\mu_2 + \langle k \rangle^2 - A\langle k \rangle}{k} \,. \tag{2.19}$$

This is now more commonly expressed in the linear form [65]:

$$k m_k = \mu_2 + \langle k \rangle^2 + \langle k \rangle (1 - \alpha)(k - \langle k \rangle) \,. \tag{2.20}$$

Equation 2.20 is known as the Aboav-Weaire law and relates the mean ring size about a given central ring to a single fitting parameter, α. The value of α describes the strength of the ring correlations, with a larger positive value indicating a greater tendency for small-large ring adjacencies. More specifically, the random limit can be deduced by evaluating $\frac{\partial m_k}{\partial x} = 0$ as [66]:

$$\alpha = -\frac{\mu_2}{\langle k \rangle^2} \,. \tag{2.21}$$

Hence all systems with $\alpha > -\mu_2/\langle k \rangle^2$ have more small-large ring adjacencies than would be expected from chance whilst conversely those with $\alpha < -\mu_2/\langle k \rangle^2$ have more small-small and large-large pairings.

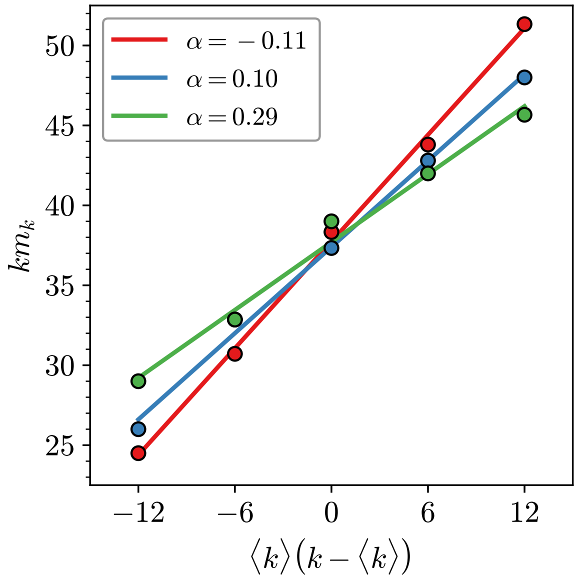


Figure 2.5: Calculation of an Aboav-Weaire fit for three configurations (shown in figure 1.1(b)-(d)). The value of the α parameter quantifies the tendency of small rings to be adjacent to large rings, with a larger value indicating stronger small-large ring correlations.

Despite the Aboav-Weaire law being purely empirical and there being no topological requirement for m_k to vary systematically k, the law does seem to hold well for a diverse set of physical systems. The law is well used for example in studies of materials, emulsions, biological tissues as well as in planetary science [28, 67–70]. As an example of the calculation of the Aboav-Weaire parameter, the plots of the fits for the systems in figure 1.1 are presented in figure 2.5, along with the corresponding α parameters. This demonstrates two contrasting aspects of the Aboav-Weaire law. Firstly the law holds very well, especially given the fact that these samples consist of just twenty rings each. However, it also demonstrates that the law is by no means exact and that some greyness is inevitably introduced during the linear regression.

3 | Computational Methods

The theoretical basis of Monte Carlo methods and their application to generating realisations of two-dimensional networks is reviewed. There is a broad discussion of Metropolis Monte Carlo methods, before specific methods are covered in detail; namely the bond switching algorithm and hard particle Monte Carlo in conjunction with the Voronoi construction. This discussion lays the groundwork for the extension of these methods and development of additional Monte Carlo algorithms in subsequent chapters.

3.1 General Monte Carlo Methods

Monte Carlo methods are a class of computational algorithms designed to solve complex problems stochastically. These normally fall into the broad categories of calculating integrals, sampling probability distributions and finding global minima of very high dimensional functions - tasks which are often incredibly hard to compute deterministically. Since their initial development in the mid-20$^{\text{th}}$ century, such methods have become an invaluable tool for solving problems in the physical sciences. Monte Carlo methods are used in this context for calculating thermodynamic averages of properties in equilibrium systems; finding the minima in potential energy surfaces of small molecules, glasses, crystals and biomolecules; as well as non-equilibrium simulations such as growth of crystals and thin-films [71–76]. In this thesis these Monte Carlo methods will be used in a variety of contexts: chapter 4 simulates the growth of bilayer materials using a random sequential growth algorithm; chapter 5 optimises a cost function to control network structure; chapter 6 samples amorphous configurations of various system topologies; chapter 7 samples the phase space of hard particle assemblies and chapter 8 finds solutions to the procrystalline problem, essentially via global optimisation. Therefore, the general theory is presented here with specific details of two established methods: bond switching and hard particle Monte Carlo given in the following section.

3.1.1 Statistical Mechanics

The total energy of a system with a fixed number of particles, $\mathcal{N}$, is given by the Hamiltonian,

$$\mathcal{H}(\mathbf{p},\mathbf{r}) = \mathcal{K}(\mathbf{p}) + \mathcal{U}(\mathbf{r}) \ , \tag{3.1}$$

where $\mathcal{K}(\mathbf{p})$ is the kinetic energy as a function of all particle momenta and $\mathcal{U}(\mathbf{r})$ is the potential energy as a function of all particle positions [77]. The positions and momenta comprise the phase space of the system. At fixed volume, $\mathcal{V}$, and temperature, T, all the the essential thermodynamic information is then provided through the classical canonical partition function:

$$Q = \frac{1}{h^{D\mathcal{N}}\mathcal{N}!} \int \mathrm{d}\mathbf{p}\, \mathrm{d}\mathbf{r}\, \exp\left[-\mathcal{H}(\mathbf{p},\mathbf{r})/k_{\mathrm{B}}T\right], \tag{3.2}$$

where D is the number of spatial dimensions. This can be factorised into kinetic and potential components as

$$Q = \frac{1}{h^{D\mathcal{N}}\mathcal{N}!} \int \mathrm{d}\mathbf{p}\, \exp\left[-\mathcal{K}(\mathbf{p})/k_{\mathrm{B}}T\right] \int \mathrm{d}\mathbf{r}\, \exp\left[-\mathcal{U}(\mathbf{r})/k_{\mathrm{B}}T\right], \tag{3.3}$$

where

$$Z = \int \mathrm{d}\mathbf{r}\, \exp\left[-\mathcal{U}(\mathbf{r})/k_{\mathrm{B}}T\right] \tag{3.4}$$

is the configurational integral [78]. As will be shown, in Monte Carlo simulations it is the energetic differences between configurations that are required, and so at constant temperature the kinetic component can be neglected and it is only the configurational integral that is of importance. In this case the probability density of the system being in the configuration $\mathbf{r}$ is given by the Boltzmann distribution:

$$P(\mathbf{r}) = \frac{\exp\left[-\mathcal{U}(\mathbf{r})/k_{\mathrm{B}}T\right]}{Z} \ . \tag{3.5}$$

This allows the expectation value of an observable of the system, $\mathcal{A}(\mathbf{r})$, to be determined from:

$$\langle A \rangle = \int \mathrm{d}\mathbf{r}\, \mathcal{A}(\mathbf{r})\, P(\mathbf{r}) \ . \tag{3.6}$$

The expectation value is then the ratio of two $\mathcal{N}D$ dimensional integrals. The next section shows how these can be evaluated by Monte Carlo sampling.

3.1.2 Importance Sampling

An integral of form (3.6) can be evaluated numerically by a number of methods. As an illustration, consider the simple example of a two-dimensional potential energy surface in figure 3.1. To calculate the expectation value of the potential energy one must evaluate the integral

$$\langle \mathcal{U} \rangle = \int_0^{L_y} \int_0^{L_x} \mathrm{d}x \mathrm{d}y\, \mathcal{U}(x, y)\, \mathcal{P}(x, y) \ . \tag{3.7}$$

One way to achieve this would be to use standard numerical methods such as the trapezium rule or Simpson's rule to calculate the potential energy over a regular grid of points, as in figure 3.1a, weighting each according to the Boltzmann distribution.

An alternative would be to take a stochastic approach. In the simplest implementation, a series of S random sampling points, (x_i, y_i), can be generated uniformly in the intervals $[0, L_x]$ and $[0, L_y]$, as in figure 3.1b. Weighting these according to the Boltzmann distribution and averaging gives an estimation to the integral:

$$\langle \mathcal{U} \rangle = \frac{L_x L_y}{S} \sum_{i=1}^{S} \mathcal{U}(x_i, y_i)\, \mathcal{P}(x_i, y_i) \ , \tag{3.8}$$

which converges to the exact value as $S \to \infty$.

However, both quadrature and Monte Carlo uniform sampling suffer from the same inefficiency. As can be seen in both schemes, many of the sampling points fall in regions of phase space where the potential energy is high and hence the weighting probability distribution is very small at reasonable temperatures. In effect, significant effort is spent calculating regions where the contribution to the total integral is negligible. A better approach is therefore to generate a series of S random sampling points, (x_i, y_i), according to the distribution $\mathcal{P}(x, y)$, as in figure 3.1b. The expectation value of the observable can then be calculated using a simple average:

$$\langle \mathcal{U} \rangle = \frac{1}{S} \sum_{i=1}^{S} \mathcal{U}(x_i, y_i) \; . \tag{3.9}$$

This is known as importance sampling and is vastly more efficient when dealing with an aggressive probability distribution like the Boltzmann, where only a small proportion of the phase space is accessible.

Whilst this scheme is ideal theoretically, it is impracticable for physical systems. This is because for any problem of real interest one lives in a “black box” where the functional form of the potential energy surface in its hundreds if not thousands of dimensions is unknown. In this case often the only way of learning about the form is by on-the-fly exploration of the surface [79]. This can be achieved by talking a random walk through configurational space using Markov chain Monte Carlo.

3.1.3 Markov Chain Monte Carlo

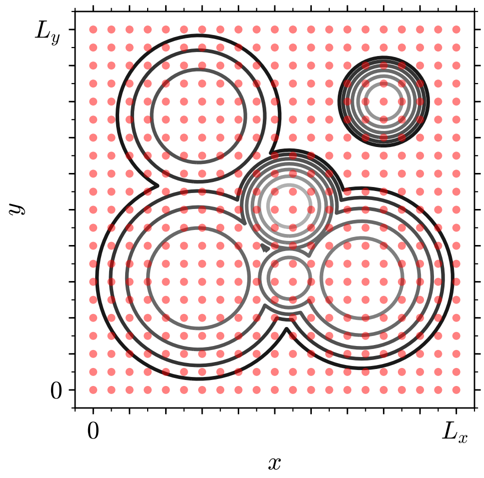
Markov chain Monte Carlo provides a framework to perform importance sampling on a potential energy surface. A system of interest can exist in a (very large) number of configurational states, $\{\mathbf{r}_0, \mathbf{r}_1, \ldots, \mathbf{r}_M\}$. A Markov chain can then be constructed from this set, whereby a sequence of states is generated stochastically across a series of steps, $s = 0, 1, \ldots, S$. In this process, the probability of moving between states at each step is given by the transition matrix, $\boldsymbol{\pi}$, where each element, π_{ij}, gives the probability of moving from the state $\mathbf{r}_i$ to another state $\mathbf{r}_j$. This leads to the two relationships:

$$0 \leq \pi_{ij} \leq 1 \, , \tag{3.10}$$

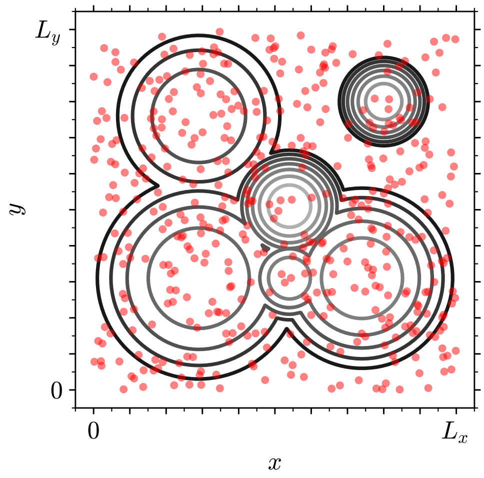
$$\sum_j \pi_{ij} = 1 \, , \tag{3.11}$$

the first being a statement of the probabilistic nature of the elements whilst the second ensures all transfer remains within the state space [77–79].

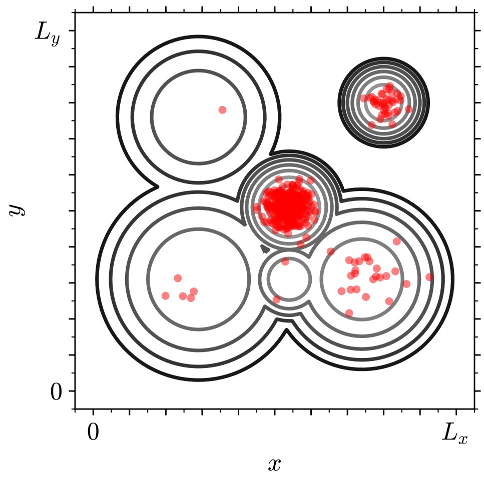
The probability that the system is in each state at a given step, s, can be represented by the row vector $\mathbf{P}_s$. This probability distribution evolves with each



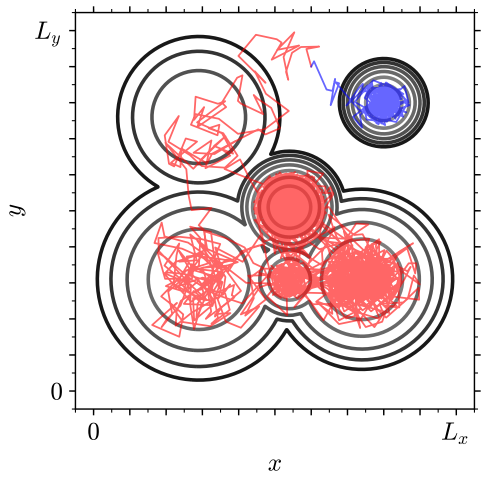
(a) Quadrature.



(b) Uniform Sampling.



(c) Importance Sampling.



(d) Markov Chain Monte Carlo

Figure 3.1: Demonstration of different sampling methods with an example two-dimensional potential energy surface (contour lines). Panels(a)-(c) display the same number of (red) sampling points. Panel (a) shows conventional quadrature where the surface is divided into a regular grid of sampling points which are then weighted by the Boltzmann distribution. Panel (b) shows Monte Carlo sampling with a uniform distribution of points which again must be Boltzmann-weighted. Panel (c) shows Monte Carlo importance sampling with points now selected according to the Boltzmann distribution. Panel (d) shows Markov chain Monte Carlo with two random walks through phase space (red and blue lines) starting from different random seeds.

step as $\mathbf{P}_{s+1} = \mathbf{P}_s \boldsymbol{\pi}$, so that starting from any initial distribution, $\mathbf{P}_0$, it follows that $\mathbf{P}_S = \mathbf{P}_0 \boldsymbol{\pi}^S$. The question is then as to the behaviour as $S \to \infty$. Provided certain criteria are met, the distribution will tend to a stationary distribution,

$\mathbf{P}$, which satisfies the eigenvalue equation

$$\mathbf{P} = \mathbf{P}\boldsymbol{\pi}\,, \tag{3.12}$$

regardless of the initial distribution (although the speed of the convergence does depend on $\mathbf{P}_0$). This will occur only if the system is *ergodic*, meaning that every state is connected to every other by some finite path.

In a discrete analogue to equation (3.6), the expectation value of an observable, A, can be calculated from the ensemble average:

$$\langle A \rangle = \sum_{i=1}^{M} A\left(\mathbf{r}_i\right) \mathcal{P}\left(\mathbf{r}_i\right)\,, \tag{3.13}$$

where $\mathcal{P}\left(\mathbf{r}_i\right)$ are the elements of $\mathbf{P}$. However, as previously mentioned the number of discrete states is usually exceedingly large and so calculating the average over all states is not possible. The solution is to take a random walk across through configurational space, sampling explicit states to form the chain $X_0, X_1, \ldots, X_S$; where each move is chosen randomly according to the transition matrix $\boldsymbol{\pi}$. In this case the expectation of the same observable can be calculated from the average over the sampled states:

$$\langle A \rangle = \frac{1}{S} \sum_{i=1}^{S} A\left(X_i\right)\,, \tag{3.14}$$

where the true value is approached as $S \to \infty$.

In this section the problem of sampling phase space efficiently has been reformulated, but as yet not solved. This is because the form of the transition matrix is still unknown. Instead only the ideal form of the limiting probability distribution, $\mathbf{P}$, is available - where the elements follow the Boltzmann probabilities in equation (3.5). A practical solution to this problem is provided by the Metropolis algorithm.

3.1.4 Metropolis Algorithm

The Metropolis algorithm gives a prescription of how to construct a transition matrix, $\boldsymbol{\pi}$, with the requisite properties that samples the Boltzmann distribution

[80]. Firstly, combining equations (3.11) and (3.12) gives a condition on the transition matrix known as global balance:

$$\sum_j \mathcal{P}\left(\mathbf{r}_i\right) \pi_{ij} = \sum_j \mathcal{P}\left(\mathbf{r}_j\right) \pi_{ji} \,. \tag{3.15}$$

Whilst it is possible to construct transition matrices which satisfy only global balance [81–83], it is practically simpler to satisfy global balance by applying the stronger condition of detailed balance:

$$\mathcal{P}\left(\mathbf{r}_i\right) \pi_{ij} = \mathcal{P}\left(\mathbf{r}_j\right) \pi_{ji} \,. \tag{3.16}$$

In the Metropolis algorithm the off-diagonal elements of the transition matrix are written as the product of two probabilities:

$$\pi_{ij} = \begin{cases} \tau_{ij} P_{ij} & i \neq j \\ 1 - \sum\limits_{j \neq i} \tau_{ij} P_{ij} & i = j \end{cases} , \tag{3.17}$$

where τ_{ij} is the trial probability of moving from state $\mathbf{r}_i$ to $\mathbf{r}_j$ and P_{ij} is the probability of accepting the trial move. To conform to detailed balance, the trial probabilities must be chosen to satisfy $\tau_{ij} = \tau_{ji}$. Then, in the crux of the algorithm, the acceptance probabilities are given by

$$P_{ij} = \begin{cases} 1 & \mathcal{P}\left(\mathbf{r}_j\right) \geq \mathcal{P}\left(\mathbf{r}_i\right) \\ \frac{\mathcal{P}(\mathbf{r}_j)}{\mathcal{P}(\mathbf{r}_i)} & \mathcal{P}\left(\mathbf{r}_j\right) < \mathcal{P}\left(\mathbf{r}_i\right) \end{cases} = \begin{cases} 1 & \mathcal{U}\left(\mathbf{r}_j\right) \leq \mathcal{U}\left(\mathbf{r}_i\right) \\ \frac{\exp[-\mathcal{U}(\mathbf{r}_j)/k_\mathrm{B}T]}{\exp[-\mathcal{U}(\mathbf{r}_i)/k_\mathrm{B}T]} & \mathcal{U}\left(\mathbf{r}_j\right) > \mathcal{U}\left(\mathbf{r}_i\right) \end{cases} , \tag{3.18}$$

which can be expressed more succinctly as

$$P_{ij} = \min\Big[1, \exp\left[-\Delta\mathcal{U}/k_\mathrm{B}T\right]\Big] , \tag{3.19}$$

where $\Delta\mathcal{U}$ is the difference in potential energy between the final and initial states. The elegance of the Metropolis algorithm lies in the fact that the acceptance probability depends only on the ratio of the configuration probabilities removing the need for a normalising factor. This means the relative probabilities can be used (which are computable) instead of the absolute probabilities (which are unknowable).

The final stage is the choice of the matrix of trial probabilities, $\boldsymbol{\tau}$. This is very flexible and one can be creative in the selection of trial moves, providing that the underlying matrix is symmetric and ergodic. An effective strategy is to choose moves in which the trial state is relatively close to the current state to trace the paths of high probability in the system. A summary of the Metropolis algorithm is therefore as follows:

1. Initialise the system in a state $X_{s=0}$ and calculate the potential energy $\mathcal{U}\left(X_s\right)$
2. Generate a trial state X_t (a perturbation of X_s) according to τ_{st}
3. Calculate the potential energy of the trial state $\mathcal{U}\left(X_t\right)$
4. Determine acceptance or rejection of the trial move according to the Metropolis criterion (3.19)
5. Update the system to the new state: if the trial move is accepted $X_{s+1} = X_t$ otherwise $X_{s+1} = X_s$
6. Repeat steps 2-5

There are a few practical factors related to the scheme above. In Markov chain Monte Carlo it was previously mentioned that it takes time for the system to evolve to the stationary distribution. Therefore it is necessary to have an equilibration period where the chain is generated but not used for sampling of observables. In addition, whilst selecting trial moves close to the current state increases efficiency, it introduces correlation into the procedure. A way around this is to not calculate observables based on every step, but rather after a number of statistically significant steps.

As an example of the Metropolis algorithm, consider again the two-dimensional potential energy surface in figure 3.1d. Here two simulation paths are displayed in red and blue, starting from the same initial state but with different starting points in the random number generators *i.e.* random seeds. As can be seen the Metropolis algorithm takes a random walk over the configurational space, conducting importance sampling as in 3.1c. However, in this example highlights a potential problem. There are two regions of phase space with non-zero probabilities which are separated by a relatively large energy barrier. Although they are in principle linked by a path, the barrier may effectively mean they are disconnected on a reasonable simulation time scale, breaking ergodicity. This manifests as the red walk sampling one region and the blue walk being trapped in the other region. Using multiple seeds in this way helps to identify if any such behaviour is present. If it leads to significant differences in the computed averages, more advanced techniques using enhanced sampling may have to be employed [84, 85].

3.1.5 Global Optimisation & Simulated Annealing

So far in this section it has been shown how Monte Carlo methods can be used perform importance sampling of potential energy surfaces. These methods can also be used to solve the related problem of finding global minima in potential energy surfaces and other more general functions. Consider the case where there is an objective function, $\Omega(\mathbf{r})$, which depends on particle positions. If it is known that there exists a solution where $\Omega(\mathbf{r}) = 0$, it may be sufficient to perform a standard random walk of the type in figure 3.1d until a solution is found, using the more general Metropolis criterion:

$$P_{ij} = \min\left[1, \exp\left[-\Delta\Omega/k_{\mathrm{B}}T\right]\right]. \tag{3.20}$$

There is of course a chance that the optimisation will not converge to the global minimum, most likely getting trapped in a local minimum (as for instance the blue path in 3.1d). One solution to this problem is just to keep restarting the algorithm with different initial conditions until the global minimum is obtained.

Often however the value of the global minimum is not known, as is the case for a potential energy surface, and this rudimentary approach is insufficient. One must then employ a more sophisticated technique to find the global minimum of a very high dimensional and potentially rough surface. This in itself is an extensive area of study and there are many approaches such as using genetic algorithms or basin-hopping [86–88]. This thesis will use simulated annealing, which can be considered an extension to Metropolis Monte Carlo [89]. In addition simulated annealing is effective for searching surfaces with many similar minima as in glasses - the name reflecting its origins in the analogous process in metallurgy to generate defect free metals.

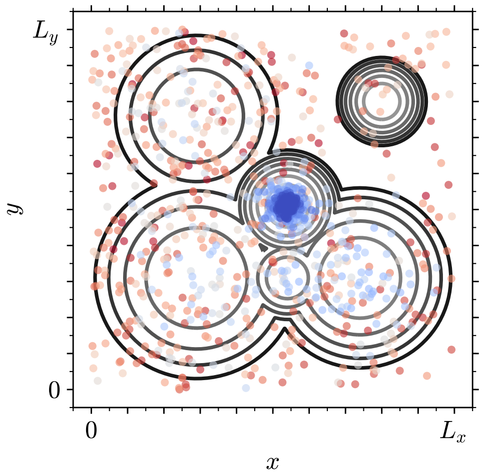


Figure 3.2: Demonstration of the simulated annealing algorithm on a two-dimensional potential energy surface, with states coloured by temperature (red→blue indicating hot→cold. As the temperature is reduced the state converges on the global minimum.

The simulated annealing algorithm proceeds as follows. The system of interest is first thermalised by performing Metropolis Monte Carlo at infinite temperature *i.e.* accepting every move. The system is then gradually cooled to zero temperature, with the Metropolis criterion (3.20) reducing the proportion of accepted moves. In theory if the cooling is infinitely slow, the system is maintained in thermal equilibrium and will eventually reach the global minimum [90]. In practice this is not realisable and so a cooling rate must be empirically selected. Still it is possible for trapping to occur in local minima, especially if the transition between low energy states is very slow. As before, one can then cycle the simulated annealing, repeatedly heating and cooling the system until the global minimum is found. The simulated annealing algorithm is demonstrated with the two-dimensional potential energy surface in figure 3.2. As can be seen at high temperature the entire surface is sampled, overcoming all energy barriers, but as cooling takes place the system settles into the low energy regions of the surface, finally terminating in the global minimum.

3.2 Bond Switching Monte Carlo

Bond switching Monte Carlo was originally developed by Wooten, Winer and Weaire to generate high quality configurations of three dimensional silica glass

[91]. The basic principle is to amorphise a crystalline lattice with a series of transformations that swap the nearest neighbours of pairs of atoms and optimise the resulting structure to generate a continuous random network which is well-relaxed. These continuous random network models replicate experimental observables with high accuracy (including bond length and angle distributions, radial distribution functions, electronic band gaps and Raman spectra) and have since been applied to alternative systems such as three-dimensional amorphous carbon, binary glasses and biological polymers [92–97]. However, the method can also be readily modified to study two-dimensional systems, as has been done for amorphous graphene and silica, and which forms the basis for much of the work in this thesis, particularly in chapters 5 and 6 [29, 98]. The basic algorithmic details are described in this section, with extensions given in sections again ref later .

3.2.1 Algorithmic Details

The two-dimensional bond switching algorithm essentially follows the prescription of simulated annealing in section 3.1.5. A skeleton algorithm structure is outlined below, followed by specific details [99]. Visualisations are provided for reference in figure 3.3.

1. Generate initial crystalline hexagonal lattice
2. Thermalise the lattice with a large number of random moves
3. Sample configurations by annealling the system slowly at finite temperature, accepting moves according to the Metropolis criterion 3.19

The Monte Carlo move for 3-coordinate atomic materials is essentially the introduction of a Stone-Wales defect into the lattice, which augments the size of two rings and decrements two others, preserving both the mean ring size and the coordination number of the individual atoms involved in the transformation [14]. As defects become more concentrated they overlap, leading to increasing diversity into the ring structure (allowing access to more than the pentagons and heptagons in a single Stone-Wales defect). Each bond transposition is followed by

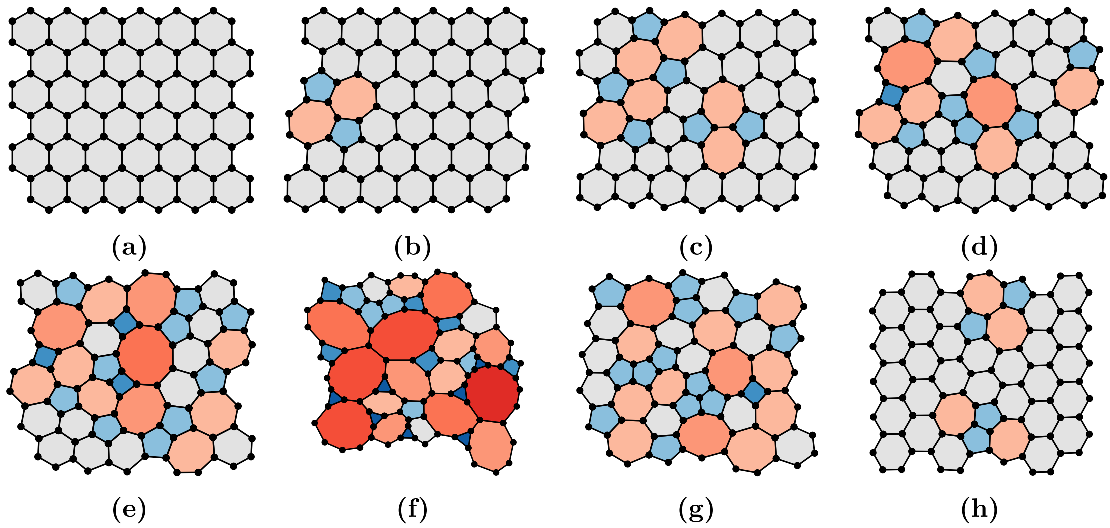


Figure 3.3: Configurations taken from stages of the two-dimensional bond switching algorithm. A crystalline lattice (a) is first thermalised to generate a random high energy network (f) by sequential overlapping Stone-Wales defects (b)-(e). Sampling then occurs as the system is slowly annealed (g)-(h), allowing access to defect states that are not initially obtainable from the crystal structure.

geometry optimisation to minimise and calculate the total energy of the system. A key aspect in the bond switching algorithm is therefore the choice of potential model. The potential models and geometry optimisation process used in this thesis can be found in subsections below.

Cooling the system slowly ensures that the material remains in thermodynamic equilibrium, allowing configurations to be sampled throughout the simulation. The ring structure of the system is then related to the temperature parameter, with more extreme ring sizes appearing at higher temperatures (compare figure 3.3f-3.3h). This simply reflects the inherent balance of enthalpic *vs.* entropic considerations. Figure 3.3h also demonstrates the importance of cooling a randomised lattice instead of heating a crystal, as some low energy defects may have a multi-step formation with a high energy barrier.

3.2.2 Potential Models

The nature of the bond switching method lends itself to the use of semi-empirical potentials which have explicit stretching and angular neighbour lists. As such a popular choice for materials modelling is the Keating potential and modifications thereof [100, 101]. For a two-dimensional system the Keating potential has the form:

$$\mathcal{U}\left(\mathbf{r}\right) = \frac{3}{16}\frac{K_S}{r_0^2}\sum_{\substack{i,j\in\\ \text{stretches}}}\left(r_{ij}^2 - r_0^2\right)^2 + \frac{3}{8}\frac{K_A}{r_0^2}\sum_{\substack{ijk\in\\ \text{angles}}}\left(r_{ij}r_{ik}\cos\theta_{ijk} - r_0^2\cos\theta_0\right)^2 , \quad (3.21)$$

where r_{ij} the distance and θ_{ijk} the angle between particles; whilst K_S and K_A are the force constants for the stretching and angular terms respectively [99]. This potential drives the system towards equilibrium values of r_0 for the bond lengths and θ_0 for the bond angles. The Keating potential has been parametrised for a range of specific materials [99, 102].

However, a more generic potential model is sometimes required which captures the same essential physics. This is provided through the simplified Keating potential [103],

$$\mathcal{U}\left(\mathbf{r}\right) = \frac{K_S}{2}\sum_{\substack{i,j\in\\ \text{stretches}}}\left(r_{ij} - r_0\right)^2 + \frac{K_A}{2}\sum_{\substack{i,j,k\in\\ \text{angles}}}\left(\cos\theta_{ijk} - \cos\theta_0\right)^2 , \quad (3.22)$$

which is harmonic in stretching and angular terms. One final modification can be made to this potential. Sometimes it is informative build models which enforce ring convexity *i.e.* maintain all angles within the range $0 \leq \theta_{ijk} \leq \pi$. This can be achieved by augmenting the simplified Keating potential with a restricted bending (ReB) potential [104]:

$$\mathcal{U}\left(\mathbf{r}\right) = \frac{K_S}{2}\sum_{\substack{i,j\in\\ \text{stretches}}}\left(r_{ij} - r_0\right)^2 + \frac{K_A}{2}\sum_{\substack{i,j,k\in\\ \text{angles}}}\frac{\left(\cos\theta_{ijk} - \cos\theta_0\right)^2}{\sin^2\theta_{ijk}} . \quad (3.23)$$

The addition of the sine term in denominator causes the potential to diverge as bond angles approach linearity, preventing bonds from “inverting”.

3.2.3 Geometry Optimisation

The purpose of geometry optimisation is to minimise the overall potential energy of a network, $\mathcal{U}\left(\mathbf{r}\right)$, as a function of all atomic positions, $\mathbf{r}$, after they have been perturbed *e.g.* by a bond transposition. As all initial configurations are well relaxed and perturbations relatively small, this can be achieved with a local minimisation routine. In addition as the potential models in this work are smooth and harmonic, a straightforward steepest descent algorithm is both sufficient and efficient.

The steepest descent algorithm is an iterative method which searches down the potential energy gradient until a minimum is reached [105]. It has the following scheme:

1. Calculate the potential energy of the system $\mathcal{U}_i = \mathcal{U}\left(\mathbf{r}_i\right)$
2. Determine the negative gradient of the potential *i.e.* the forces acting on the particles $\mathbf{F}_i = -\nabla\mathcal{U}_i$
3. Find the optimal distance to displace the particles along the lines of force $\mathcal{U}_{i+1} = \min\left[\mathcal{U}\left(\mathbf{r}_i + \lambda\mathbf{F}_i\right)\right]$
4. Set $\mathbf{r}_{i+1} = \mathbf{r}_i + \lambda_{\min}\mathbf{F}_i$
5. Evaluate convergence and repeat steps 1-4 if $|\mathcal{U}_{i+1} - \mathcal{U}_i| > \gamma$

The calculation of forces in stage 2 will depend on the potential model used, details of which are given in appendix **??**. Note that stage 3 also requires a minimisation routine, which may seem counter-intuitive. However, this is a one-dimensional minimisation which trivial to estimate with a line search method appendix? . The tightness of the convergence condition is set through the parameter γ.

One final performance improvement arises from the fact that the Monte Carlo are inherently local. Therefore geometry optimisation can be employed such that only the atoms in the immediate vicinity of the switching move need to be minimised to obtain an accurate structure. Typically this would extend to all atoms within five coordination shells of those directly involved in the switch move [106].

3.3 Hard Particle Monte Carlo

Hard particle Monte Carlo is one of the most well-established computational methods in statistical physics. Through its simplicity it is able to provide insight into the fundamental behaviour of particle systems and simulations of increasing size are still performed this century [107–110]. In this thesis it will be used to generate ring systems in the form of Voronoi tessellations (see section **??**), in analogy to experimental colloidal systems [50].

3.3.1 Hard Particle Model

Hard particle models are applicable over a range of dimensions. In two dimensions the system consists of an arrangement of hard disks and in three dimensions hard spheres. One can also take a quasi two-dimensional system, which comprises hard spheres confined to a plane. Regardless of the dimensionality, the central principle is that no two particles in the system can have any degree overlap. Formally, if the particle radii are denoted by R_i and the distance between any pair of particle centres by r_{ij}, the pair potential is:

$$\mathcal{U}_{ij} = \begin{cases} \infty & r_{ij} < R_i + R_j \\ 0 & \text{otherwise} \end{cases} . \tag{3.24}$$

As the total energy is simply then

$$\mathcal{U}\left(\mathbf{r}\right) = \sum_{i<j} \mathcal{U}_{ij} , \tag{3.25}$$

it follows that if any pair of particles have overlap the system energy is infinite and the Boltzmann weighting is zero. Hard particle models are typically quantified in terms of the packing fraction, ϕ, which in two dimensions has the form

$$\phi_{2D} = \rho\pi\langle R^2\rangle , \tag{3.26}$$

where $\rho = \mathcal{N}/V$, the number density.

3.3.2 Algorithmic Details

Hard particle systems can be simulated using the Metropolis algorithm outlined in section 3.1.4. The system is initialised by selecting a random non-overlapping configuration. This can be achieved easily for low to medium densities by a greedy algorithm like random sequential addition, where particles are added successively in a manner which does not overlap with any previous particles [111]. For higher packing fractions a more sophisticated algorithm is needed Find refs .

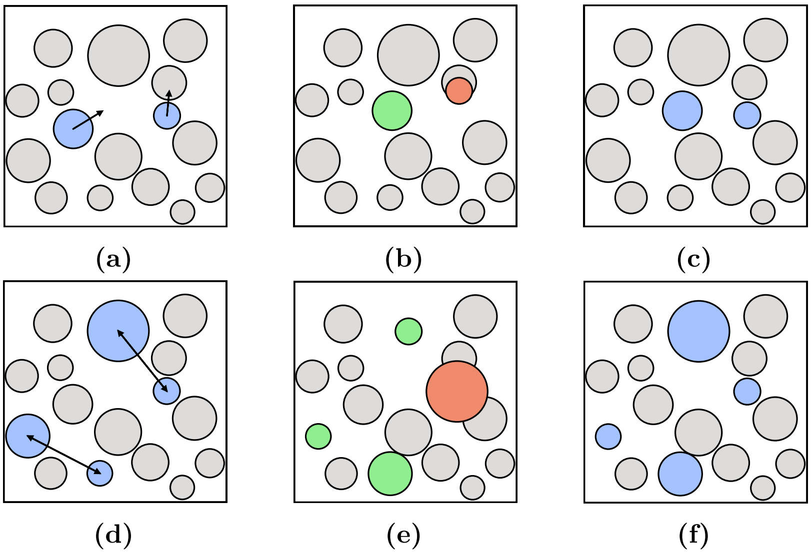


Figure 3.4: Demonstration of two displacement (a)-(c) and two swap (d)-(f) moves in hard particle Monte Carlo. In displacement moves, particles are randomly selected and assigned a trial random displacement vector (a). In swap moves, two particles are randomly selected and their radii trial swapped (b). The trial move is then examined to see if it introduces any particle overlaps (b),(e). If there are no overlaps (green), then the trial move is accepted and the system updated but otherwise (red) the move is rejected and the system returns to the previous state (c),(f).

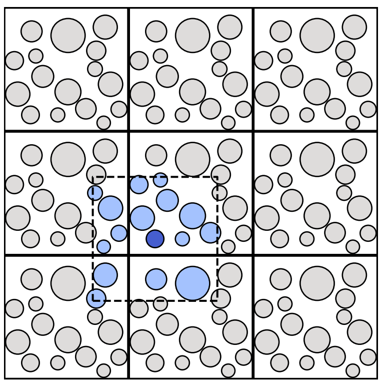


Figure 3.5: Simulation of bulk system is achieved using periodic boundary conditions, where a central cell is surrounded by repeated images of itself. A particle of interest (dark blue) then interacts with the nearest images of every other particle (light blue).

Once the initial configuration has been generated, it is evolved via two Monte Carlo moves. The first is the displacement move, whereby a random particle is selected and translated according to a random vector with elements generated uniformly in the range $[-\delta, \delta]$. If the displacement introduces any particle overlaps it is rejected, otherwise the system is updated to the new configuration, as illustrated in figure 3.4a-3.4c. The value of δ is chosen for each simulation such that the proportion of accepted moves is $\sim 50\%$, allowing for efficient searching of configurational space. The optimal value can be determined by continuous adjustment during equilibration.

The second is the swap move, where two random particles are selected their radii exchanged [112, 113]. Once again a swap move is only accepted if it does not lead to any overlapping particles and is demonstrated in figure 3.4d-3.4f. The swap move is used to increase the efficiency in simulations of polydisperse particles and is an example of how the design of Monte Carlo moves can be flexible and they do not have to have a direct physical basis. The swap move is attempted for every ten displacement moves.

Finally, to remove the presence of an interface in the system, simulation is performed with periodic boundary conditions. In this scheme the central simulation cell is repeated to form an infinite lattice, so that every particle experiences a bulk environment. Coupled with this is the use of the minimum image convention, where each particle then only interacts with the nearest repeated image of all the remaining particles. This is illustrated in figure 3.5.

3.3.3 Voronoi Construction

The hard particle configurations produced by Monte Carlo simulations are not in themselves network structures, rather simply a collection of correlated points. The network structure is revealed by construction of a Voronoi diagram, which partitions the sample into a system of tessellating cells, where each cell encapsulates all the space closest to the associated particle [114]. A two-dimensional Voronoi diagram is formed through the placement of dividing lines between the centres of neighbouring particles. The intersection of these lines forms the characteristic tessellating polygons.

In the simplest unweighted approach, the dividing line between two neighbouring particles separated by the Euclidean distance r_{ij}, is simply located midway between the particles at a distance $r_{ij}/2$. The elegance of the unweighted Voronoi diagram is that only the particle centres are required for its construction, with no requirement

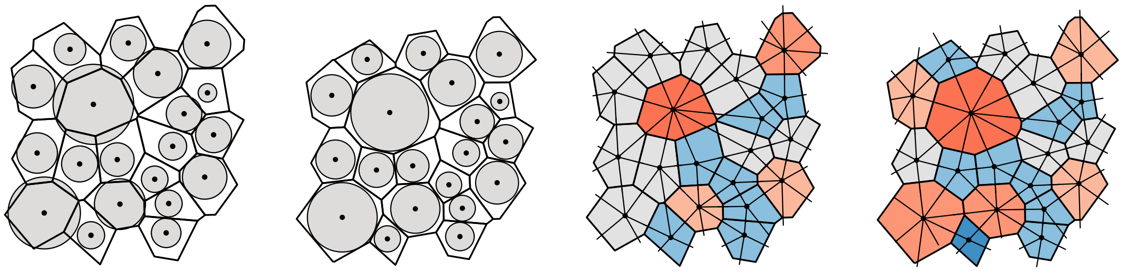


Figure 3.6: Voronoi construction of a polydisperse hard disk system. Panels (a) and (b) compare the unweighted and weighted (radical) Voronoi tessellations respectively. The radical Voronoi assigns more volume to the larger particles to ensure a more equitable distribution of space, which can affect the underlying ring structure, shown in panels (c) and (d). The dual network, known as the Delaunay triangulation, is also overlaid.

for a cut-off parameter. Whilst the unweighted Voronoi tessellation is very effective for studying monodisperse particles, there are some limitations for polydisperse species. Specifically, the Voronoi partition underestimates the space assigned to large particles and overestimates that for small particles - a simple reflection of the lack of information on particle radii (see figure 3.6a). To rectify this, weighted modifications have been suggested which take account of the differences in radii [115].

To construct a weighted Voronoi diagram, one simply adjusts the position of the dividing line, such that it is further from the particle with the greater weight. The weighting method used in this work is the so called radical tessellation, or power diagram [116, 117]. In this modification, the dividing line is placed a distance r_i from particle i, given by:

$$r_i = \frac{w_i^2 - w_j^2 + r_{ij}^2}{2r_{ij}}, \tag{3.27}$$

where w_i and w_j are the weights for each particle. The benefit of this method is that is adjusts the partitioning of space so that greater volume is assigned to the particles with larger weight, and is well designed so that all of the sample space remains accounted for - unlike some alternative constructions [118]. In terms of the particle weights, the logical choice is simply the disk radii. This is because at the contact distance, $r_{ij} = R_i + R_j$, equation (3.27) shows that $r_i = R_i$ *i.e.* the radical dividing line sits exactly between the two disks, producing the most equitable distribution

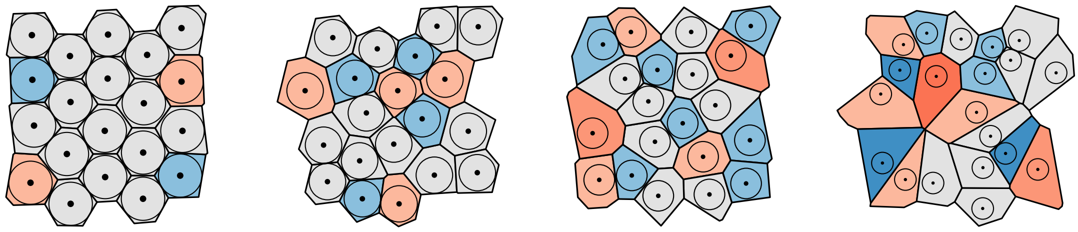

(a) $\phi = 0.8$ (b) $\phi = 0.6$ (c) $\phi = 0.4$ (d) $\phi = 0.2$

Figure 3.7: The ring structure in Voronoi diagrams is controlled through the packing fraction, ϕ, of the underlying hard particle system. Ring diversity increases as packing fraction is lowered from $0.8 \to 0.2$ in (a)-(d).

of volume (see figure 3.6b). Furthermore, when the radii are equal, $r_i = r_{ij}/2$ and the result from the standard unweighted Voronoi is regenerated as expected. It is worth noting here that the weighting method can affect the ring sizes (*i.e.* number of vertices) as well as the ring areas, as demonstrated in figures 3.6c,3.6d.

The outcome of the Voronoi construction is a system of percolating rings not dissimilar to those seen in materials. The dual network, known as the Delaunay triangulation, is also obtained, which defines the nearest neighbours for each particle. The main difference with atomic materials is that the polygon edge lengths and angles are not constrained by a potential model the ring structure is therefore completely entropically controlled. The degree of disorder is then determined by the packing fraction, ϕ, where decreasing the packing fraction leads to increased diversity in the ring statistics, as illustrated in figure 3.7. As can be seen there are some defects which are analogous to those seen in materials, such as the Stone-Wales defect in figure 3.7b, but others are not, as in figure 3.7a which arise from very small perturbations in the crystalline lattice. The limiting value as $\phi \to 0$ is well studied as the Poisson Voronoi diagram [119, 120]. This corresponds to the Voronoi diagram formed from a random uniform array of points. In this way Voronoi systems provide a good complement to compare and contrast with materials.

cite [121] somewhere

3.4 Analysis Methods

3.4.1 Bond Length and Angle Distributions

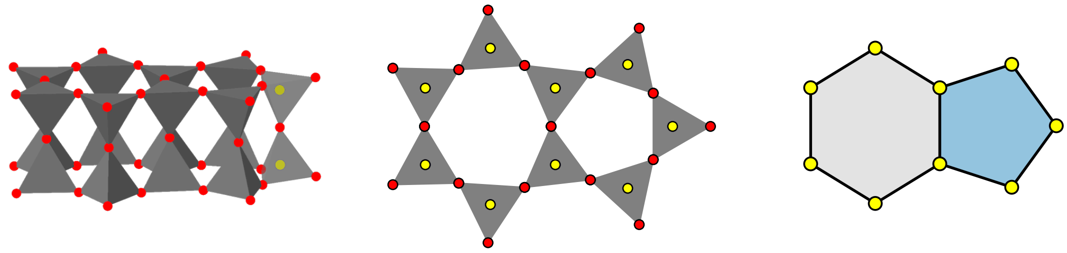
3.4.2 Radial Distribution Functions

4 | Modelling Bilayer Materials

A computationally tractable Monte Carlo method using triangle rafts is developed to generate configurations for bilayers of SiO_2 and related materials. The method allows defect free networks of any given shape to be grown with both tuneable ring statistics and topologies, controlled by a combination of the choice of the "allowed" rings and the effective growth "temperature". Configurations are generated with Aboav-Weaire parameters commensurate with those obtained from an analysis of experimental configurations, improving significantly on previous methods for generating these networks (which systematically underestimate this parameter). The ability to efficiently grow configurations allows exploration of the structural basis of Lemaître's law, where the commonly observed value of $p_6 \approx 0.4$ is presented as a balance between entropic and enthalpic contributions to the free energy. The deviations of ring areas from the ideal values are discussed and the relative insensitivity of the ring area to relatively strong distortions is highlighted.

4.1 Bilayer Materials

An important class of two-dimensional materials which have emerged in the 21$^{\text{st}}$ century are bilayers of silica, SiO_2, and related species [11]. These can be prepared experimentally by chemical vapour deposition on metal and graphene supports [4, 5]. As in the three-dimensional glass, the basic building blocks of silica bilayers are vertex sharing SiO_4 tetrahedra, maintaining full coordination for all atoms in the bulk [22]. These are arranged such that three of the vertices are connected to tetrahedra in the same layer, with the final vertex being shared between layers acting as a "bridge" (figure 4.1a). A consequence of these bridging oxygen atoms is to enforce a symmetry plane between the upper and lower layers.



(a) Tetrahedral bilayer (b) Triangle raft (c) Ring structure

Figure 4.1: Silica bilayers of vertex sharing tetrahedra, (a), can be represented as a two-dimensional triangle raft, (b), (silicon and oxygen atoms are coloured yellow and red respectively). The ring structure then emerges from the three-coordinate network comprising the silicon atoms, (c).

Topologically, the symmetry plane means that these materials can be viewed as effective two-dimensional networks. Taking one of the layers, without the bridging oxygens, and projecting the atoms onto the horizontal plane reveals a representation of vertex sharing triangles, referred to as a triangle raft (figure 4.1b). The ring structure then emerges from the three-coordinate network formed by connecting the silicon atoms of adjacent triangles as in figure 4.1c. Indeed, scanning tunnelling microscopy (STM) has been used to directly visualise the ring structure in silica bilayers, revealing both crystalline and glassy arrangements and even the interface between the two [122, 123].

More recently experimentalists have also succeeded in synthesising bilayers of germania, GeO_2 [7, 8]. These have the same fundamental structure as SiO_2, but with more distorted tetrahedra Why again...some inorganic stuff... . This can lead to a build up of strain and rumpling of the tetrahedral layers. Have you discussed somewhere why they are doing this? Control of the pore size and pote density critical for gas separation applications.... Note somewhere that making the GeO2 analogue is more difficult experimentally.....

4.2 Review of Existing Methods

As mentioned in the introduction, both *ab initio* methods and classical molecular dynamics have been used in computational studies of silica bilayers, which often require a starting atomistic configuration [20–22, 28]. One approach is to simply take an experimental sample as the starting structure. Whilst this is on the surface the best solution, the experimental configurations may contain defects or areas where the image is corrupted *i.e.* the configuration may not be “pristine”. Additionally, the location of each atom has an associated uncertainty which leads to discrepancies in the observed bond lengths and angles, which can be compounded by any out-of-plane distortions. Whilst computational refinement can attenuate these problems [23, 124], there remains the more fundamental question of how “typical” the available images are from experiment, as STM provides exceptional information but only on relatively small sample sizes. Computational techniques can therefore prove a valuable tool for generating a large number of high-quality configurations, and corroborating experimental information.

One current approach is to transform amorphous graphene configurations [22]. Here amorphous samples of carbon are generated using a bond switching method (as outlined in section 3.2), before the carbon atoms are swapped from silicon and decorated with oxygens. Whilst this is a valid approach, the method assumes that the two materials are topologically equivalent. This is likely an oversimplification, as the presence of the bridging oxygens in silica afford the structure increased flexibility when compared to the carbon analogue. This likely explains why this method has struggled to mirror experimentally observed values of the ring statistics and Aboav-Weaire parameter, with small and large rings being under-estimated [29].

An alternative approach is to use molecular dynamics coupled with an effective pair potential to obtain viable configurations [28]. Such methods are relatively common, having been employed previously to study amorphous graphene [99]. Such methods offer the potential for generating realistic configurations but are difficult to control as the cooling rates which must be applied are necessarily huge compared to experimental rates. A potential artefact of the high cooling rates is the effectively freezing in of defect states, either in terms of local coordination environments or highly-strained (three-membered) rings. In addition, as with the method above, such methods appears to systematically underestimate the Aboav-Weaire parameter, indicative of too little structural ordering.

4.3 Triangle Raft Method

The motivation of this work was to develop a construction algorithm to generate samples of silica bilayers which can capture the full two-dimensional network topology; both the ring distribution *and* correlations. The model should be able to explore all phases from crystalline to amorphous yet computationally efficient enough to produce configurations suitable for further high throughput calculations. To achieve this a grow-from-seed Monte Carlo algorithm has been developed, where rings are individually added to build a triangle raft. This approach takes inspiration from the first hand-built models, which have been noted to bear good similarity to experimental structures [125, 126]. Such models were superseded by computational techniques designed to generate periodic configurations. However, the recent development in techniques to simulate aperiodic samples, such as sliding boundary conditions for molecular dynamics [127], makes this constraint no longer essential, and benefit may be gained from the added freedom of an aperiodic model.

4.3.1 Potential Model

As explained in figure 4.1 it is possible to capture the full topology of silica bilayers with a simplified representation consisting of a network of vertex-sharing SiO_3 triangles. As the focus of this chapter is on generating a large number of samples with varying ring statistics, to be used as a base for further calculations, working with this reduced representation is sufficient, as it provides a computationally efficient way to produce networks with the required *topology*. The precise *geometry* of the bilayer can be refined with advanced optimisation techniques if required [128].

In order to simulate bilayer systems in two-dimensions, a suitable potential model is needed which captures the essential physics of the system: the local triangular environment of the SiO_3 units and the relative energies of rings of different sizes. The model used here is modified from a relatively simple potential used in all-atom bilayer calculations [22, 23]. Each SiO_3 unit has a harmonic potential acting between all three Si–O pairs, and the three nearest-neighbour O–O pairs, given by:

$$\mathcal{U}_{ij} = \frac{K}{2}\left(r_{ij} - r_{ij}^{0}\right)^2, \tag{4.1}$$

where K is a constant, r_{ij} is the interatomic separation and r_{ij}^0 the equilibrium interatomic separation between i, j. The spring constant, k, is set to be very stiff, whilst the equilibrium separations are set according to elemental species such that $r_{\mathrm{OO}}^0 = \sqrt{3}\, r_{\mathrm{SiO}}^0$, maintaining a set of ideal $\mathrm{SiO_3}$ triangles.

The Si–O–Si angle, which determines the strain associated with different ring sizes, is controlled by a shifted and cut 24-12 potential of the form:

$$\mathcal{U}_{ij} = \begin{cases} \epsilon\left[\left(\frac{r_0}{r_{ij}}\right)^{24} - 2\left(\frac{r_0}{r_{ij}}\right)^{12}\right] + \epsilon & r_{ij} \leq r_0 \\ 0 & \text{otherwise} \end{cases} \tag{4.2}$$

where ϵ is a constant and r_{ij} is now the Si–Si separation between atoms in adjacent triangles. It is the value of r_0 which sets the Si–O–Si angle at which strain begins to be felt and therefore the relative ring energies. Taking the hexagonal lattice as being the zero in energy it follows that $r_0 = 2r_{\mathrm{SiO}}$. Rings which deviates increasingly from the ideal hexagon will therefore incur an increasingly energetic penalty.

To summarise, the primary aim here is to generate topologies suitable for later investigation using more detailed (and hence more accurate but more computationally-demanding) potential models. As a result, the harmonic springs simply control the local (triangular) geometries whilst the 24-12 potential controls the repulsion between these local polyhedra. These functions are chosen as deliberately simple to improve computational efficiency and achieve high throughput of idealised networks. Furthermore, the parameters k and ϵ need have no direct physical meaning, simply controlling the meaning of the system "temperature" as discussed below. The only requirement is that they generate energies of the same magnitude to allow for efficient structural evolution. Accompanying figure

4.3.2 Algorithmic Details

Using the model detailed above, a Monte Carlo construction algorithm has been developed which allows two-dimensional networks to be built ring by ring in the shape of a specified function. The main steps of the algorithm are outlined below:

1. Take a starting seed, such as a single ring or experimental configuration
2. Select triangles on which to build the next ring (see figure 4.2)
   (a) Overlay a function on the network (*e.g.* circle, square)
   (b) Check for atoms with dangling bonds lying inside the function region
   (c) If no such atoms exist, systematically increase the function size until an atom is found
   (d) Find the next nearest atoms which also have a dangling bonds
   (e) Choose the two triangles that correspond to the largest starting ring size
3. Determine the probability of constructing rings of different sizes
   (a) Build trial rings in the range k_{min} to k_{max} (see figure 4.3)
   (b) Geometry optimise the local structure and calculate minimised potential energy (as explained in section 3.2.3)
   (c) Calculate the probabilities of each ring occurring, P_k, equation (4.3)
4. Accept single trial ring according to the probability distribution
5. Repeat steps 2 $\rightarrow$ 4 until the target number of rings is reached

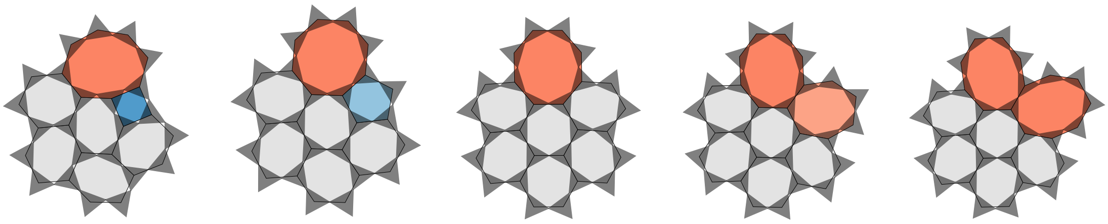
The probability of a ring of size k being accepted, P_k, is given by the equation:

$$P_k = \frac{\exp\left[-\left(\mathcal{U}_k - \mathcal{U}_k\right)/T\right]}{\sum\limits_k \exp\left[-\left(\mathcal{U}_k - \mathcal{U}_0\right)/T\right]}, \tag{4.3}$$

where $\mathcal{U}_k$ and $\mathcal{U}_0$ correspond to the energy of the trial structure and lowest energy of all trial structures respectively, and T is a "temperature". The parameter T controls how easily the potential energy landscape can be explored, and therefore how accessible strained rings become. In the low T limit, the acceptance probabilities are dominated by the energy term, and the rings which are selected will be those with the lowest energy. Note that this is not necessarily the 6-ring, but rather is dependent on the local environment. On the other hand, in the high T limit, the acceptance probabilities are approximately equal, and rings are selected on a more

(a) **(b)**

Figure 4.2: Panel (a) shows how triangles used to construct a ring are initially selected. There are no atoms with dangling bonds within the first search region (blue dashed line), and so the search area is extended (red dashed line), where triangles A and B are found. Panel (b) gives the three possibilities for the triangles that will form part of the constructed ring: A–C–D–B, A–E, B–F. As A–C–D–B corresponds to the largest starting ring size this is selected.



(a) **(b)** **(c)** **(d)** **(e)**

Figure 4.3: Geometry optimised structures for trial rings in the range $k = 4 - 8$. The ring structure is shown along with the SiO_3 triangle

Table 4.1: Variation of acceptance probabilities with temperature for the configurations in figure 4.3.

P_k	4	5	6	7	8
$T = 10^{-4}$	0.0000	1.0000	0.0000	0.0000	0.0000
$T = 10^{-3}$	0.0000	0.8837	0.1162	0.0001	0.0000
$T = 10^{-2}$	0.0336	0.4104	0.3351	0.1659	0.0550
$T = 10^{-1}$	0.1734	0.2227	0.2183	0.2034	0.1822
$T = 10^{0}$	0.1973	0.2023	0.2018	0.2004	0.1982

random basis. This is demonstrated in table 4.1, using the example configurations from figure 4.3. The “temperature” parameter is therefore the primary method for controlling the distribution of ring sizes in constructed networks.

4.4 Properties of Triangle Rafts

The triangle raft method is evaluated in terms of its effectiveness in producing configurations which accurately replicate the network properties of experimental silica bilayers *i.e.* the ring statistics and Aboav-Weaire parameter. It is also compared against the existing methods introduced in section 4.2, namely generation from amorphous graphene or molecular dynamics. This is performed in wider context of systematically varying the model parameters to explore the behaviour of generic networks of this type.

4.4.1 Network Growth

The triangle raft method is robust and controllable, and is able to generate configurations with tuneable ring statistics and topologies. Results will largely focus on the system where $k = 4 - 10$, denoted $\{4, 10\}$, mimicking the experimentally observed range for silica bilayers. Six example configurations are given in figure 4.4, which are generated with a range of temperatures and growth geometries. Figures 4.4a-4.4d provide a good qualitative analysis of the effect of temperature on the ring structure. At low temperature a phase boundary can be seen separating crystalline and amorphous regions, as seen in experimental silica bilayers [123]. In these samples although the proportion of small and large rings is low, their positions are highly correlated and chain structures of alternating rings sizes are clearly present. These motifs are reminiscent of defects found in a wide range of materials, including amorphous graphene and thin silicon and germanium oxides [3, 7, 11, 20]. The increase in temperature is coupled with the emergence of rings of more extreme sizes and regions which could be viewed as nano-crystalline are dispersed. The high temperature limit reveals a fully amorphous structure.

Figures 4.4e and 4.4f give examples of the diverse geometries in which samples may be constructed. It is interesting to note that even “difficult” shapes, such as those containing concave regions and cusps, do not prevent growth. Although the shape does not affect the network topology and is in a sense arbitrary, certain calculations may benefit from the different configurational shapes. For instance, molecular dynamics with sliding boundary conditions requires fitting of a smooth function to the sample perimeter, which is facilitated by having a near-circular form. Other areas such as percolation problems may benefit from square samples.

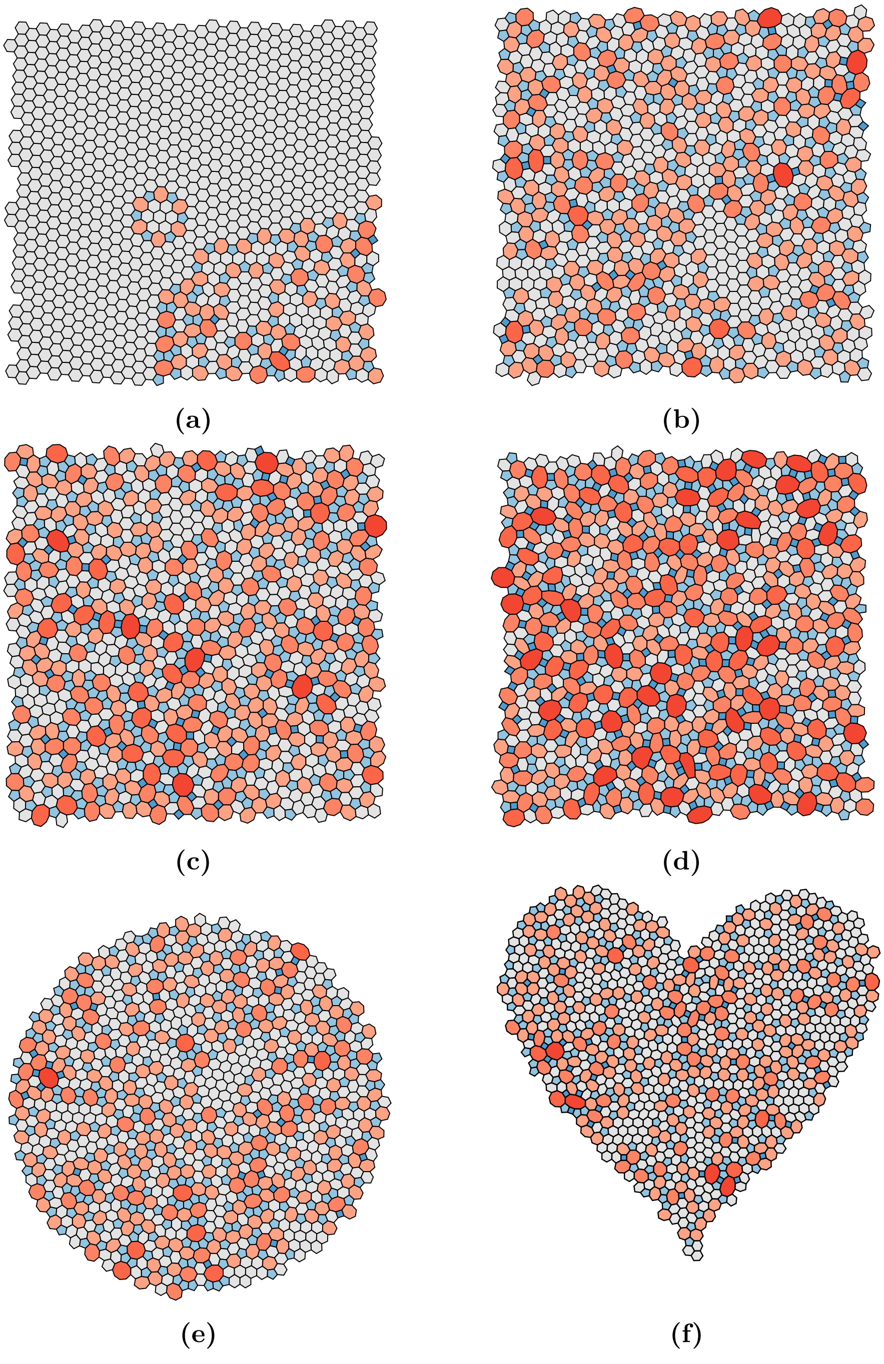


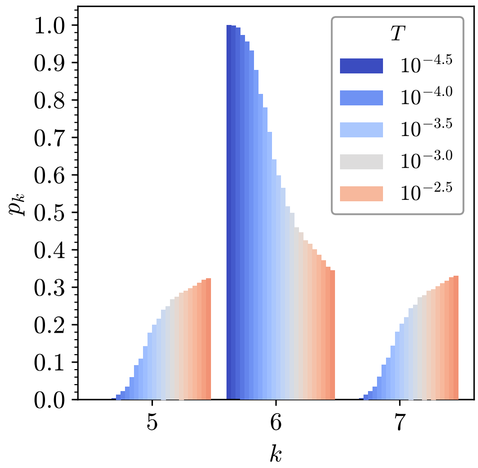
Figure 4.4: Example 1,000 ring configurations generated with different temperatures and shapes. Panels (a) through (d) show square lattices grown at $T = 10^{-4.0}$, $10^{-3.0}$, $10^{-2.5}$, $10^{-2.0}$ respectively. The samples show the increasing diversity in ring structure as temperature is increased. Panels (e), (f) show configurations with alternative lattice shapes at $T = 10^{-3.0}$, demonstrating the flexibility of the method in growing samples with variable geometries. Rings are coloured according to size with $k < 6$ as blue, $k = 6$ as grey and $k > 6$ as red.

4.4.2 Network Properties

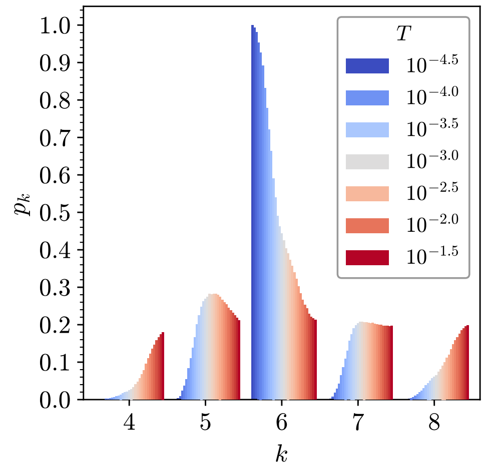
The quantitative relationship between temperature and ring structure was investigated for three systems of varying ring size ranges; $\{5,7\}$, $\{4,8\}$ and $\{4,10\}$. For each system, 100 samples consisting of 1000 rings were grown at temperatures between $T = 10^{-4.5} \rightarrow 10^{-1.5}$. The evolution of the combined ring statistics with temperature is presented in figure 4.5. Figures 4.5a-4.5c give bar representations of the ring size distributions for the three systems, which show different behaviours. For $\{5,7\}$ the individual p_k are all monotonically increasing ($k \neq 6$) or decreasing ($k = 6$) functions, but both $\{4,8\}$ and $\{4,10\}$ have p_k containing maxima. Additionally, both $\{5,7\}$ and $\{4,8\}$ achieve uniform distributions in the high temperature limit but $\{4,10\}$ does not.

This disparity in behaviour can largely be traced back to the constraint of Euler's theorem. As $\{5,7\}$ comprises of just three ring sizes, Euler's formula demands that $p_5 = p_7 = (1 - p_6)/2$ and so the system is relatively well defined. Hence as the 5 and 7-rings are more strained than the 6-ring, p_5 and p_7 show a systematic increase with temperature. Furthermore, the uniform equilibrium distribution can only satisfy Euler's formula when the ring size range is symmetric about 6, as is observed for $\{5,7\}$ and $\{4,8\}$. The form of the ring statistics at intermediate temperatures and for $\{4,10\}$ follow the maximum entropy solutions according to Lemaître's law, discussed in section 2.2.2 and later in this section.

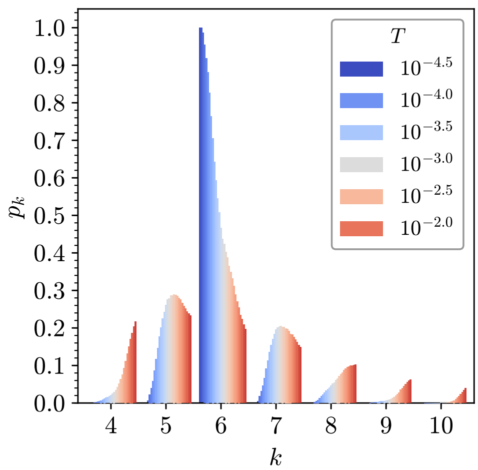
The ring distribution for $\{4,10\}$ is also shown as a function of temperature in figure 4.5d, along with the value of the Aboav-Weaire parameter, α, allowing for more facile comparison with experiment. The temperature which gives the best agreement between our model and amorphous experimental samples is highlighted by the vertical dashed line. The values of p_k and α are provided in table 4.2, alongside results from two experimental samples. It is evident that the model can be successfully tuned to match the topology of the experimental system. Not only are the ring distributions in very good accordance, but also the ring correlations, which have until now proved difficult to capture. This provides confidence that



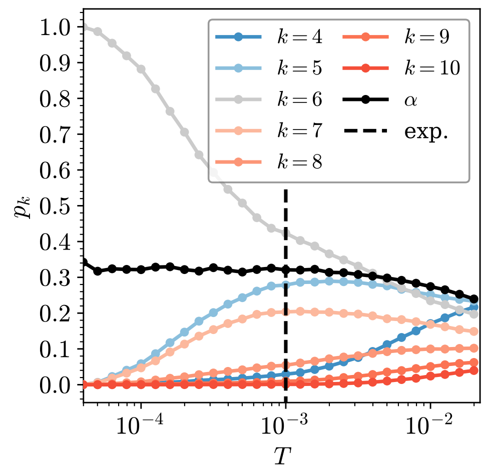
(a) $\{5, 7\}$



(b) $\{4, 8\}$



(c) $\{4, 10\}$



(d) $\{4, 10\}$

Figure 4.5: Variation in ring statistics with temperature over a given allowable k-range. Panels (a)-(c) show bar graph representations of the ring statistics, coloured by temperature, for the $\{5, 7\}$, $\{4, 8\}$ and $\{4, 10\}$ systems, respectively. Panel (d) gives an alternative line graph representation of the ring statistics for $\{4, 10\}$, coloured by ring size, along with the Aboav-Weaire parameter. The temperature which gives the best match to the experimentally observed amorphous region is also highlighted (vertical black dashed line).

this simplified but physically motivated triangle raft model is able to reproduce the behaviour of real systems.

Table 4.2: Comparison of silica bilayer samples from experiment, computational modelling and theory.

	Experiment		Computation				Theory
	Ru(0001) [126]	Graphene [4]	MC[a] [29]	MC[a] [29]	MD[b] [28]	TR[c]	Lemaître [57]
$\mathcal{N}$	317	444	216	418	16×85000	1000×100	–
p_3	0.0000	0.0000	0.00	0.00	0.0038	0.0000	0.0000
p_4	0.0379	0.0383	0.02	0.00	0.0537	0.0295	0.0280
p_5	0.2744	0.2725	0.33	0.37	0.2686	0.2786	0.2834
p_6	0.4448	0.4189	0.37	0.32	0.3773	0.4234	0.4200
p_7	0.1609	0.2117	0.21	0.25	0.2224	0.2034	0.2077
p_8	0.0757	0.0495	0.07	0.06	0.0602	0.0544	0.0518
p_9	0.0063	0.0068	<0.01	0.00	0.0118	0.0097	0.0082
p_{10}	0.0000	0.0023	0.00	0.00	0.0018	0.0010	0.0009
$p_{>10}$	0.0000	0.0000	0.00	0.00	0.0004	0.0000	0.0000
μ_2	0.9460	0.9333	0.94	0.86	1.1302	0.9208	0.8985
α	0.32	0.33	0.18	0.23	0.25	0.32	–

Note: Each method is given alongside the number of rings in the sample, $\mathcal{N}$, followed by the ring statistics, p_k, the second moment of the ring statistics, μ_2, and the Aboav-Weaire parameter, α

[a] Bond Switching Monte Carlo (graphene potential) [b] Molecular Dynamics

[c] Triangle Rafts, this work, $T = 10^{-3}$

Table 4.2 also lists the ring statistics obtained from previous computational studies which used both Monte Carlo and molecular dynamics methods. As mentioned in the review of these methods above, neither fully succeeds in accurately capturing the topology of silica bilayers. Kumar *et al.* attempted to transform an amorphous graphene structure generated from bond switching Monte Carlo into a silica bilayer. The ring statistics of the resulting structure were approximately correct, but the proportion of 5- and 6- rings over- and under-estimated respectively. In addition the Aboav-Weaire parameter was substantially lower than experiment, indicating a relative lack of structure in the ring ordering. The origin of these discrepancies is likely the use of a graphene potential model. The increased stiffness of the carbon network (which unlike silica lacks bridging oxygens) means a high temperature must be used to obtain an amorphous structure with the required disorder. This leads to heavily distorted rings (as noted in the original paper) which reduces the requirement for small rings to be adjacent to large.

Roy *et al.* have an alternative approach of generating configurations with an effective pair potential and molecular dynamics. As can be seen the ring statistics are closer to the experimental values, but now contain artefacts, with a significant fraction of highly strained 3-membered rings and large rings up to $k = 14$. These manifest as a result of the artificially high cooling rates in the computational studies which trap defect states in the configurations. Once again the final Aboav-Weaire parameter, α, is underestimated.

It is worth re-emphasising here that the triangle raft method is able to replicate experimental values of both p_k and α, due to its tuneable approach and "organic" growth mechanism, where sample formation is not influenced by enforced periodicity. Beyond this, the controllable nature of the method also allows insight into key questions about silica bilayers, for instance the form of the ring distribution in this amorphous phase. As detailed in section 2.2.2, the maximum entropy ring distribution can be calculated numerically given the value of p_6. For example, table 4.2 gives the maximum entropy solution for $p_6 = 0.42$, which agrees very well with the results from triangle rafts and experiment. This second moment

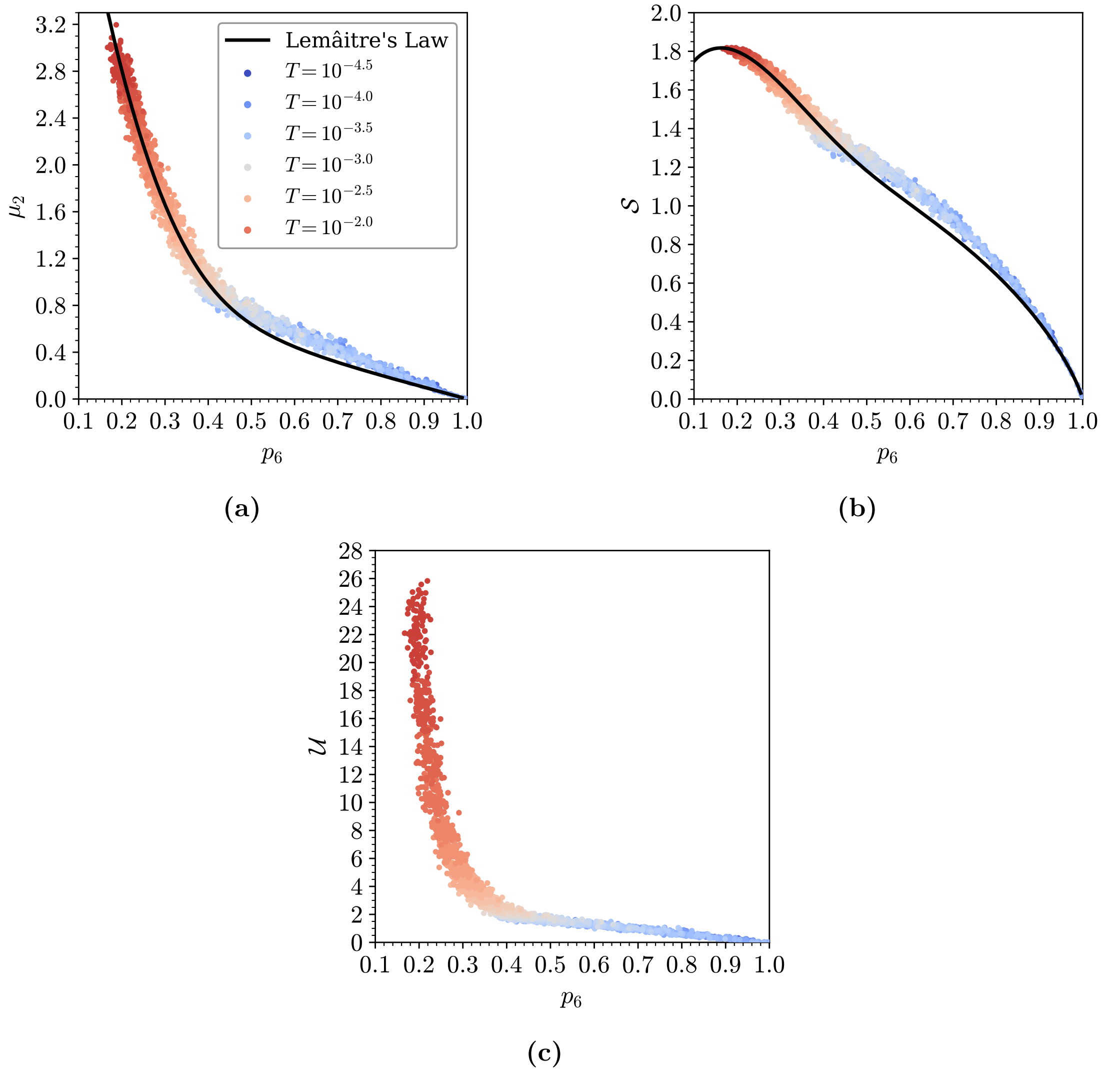


Figure 4.6: Evolution of ring statistics (a), entropy (b) and potential energy (c) of triangle rafts with temperature. The experimental value of $p_6 \approx 0.4$ occurs just before the exponential increase in potential energy, reflecting the balance of energetic and entropic factors.

of the distribution, μ_2, is then uniquely related to p_6 via Lemaître's law, shown as the black line in in figure 4.6a.

However, Lemaître's law gives no information on why a particular maximum entropy distribution is found for a given system. The triangle raft method allows systematic generation of configurations with different p_6 values by tuning the temperature parameter. The resulting configurations follow Lemaître's law across the entire temperature range. Figures 4.6 gives the results from the individual 1000 ring samples, coloured by temperature. Figures 4.6a and 4.6b compare the observed μ_2 and S (entropy) of the generated configurations to those expected from Lemaître's law, showing the law provides a good fit, with only a small deviation observed for $p_k > 0.5$.

Figure 4.6c plots the geometry optimised potential energy of the samples against p_6, which increases as the ring sizes become more diverse. The curve is split into two regimes, with gradual increase in energy from $p_6 = 1.0 \rightarrow 0.4$ followed by exponential increase for $p_6 < 0.4$. This is consistent with the information in figure 4.5d which shows that below $p_6 \approx 0.4$, not only does the number of extreme ring sizes increase rapidly, but they become less correlated with a lower α, decreasing the number of favourable small-large ring pairings.

It can now be proposed why the experimental amorphous distributions are found with a value of $p_6 \approx 0.4$. The system aims to maximise entropy by obtaining a ring distribution along the Lemaître curve with the minimum p_6 possible. However, for $p_6 < 0.4$ the energetic cost becomes prohibitively large, as higher entropy distributions can only be achieved by increasing the proportion of extreme ring sizes at the expense of relatively low strain 5- and 7- rings. Interestingly it is also evident why no configurations are present below $p_6 \approx 0.16$, even at the highest temperature. Below this point, the entropy of the $\{4, 10\}$ system decreases whilst the energy continues to rise and so there is no driving force to sample this area of phase space.

4.4.3 Physical Properties

As an additional check that the developed triangle raft model behaves physically, the angle distribution between adjacent SiO_3 units, $f(\theta)$, was calculated for the $\{4, 10\}$ system across the range of temperatures studied. The results are summarised in figure 4.7a. The angle distributions are necessarily symmetric about 120°, as each triangle pair contributes two complementary angles. At lower temperatures the distribution is dominated by angles close to 120°, as a consequence of the large proportion of near strainless six membered rings. Furthermore, at the temperature corresponding to the amorphous experimental region, $T = 10^{-3}$, the distribution has a similar extent to the angle distribution found in experimental samples (see for example figure 7 reference [28]). However, as the temperature increases, the form of $f(\theta)$ does not simply broaden as might be expected, but becomes bimodal. This can be rationalised by considering the angles that would be present in regular polygons of different sizes, marked by vertical lines in figure 4.7a. These ideal angles are clustered away from the mean value of 120°, and hence increasing the diversity of ring sizes through temperature acts to shift the most commonly observed angles from the central value of 120°. It is therefore interesting to note that increasing structure in the angle distribution does not necessarily translate to increased order in the atomic configurations.

A final check comes from examining the ring areas in the generated configurations. Inspection of amorphous experimental samples reveals that the rings appear highly regular in shape. This can be quantified by determining the average dimensionless area for each ring size, A_k, and comparing it to the area of the corresponding regular polygon, A_k^0, where:

$$A_k = \frac{\langle Area\,(k) \rangle}{\left(r^0_{\mathrm{SiSi}}\right)^2}, \tag{4.4}$$

$$A_k^0 = \frac{k}{4\tan\left(\pi/k\right)}. \tag{4.5}$$

As the regular polygon has the maximum achievable area for a given ring size, the ratio A_k/A_k^0 is expected to lie in the range $0 \to 1$, with a lower value corresponding to increased deviation from regularity, and assuming r^0_{SiSi} to be fixed.

The study by Kumar *et al.* found that whereas for experimental configurations, $A_k/A_k^0 \approx 1$, configurations generated using thier bond switching method generally displayed ratios much less than unity [99], indicative of large distortions in the ring structure. For larger rings, a value of $A_k/A_k^0 > 1$ was also found, which can only be achieved if there is appreciable bond stretching (see equations (4.4), (4.5)).

The analogous results for the method presented in this chapter can be found in figure 4.7b, for $T = 10^{-3}$. This figure demonstrates that there is now good agreement between experimental and computational results. In both cases the deviation from regularity increases with ring size, as the flexibility of the rings

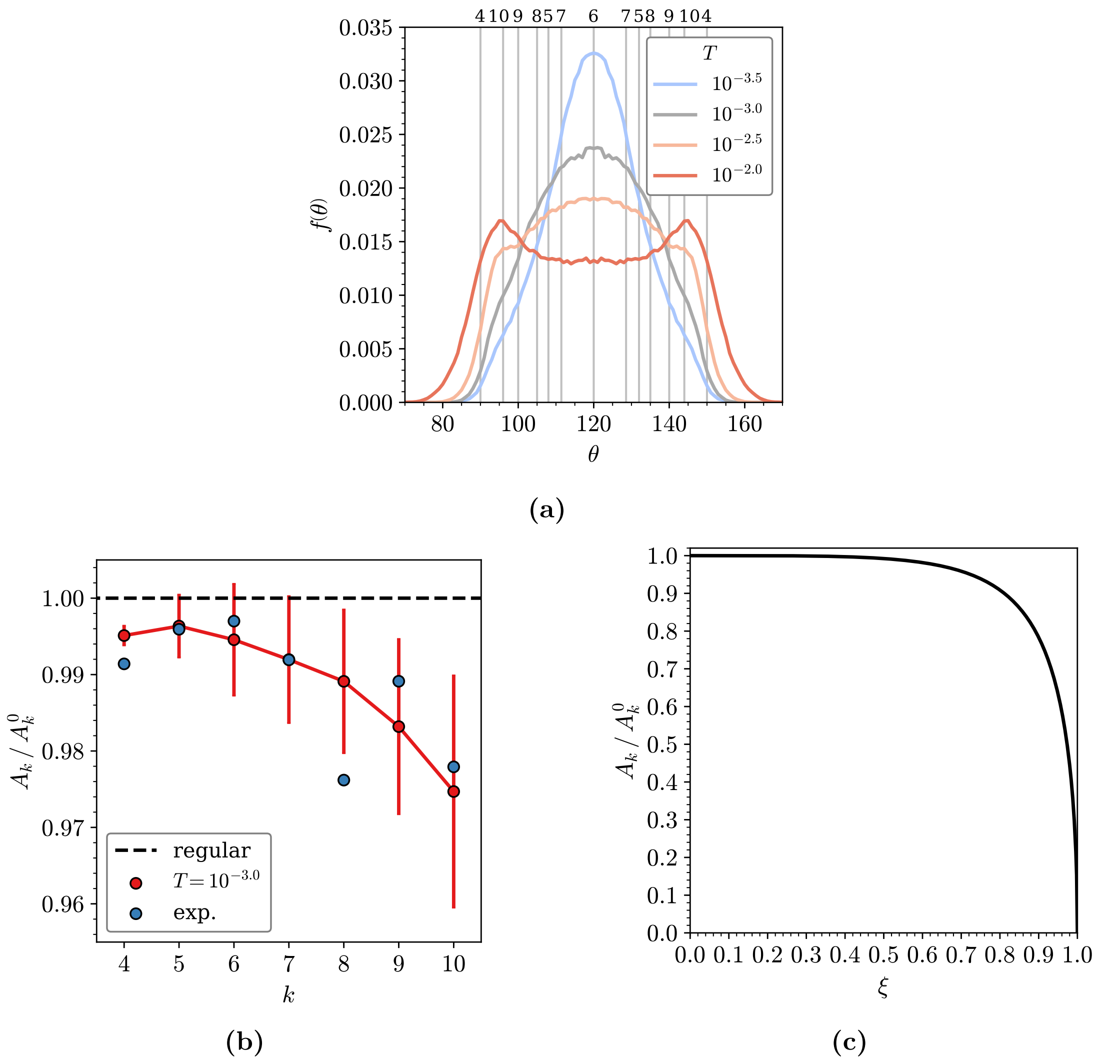


Figure 4.7: Panel (a) gives the ring angle distribution function for triangle rafts formed at different temperatures. Panel (b) compares the regularity of rings in computational and experimental amorphous configurations, with points indicating the mean value and bars corresponding to the standard deviation. Experimental data is taken from ref. [29]. Panel (c) shows the effect on the area when distorting a circle to an ellipse whilst maintaining a constant perimeter length.

increases. Again it can be proposed that the difference between current and previous methods could be due to the lack of enforced periodicity on the system. By allowing the network to grow relatively freely, the system can avoid a build up of strain associated with maintaining periodic boundaries.

Even with this analysis, an argument can be made that by visual inspection the rings in the experimental configurations are still more regular than those generated from computational samples. Therefore one can consider if deformation of a ring should be expected to lead to significant reduction in area. This can be explored by considering the distortion of a circle to an ellipse. The degree of distortion can be described by the eccentricity of the ellipse,

$$\xi = \left(1 - \frac{b^2}{a^2}\right)^{1/2}, \tag{4.6}$$

where a, b are the major and minor axis radii respectively. This change in area with distortion is shown in figure 4.7c, the calculation of which can be found in appendix add ellipse appendix . As can be seen, a large degree of eccentricity is needed for a significant change in the observable area. For example, if $a = 1.5b$, the area is still $\approx 0.94\%$ of the area of the corresponding circle.

For silica networks the Si–Si distances lie in a narrow range because of the covalent nature of the atomic bonding and the near-linear Si-O-Si bridges which join the two layers. Hence we would expect similar behaviour to occur, with ring areas relatively invariant to distortions in the ring shape (this same analysis would not be expected to hold for foams for example, where the length of the boundary is much more flexible). This suggests that the ring area is not the most suitable metric for quantifying the regularity of rings in systems such as this, and could explain any disagreement between the seemingly near ideal ring areas and the visual evidence. As previously stated, although the potential model used is physically motivated, it is lightweight in order to facilitate generation of a large number of configurations with the correct network topology. In future it would be informative to see if the required regularity can be achieved by geometry optimising the resulting bilayer configurations with a more accurate potential, such as the TS potential which includes potentially significant electrostatic interactions including many-body polarisation effects [128].

4.5 Chapter Summary

In this chapter a method for the effective growth of two-dimensional networks from a given seed has been developed, allowing for control over the ring size distributions and the system topologies. The latter is often characterised by the Aboav-Weaire parameter, α, and the values obtained here are more commensurate with those obtained from experimental imaging compared with previously constructed configurations. The high throughput method has allowed a detailed analysis of Lemaître's law and has highlighted why the fraction of six-membered rings observed in real systems is often $\sim$0.4. Finally, a consideration of the ring areas show our configurations to contain more regular polyhedra than a number of previous configurations. However, the area itself is shown to be a relatively poor measure of a deviation from ideality for systems of this type.

5 | Targeted Optimisation of Atomic Networks

A targeted optimisation method is presented which enables two-dimensional networks to be constructed by reference to a set of ring statistics and Aboav-Weaire parameter, α, which controls the preferred nearest-neighbour spatial correlations. The method efficiently utilises the dual lattice and allows systematic exploration of configurational space. Three different systems are considered; a system containing 5-, 6- and 7-membered rings only (a proxy for amorphous graphene), the configuration proposed by Zachariasen, and those observed experimentally for ultra-thin films of SiO_2. The system energies are investigated as a function of the network topologies and the range of physically-realisable structures established and compared to known experimental results. The limits on the parameter α are discussed and compared to previous results, whilst the evolution of the network structure as a function of topology is discussed in terms of the ring-ring pair distribution functions. A short study on ring percolation in amorphous graphene is also presented.

5.1 Disorder in Two-Dimensional Networks

As mentioned in previously, the characterisation of the disorder in two-dimensional networks can be achieved through the ring structure. For three-coordinate atomic materials the mean ring size is constrained to six by Euler's law, which allows the variance of the ring size distribution, μ_2, to act as a proxy measure for disorder (see sections 2.2.1, 2.2.2). The same set of ring statistics can however lead to a large number of different ring arrangements, as shown in figure 1.1. These can be further quantified by the Aboav-Weaire parameter, which measures the ring-ring correlations. An interesting observation across a wide range of experimental systems, is that the measured value of the Aboav-Weaire parameter lies in a tight range of $\alpha \approx 0.15 \rightarrow 0.3$ [129]. This is also effect is also seen in computational studies, including for example the previous chapter.

Whilst it is necessary for good computational models to capture these measures accurately, they do not give insight into *why* such configurations are preferred. To answer this question a different approach is required, where configurations can be systematically generated, covering a parameter space which exceeds the experimentally accessible region. To achieve this a targeted optimisation method can be employed, whereby configurations are produced to fit network properties, and not driven by an underlying potential model. This allows the experimentally occurring structures to be viewed in the context of the wider configurational landscape.

5.2 Targeted Optimisation Algorithm

The primary remit of the targeted optimisation algorithm is to generate plausible network configurations based on the supplied network properties of ring statistics and Aboav-Weaire parameter. A secondary requirement is for the method to be efficient enough to produce samples for further high-throughput calculations. Both these goals can be successfully accomplished with the method presented here: a Monte Carlo search algorithm, using the machinery of bond switching.

The bond switching algorithm (described in detail in section 3.2), amorphises a crystalline hexagonal lattice by exchanging the neighbouring interactions between pairs of bonded atoms and geometry optimising the structure. Moves are accepted according to the resulting change in the potential energy, where those with lower energy are accepted with increasing probability. The driving force is therefore always towards a structure which is physically motivated. In targeted optimisation, the same Monte Carlo moves are proposed as in bond switching, but crucially moves are not accepted on the basis of the energy of the network, but rather its agreement with a target ring distribution and Aboav-Weaire parameter. This agreement is measured by a cost function of the form:

$$\Omega = K_\alpha|\alpha - \alpha^t| + \frac{|\mu_2 - \mu_2^t|}{\mu_2^t} + \sum_k \frac{|p_k - p_k^t|}{p_k^t}\,, \tag{5.1}$$

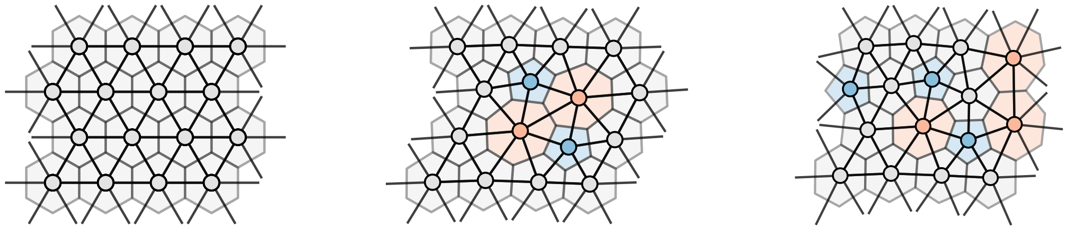
where K_α is a scaling constant; p_k^t, μ_2^t and α^t are the input target values; p_k are the system ring statistics; and μ_2 and α are calculated from an Aboav-Weaire fit on the current state. In the cost function the relative difference is used for the ring distribution, as the same accuracy is required for all p_k^t, which may differ by several orders of magnitude. This is not a concern for α^t, which must also have the flexibility to take a zero value, and hence the absolute difference is used in the first term.

Moves in targeted optimisation are accepted with probability given by the Metropolis condition:

$$P_{ij} = \min\left[1, \exp -\Delta\Omega/T\right]\,, \tag{5.2}$$

where $\Delta\Omega$ is the difference in cost functions before and after the proposed move, and T is a temperature parameter. In contrast to bond switching which is concerned with sampling, this is a global optimisation algorithm and moves are proposed until the network has converged to the target properties and the cost function is zero. As is the case with such optimisation techniques, steps must be taken to avoid becoming trapped in local minima, and the calculation not converging. This is achieved through selection of the parameters K_α and T. The parameter K_α changes the relative costs of satisfying the α^t and p_k^t conditions, and must be chosen so that neither is overweighted. The parameter T controls the proportion of moves which are accepted. Some temperature is required to overcome local minima, but if set too high the algorithm will no longer move downhill in cost and the search becomes effectively random - invariably leading to non-convergence. Values for K_α and T can be determined from a parameter search checking for convergence of target systems; but $K_\alpha = 10$ and $T \sim 10^{-4}$ were appropriate for systems of the type and size described in this work.

One key point which arises from using a cost function in this way is that there becomes no requirement for accurate on-the-fly geometry optimisation of the atomic



(a) (b) (c)

Figure 5.1: Bond switching Monte Carlo moves can be performed solely through the dual lattice. Two successive moves are shown from (a)-(b) and (b)-(c). In the dual lattice (bold circles and lines) two edge-sharing triangles are selected and the shared edge transposed. The atomic network is also shown (faded rings) to illustrate the corresponding effect on the atomic structure.

positions (as there is no need to calculate potential energies). It is the underlying topology of the network which determines the system properties, which is invariant to the geometry. The final energy of the system may well be of interest, but this can be evaluated just once at the end of the calculation. This opens the door for significant speed-ups through efficient use of the dual lattice.

5.2.1 Dual Space Implementation

Whilst the targeted optimisation algorithm can be employed using atomic positions, there are significant advantages to a dual space implementation. As discussed in section 2.1.2, the ring structure is better described through the use of the dual network. In this representation the ring statistics in equation (5.1) are simply given by the node degree distribution. In addition, the mean ring sizes about each ring, m_j, required for the Aboav-Weaire fit, equation (2.20), can be easily calculated from the joint degree distribution:

$$m_j = \sum_k \frac{k e_{jk}}{q_j} \,. \tag{5.3}$$

Hence, by utilising the ring network, the book-keeping to track the network properties becomes trivial.

The implementation of the bond switching move itself is also straightforward in dual space. Figure 5.1 shows how an atomic system can be manipulated *solely* through the dual lattice. Here the triangular nature of the dual (reflecting the trivalency of the atoms) can be exploited to good effect. By selecting edge sharing triangles in the ring network and transposing the shared edge connection, a perturbation equivalent to the Stone-Wales defect can be enacted. This process can be continued to generate an amorphous network.

In addition, although there is no strict requirement for geometry optimisation after each step, the triangle lattice can be used to maintain a reasonable physical structure in a cost efficient manner. By applying a harmonic potential, equation (4.1), between all pairs of linked nodes the ring centroids can be maintained at a reasonable separation. The atomic positions can then be regenerated by reversing the triangulation, placing species at the centre of each triangle, relatively close to the minimum in the atomic potential energy surface. Specifically, in this chapter a Keating potential, equation (3.21), is used with an interatomic separation of r_0 and $K_S = 5K_A$ (as in previous studies of amorphous graphene [99]). If the resultant polygons are assumed to be regular, the equilibrium separation for two polygons in the dual of sizes, k_i and k_j, can be expressed:

$$r_{ij}^0 = \frac{r_0}{2}\left(\frac{1}{\tan\left(\pi/k_i\right)} + \frac{1}{\tan\left(\pi/k_j\right)}\right) . \tag{5.4}$$

The extreme computational efficiency of evaluating the forces of the harmonic potential enables the targeted optimisation algorithm to complete rapidly whilst retaining the essential physics of the system. The final geometry can then be refined.

5.3 Mapping Configurational Space

The targeted optimisation algorithm provides a opportunity to gain insight into the physical meaning of the Aboav-Weaire and its effect on network topology. For this, a variety of test systems are used, the principle of which contains only $5 \to 7$ membered rings, a proxy for amorphous graphene, aG. This system represents a useful framework for investigating the Aboav-Weaire law due to the presence of additional constraints which make it highly controllable. As a consequence of

Euler's law the proportion of 5- and 7- rings must be equal, which leads to a trivial relationship between the second moment and proportion of 6- rings,

$$p_5 = p_7 = \frac{1}{2}\left(1 - p_6\right), \qquad \mu_2 = 1 - p_6 . \tag{5.5}$$

In addition, this allows the α parameter to be explicitly defined in terms of the difference between the $5-5$ and $5-7$ ring adjacencies:

$$\alpha = \frac{12\chi^5_{75} - \left(1 - p_6\right)^2}{6\left(1 - p_6\right)}, \tag{5.6}$$

where $\chi^5_{75} = e_{57} - e_{55}$ (details the derivation can be found in appendix appendix for derivation of aG α . This makes the aG model the first example of a system where the α parameter is well defined in terms of the underlying ring structure. It also highlights the relative complexity in the Aboav-Weaire parameter for even a seemingly simple case.

Two further systems with fixed ring statistics are also used to provide supplementary results. These are based on the Zachariasen configuration, figure 1.1, and experimental samples of silica glass, which are chosen to provide examples of increasing ring diversity, with the Zachariasen sample containing ring sizes in the range $k = 4 \rightarrow 8$ and silica $k = 4 \rightarrow 10$. The ring distributions for all the systems used in this chapter are summarised in table 5.1. In addition whereas the silica distribution should be easily achievable by the targeted optimisation algorithm (essentially following Lemaître's maximum entropy distribution), the Zachariasen distribution provides a more "extreme" case, where the distribution is not unimodal and the proportion of 5-rings is greatest.

Table 5.1: Ring statistics for systems used with the targeted optimisation algorithm.

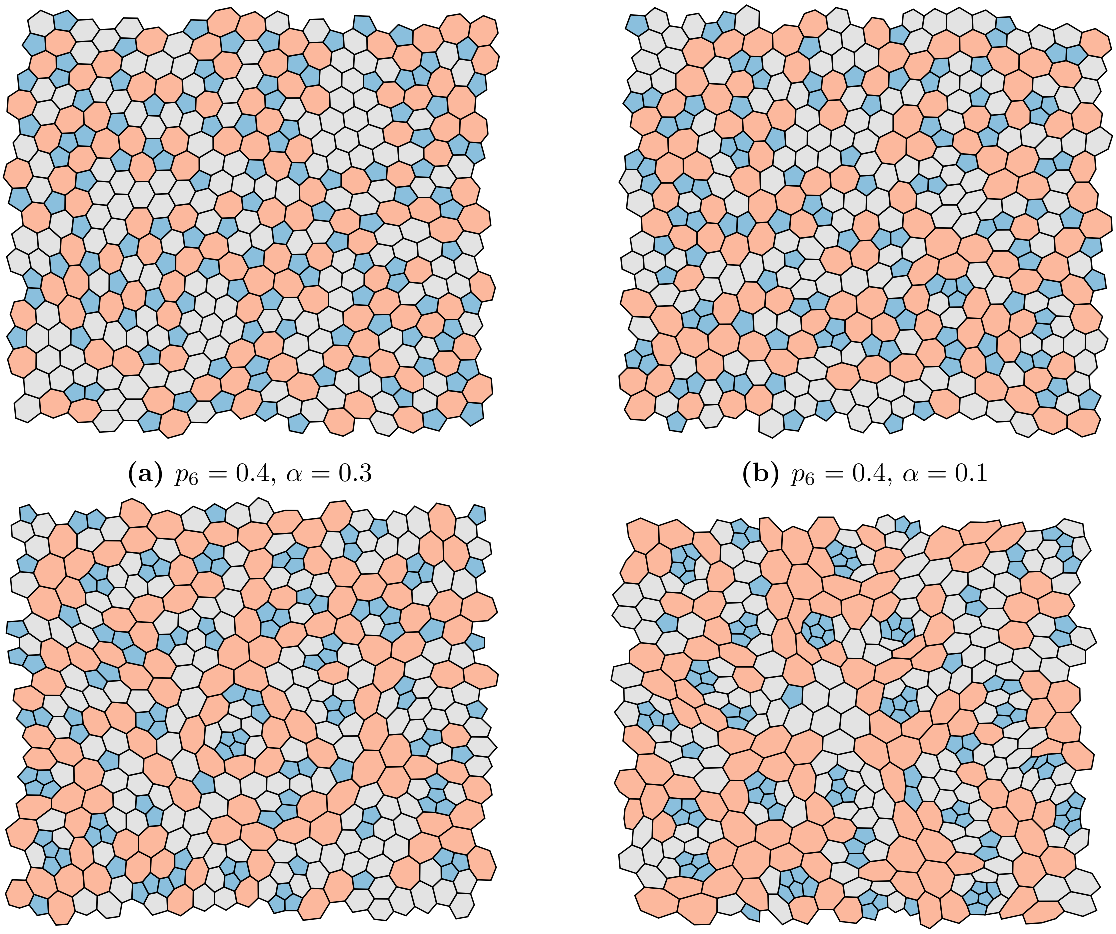
	p_4	p_5	p_6	p_7	p_8	p_9	p_{10}
aG	-	$\left(1 - p_6\right)/2$	p_6	$\left(1 - p_6\right)/2$	-	-	-
Zach.	0.10	0.35	0.15	0.25	0.15	-	-
SiO_2	0.040	0.268	0.420	0.210	0.050	0.010	0.002

5.3.1 Limits of the Aboav-Weaire Parameter

To begin mapping the configurational space of these atomic networks, the range of accessible α values for the aG system was determined by generating periodic networks containing 10,000 rings with $0.1 \leq p_6 \leq 0.9$. The aim of these simulations was to try and probe the topological limits of α, and so a high number of Monte Carlo steps was used, 10^9, without the need for geometry optimisation. Visualisations of the output of the targeted optimisation algorithm are given in figure 5.2 for $p_6 = 0.4$ and $\alpha = -0.3 \rightarrow 0.3$. These images give a good qualitative feel for the physical meaning of the Aboav-Weaire parameter: at low α similar sized rings tightly cluster together, dispersing as α increases to favour dissimilar ring pairings. Figure 5.3 shows the range of accessible α values as a function of p_6 *i.e.* those for which the targeted optimisation algorithm converges. The upper limit, α_{max}, appears a relatively weak function of p_6 whilst the lower limit, α_{min}, shows a much stronger dependence. In addition, the range of accessible values, $\Delta\alpha = \alpha_{\mathrm{max}} - \alpha_{\mathrm{min}}$, broadly mirrors the system entropy, although there is deviation around $p_6 = 1/3$.

5.3.2 Structure and Energetics

To explore the structural properties of the aG networks at different values of p_6 and α, 100 periodic networks containing 10,000 rings, were constructed for $p_6 = 0.2, 0.4, 0.6, 0.8$. These simulations were performed with geometry optimisation and so also provide information on the physical limits on α. Figure 5.4a displays the mean and standard deviation of the total potential energy for each p_6 atomic network across a range of α values. It can be seen that the energy minimum in each case is only weakly dependent on the value of p_6, varying from $\alpha \simeq 0.23$ at $p_6 = 0.8$ to $\alpha \simeq 0.27$ at $p_6 = 0.2$, and close to the value of α seen across many natural systems. Whilst there is little cost for small deviations from the minimum, decreasing α rapidly incurs a relatively large energetic penalty. Figure 5.4b shows the analogous energies when minimising through the dual lattice alone. The curves have a very similar form with the minima aligned, suggesting that



(a) $p_6 = 0.4$, $\alpha = 0.3$ **(b)** $p_6 = 0.4$, $\alpha = 0.1$

(c) $p_6 = 0.4$, $\alpha = -0.1$ **(d)** $p_6 = 0.4$, $\alpha = -0.3$

Figure 5.2: Configurations produced via targeted optimisation of an aG network with 400 rings. Each has the same ring statistics ($p_5 = 0.3$, $p_6 = 0.4$, $p_7 = 0.3$) but a variable α parameter.

working in dual-space can be sufficient to capture all system properties, with a much lower computational overhead.

Partial radial distribution functions (RDF) link to methods rdf can be used to further quantify any ordering imposed on the generated configurations. These partial RDFs are constructed in reference to the distance of the centroids of a k-ring from a central j-ring, denoted $g_{jk}(r)$. They can therefore equivalently be thought of as the dual-space RDFs between nodes of degrees j,k. The Euclidean distance is used as opposed to the topological distance (*i.e.* the number of links from a given node) as the latter has been shown to lead to artificial long range correlations [130].

Figures 5.5a and 5.5b show the partial RDFs for the 5-5 and 5-7 ring pairings,

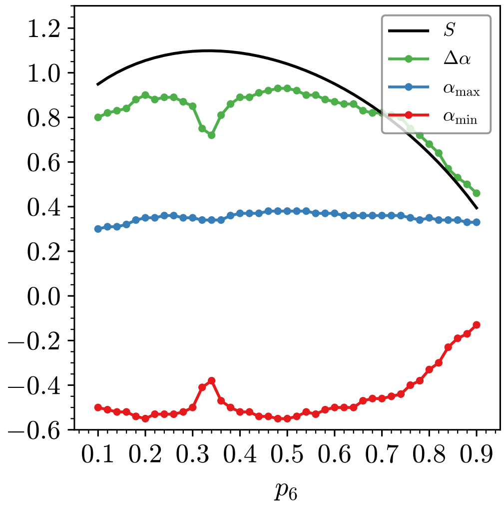
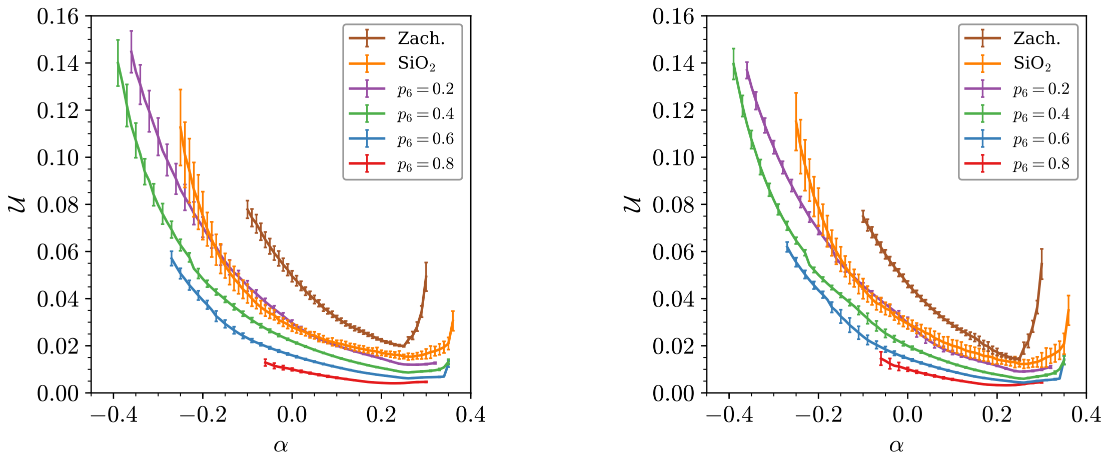


Figure 5.3: Accessible range of the Aboav-Weaire parameter in the aG system, for variable p_6.



(a) Minimisation through Atomic Network

(b) Minimisation through Ring Network

Figure 5.4: Geometry optimised potential energy of configurations produced via targeted optimisation for a range of systems with variable α parameter, with bars indicating one standard deviation from the mean. Panel (a) gives the results of optimisation through the atomic network with the Keating potential, whilst panel (b) gives the optimisation through the ring network with a simple harmonic potential.

$g_{55}(r)$ and $g_{57}(r)$ respectively add remaining to appendix? . As is consistent with its intuitive meaning, increasing α causes a reduction in intensity in the first peak of $g_{55}(r)$ and a concomitant increase in intensity in the first peak of $g_{57}(r)$, as 5-5 adjacencies are replaced with 5-7. In addition, the position of the first peak shifts to smaller r as α is reduced, reflecting both the increased distortion in the rings and the deviation from the ideal $2\pi/3$ bond angle, which translates to the higher observed potential energy.

These figures also show significant structural evolution beyond the nearest-neighbour length scale. As α becomes more positive, peaks emerge in $g_{55}(r)$ at $r/r^0_{55} \sim 1.8$ and ~ 2.3. An increase in α corresponds to a greater tendency for 7-rings to be near-neighbours to 5-rings and, in turn, increases the probability of the same 7-ring having a second 5-ring near-neighbour. In simple geometric terms, the second 5-ring can occupy three possible sites around the 7-ring fig here maybe, and for 8-4-8 , the non-adjacent positions corresponding to the developing peaks. Note that one might naïvely assume that driving α to more positive values would tend to eliminate the nearest-neighbour 5-5 spatial correlations. However, figure 5.5a indicates this not to be the case, reflecting the balance between retaining these units and facilitating nearest-neighbour 5-7 ring interactions via the formation of 5-7-5 triplets.

Similar analysis was performed on 100 generated Zachariasen and SiO_2 networks. Although our algorithm requires the fit to equation (2.20) to be exactly linear for the aG system, for broader ring distributions this is no longer the case. However, for the Zachariasen configuration the linear regression (R^2) coefficient was always in excess of 0.995, and for the silica the average R^2 was 0.979, representing a very good fit. Figure 5.4a shows the energies of both the Zachariasen and SiO_2 systems as a function of α. Both cases resemble those for the aG with energy minima at $\alpha \sim 0.25$. The silica curve shows smaller curvature reflecting the broader distribution of ring sizes whilst the Zachariasen curve shows a greater curvature reflecting the "extreme" *i.e.* physically unrealistic) nature of the distribution. In addition it proved difficult to generate low α configurations ($\alpha < -0.1$) for the Zachariasen network.

Figures 5.5c and 5.5d show two key RDFs for the Zachariasen configuration, $g_{44}(r)$ and $g_{88}(r)$, highlighting the spatial correlations between the smallest and largest rings in the system. The effects of changing α on $g_{44}(r)$ are dramatic with strong nearest-neighbour clustering at negative values. In this case, however, the nearest-neighbour 4-4 correlations do vanish at high α as 4-8 nearest-neighbour correlations dominate but the 8-ring is large enough to accommodate up to four 4-ring

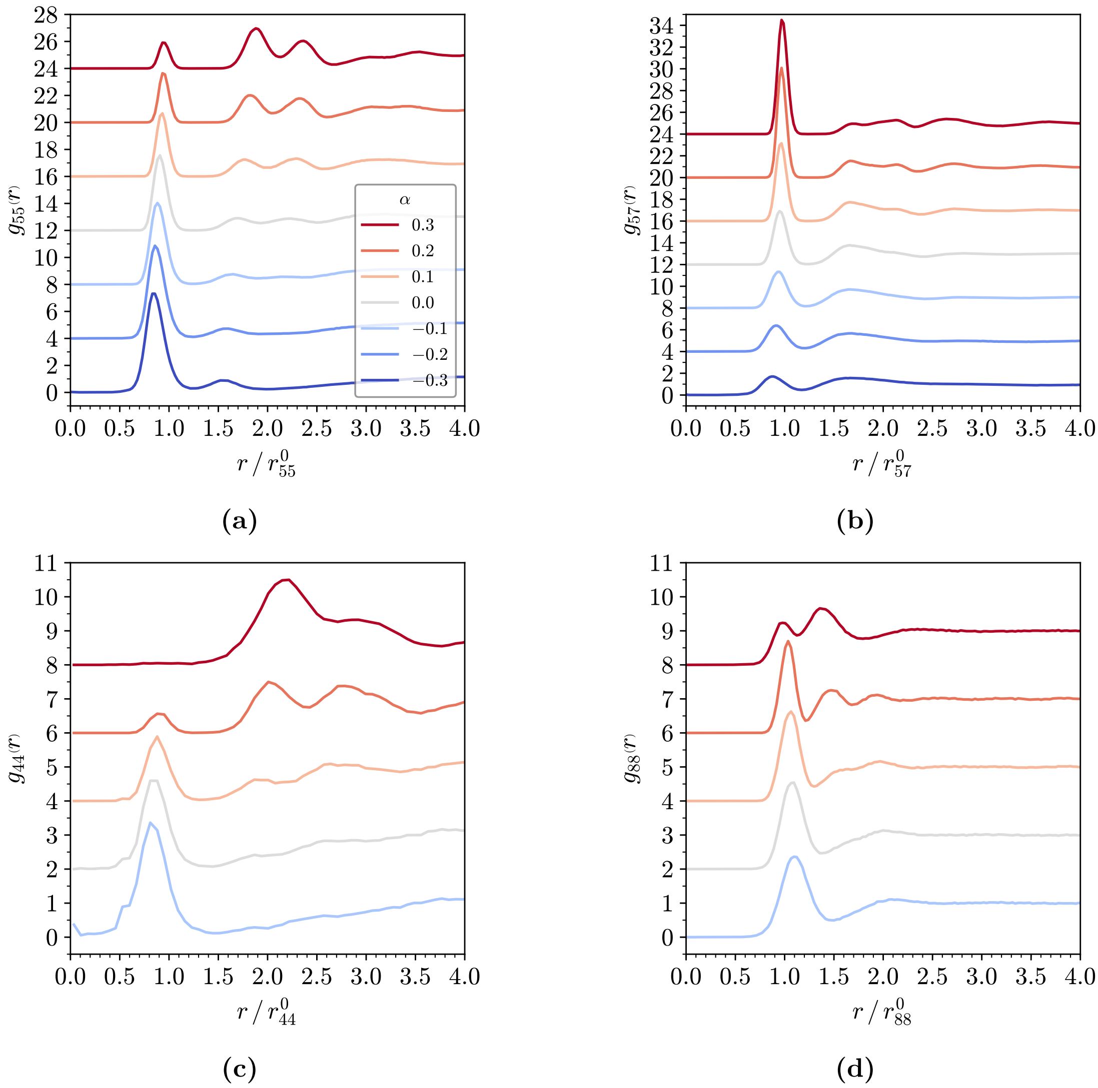


Figure 5.5: Partial RDFs for the aG (a)-(b) and Zachariasen (c)-(d) systems illustrate the evolution in ring structure with varying α parameter.

nearest-neighbours without any 4-4 neighbouring pairs. Again this is demonstrated through the next nearest neighbours by the 8-4-8 peak developing at ~ 1.4.

5.4 Ring Percolation in Amorphous Graphene

As a further demonstration of the utility and scope of the targeted optimisation algorithm, a short study is presented on the percolation of different ring sizes in aG systems. Owing to the fact that this is a standalone section, the theory pertinent to this investigation is first presented, followed by results.

5.4.1 Percolation Theory and Clustering

Percolation theory has its roots in problems concerning the flow of fluids through porous media [131], but now it can more generally be thought of as relating to the connectedness of components in a network (also referred to as *robustness*) [132]. The theory of clustering and percolation is an extremely rich field, which this thesis will merely dip its toe into, and so the discussion of the underlying theory be framed in the context of the aG networks already introduced in this chapter.

As an introductory example, consider a pristine hexagonal lattice for which the dual structure is a triangular net. It is clear that in this lattice all the nodes are connected *i.e.* there is some continuous path linking any two given rings. Equally, one could say that all the rings belong to the same cluster. Now imagine the process of removing nodes sequentially and at random from the original lattice, as shown in figure 5.6. Initially, removing nodes will have little effect on the network structure, but after a sufficient number are deleted, the interconnectivity of all the nodes will likely be broken, and the original large cluster will fragment into smaller clusters. At some point, the lattice will undergo a phase transition, from one in which there is a single "giant" component to one which has many small components. Quantifying this behaviour is the essence of percolation theory - exploring this transition and determining at what point this "percolation threshold" occurs.

To formalise this slightly, let's say there is an infinite triangular lattice, of which a random proportion, p, of nodes are occupied. The size of a cluster (*i.e.* the number of nodes which comprise it) can be denoted, s. The probability of a cluster of given size being found in the lattice is then P_s, and so the probability of an infinitely sized cluster as P_∞ [133]. The percolation threshold, p_c, is then the critical occupancy at which a giant component appears *i.e.*

$$P_\infty = \begin{cases} 1 & p \geq p_c \\ 0 & p < p_c \end{cases} . \tag{5.7}$$

Additionally, at this critical point, measures such as the average finite cluster size and correlation length diverge. For the example given above, which is the classic example of site percolation on a triangular lattice, the percolation threshold is $p_c = \frac{1}{2}$ [134].

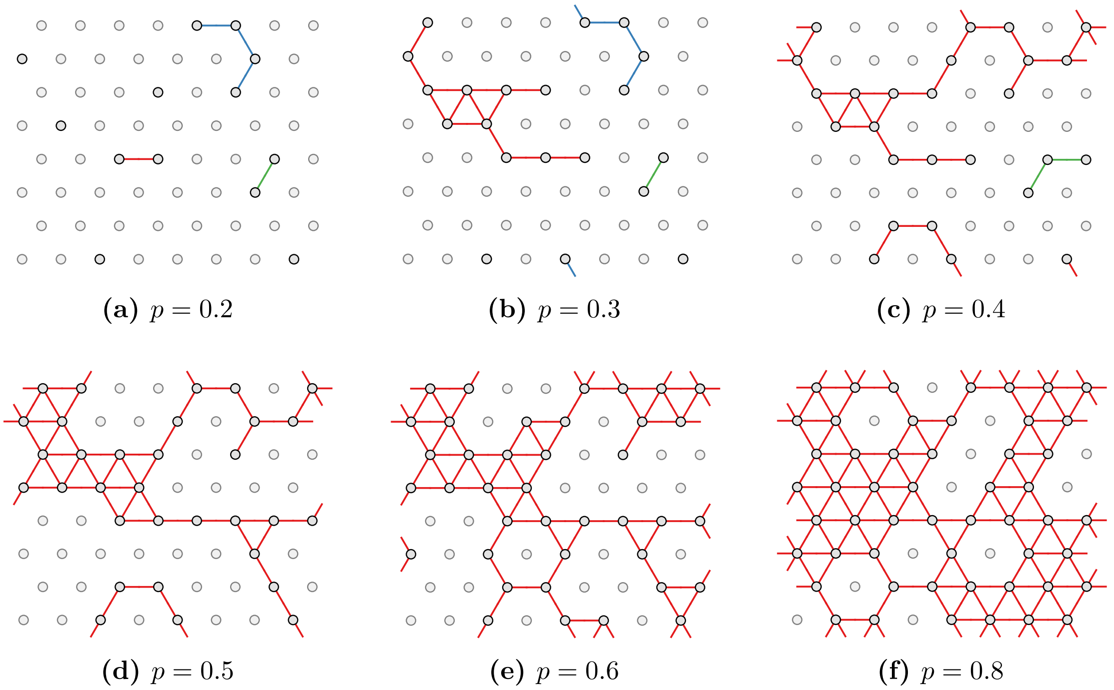
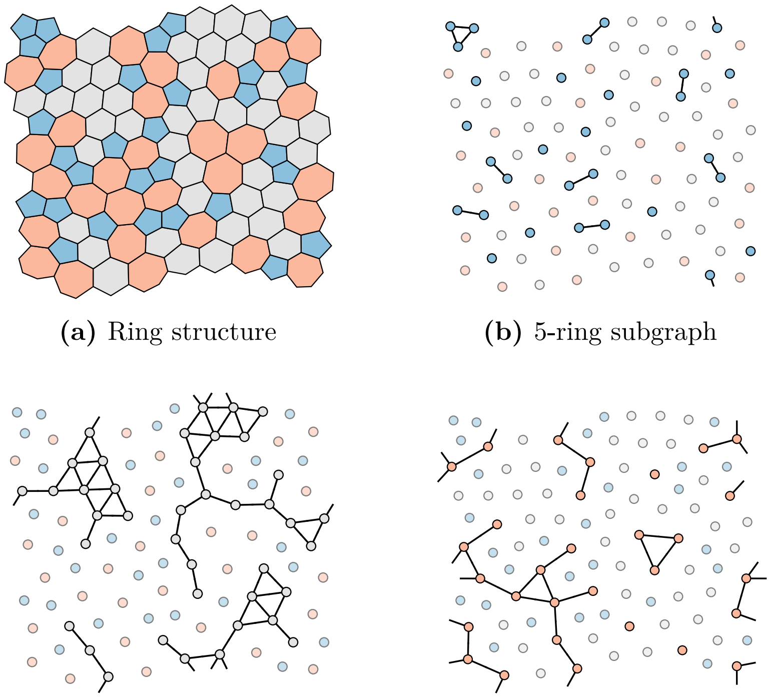


Figure 5.6: Site percolation on a triangular lattice. Panels (a)-(f) show network structure as site occupancy is increased, as indicated in the captions. Full circles signify occupied sites whilst connections are given by coloured lines, with the colour indicating nodes forming part of the same cluster.

The example of the triangular lattice above is one of the few examples of problems in percolation theory that can be solved analytically [135]. In order to find percolation thresholds for all but the simplest cases, numerical methods must be used. This problem is an ideal candidate for solution using a Monte Carlo method [136, 137]. One potential concern with a numerical method is that the lattices involved must be finite. The solution is to approximate the probability of a node residing in the infinite cluster as the probability of a node residing in the largest cluster. This is to say, if there are N nodes in the lattice and the maximum lattice size is s_{max}, then

$$P_\infty \approx \frac{s_{\mathrm{max}}}{N} . \tag{5.8}$$

This expression will hold in the limit of $N \to \infty$. As will be seen in section ref , for finite size lattices this approximation leads to smoothing of the step-like nature of P_∞. The percolation threshold in this case is then approximated by as the occupancy, p, for which $P_\infty = \frac{1}{2}$.



(a) Ring structure **(b)** 5-ring subgraph

(c) 6-ring subgraph **(d)** 7-ring subgraph

Figure 5.7: Panel (a) gives an example disordered aG ring structure and panels (b)-(e) the associated ring subgraphs, as indicated in the figure captions. Each subgraph contains only vertices and edges pertaining to the given ring size.

5.4.2 Percolation in Disordered Networks

The example of preceding section concerns site percolation on a regular lattice, where each site is equivalent. However, the ring networks of interest in this work are disordered, where sites have different node degrees (reflecting the underlying ring sizes). Disordered lattices therefore have an extra degree of complexity when compared to their ordered analogues. This allows the study of the percolation of different ring sizes in the network. To achieve this one must first construct the subgraphs for each ring size, which contain only the vertices and edges which relate to a given node degree, as shown in figure 5.7. The percolation threshold can then be studied for each of these subgraphs. For each k-ring subgraph, the percolation threshold will naturally depend on the global ring statistics, p_k. However, unlike the regular lattices, the percolation threshold must also depend on the ring correlations, which must influence the clustering [138]. As seen throughout this chapter, this property is controllable through the Aboav-Weaire parameter. Therefore, for each

k-ring subgraph, the percolation threshold will be a function of both a critical ring frequency, p_k^c, and a critical Aboav-Weaire value α_k^c.

5.4.3 Phase Diagram of Amorphous Graphene

The percolation phase behaviour was investigated for the aG system, containing only 5-, 6- and 7-rings. Again this system is relatively well defined in terms of the ring statistics, as shown by equation 5.5. As the accuracy of the percolation transition is dependent on the system size, very large networks were generated with 1×10^6 rings. This remained computationally tractable as the calculation of percolation requires only the node connectivities, not their positions, and so there is no need for geometry optimisation. Networks were constructed using targeted optimisation across the full spectrum of p_6 values and with α in the range $-0.1 \rightarrow 0.3$. For each state point, 100 networks were sampled starting from different random seeds.

The results of these simulations are presented in figure 5.8. Figures 5.8a, 5.8b show the evolution in P_∞ for selected α values across the p_k range for the 6- and 7-ring subgraphs, which demonstrate slightly different behaviours. Neglecting the effect of the Aboav-Weaire parameter initially, it can be seen that as the proportion of a given ring increases, the probability of a giant component forming increases. This is process is visualised in figures 5.9a-5.9c. In the case of the 6-ring subgraph, it can be seen that it bears similarity to the triangular site percolation problem discussed in section 5.4.1, with the percolation threshold oscillating around $p_6^c \approx 0.5$. The 7-ring case on the other hand displays a percolation threshold at a lower value of $p_7^c \approx 0.35$. This is intuitive as each node has a greater number of edges emanating from it, and so a greater probability of connecting to other ring sizes. It is also for this reason that there is *no* percolation threshold for the 5–ring subgraph in aG. This can be rationalised by realising that as $p_5 = p_7$, and a 7-ring by definition has more connections to adjacent rings than the 5-ring, there can never be a point where the 5-rings can form a giant component in preference to the 7-rings. In the most extreme case, one can see this in the example of haeckelite, in which $p_5 = p_7 = \frac{1}{2}$ and all 7-rings are connected, yet the 5-rings remain isolated from one another.

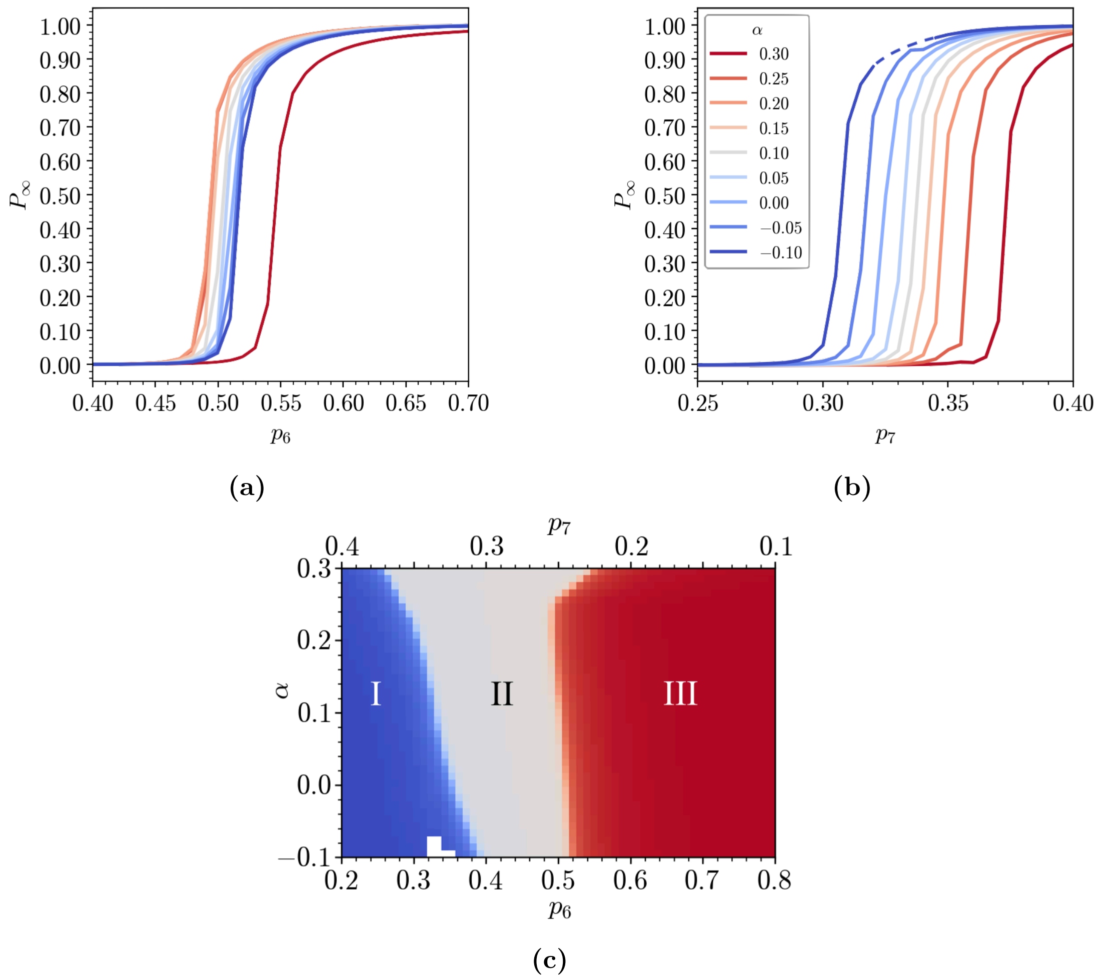
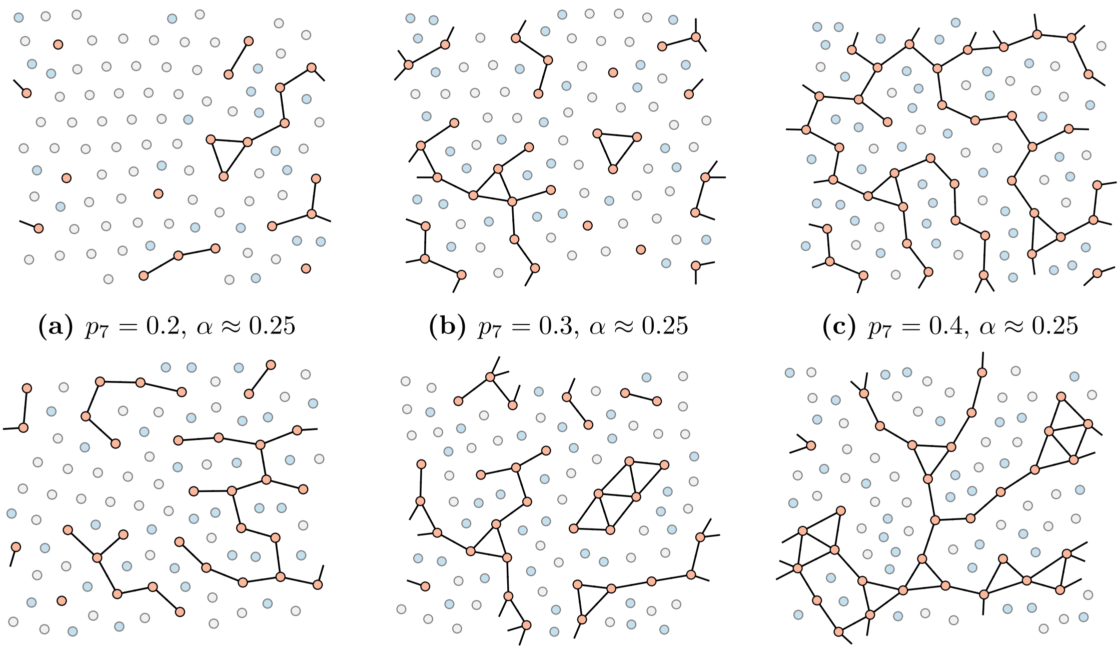


Figure 5.8: Percolation in aG configurations generated via targeted optimisation. Panels (a),(b) maps percolation as a function of α and p_6, p_7 respectively (dashed line indicates interpolated data). The percolation threshold is defined as when $P_\infty = \frac{1}{2}$. Panel (c) gives the phase behaviour of these systems: phase I contains a giant component in the 7-ring subgraph; phase II no giant components in any subgraph; phase III a giant component in the 6-ring subgraph.

The behaviour of the network percolation threshold is also subtly related to the node correlations, as expected [139, 140]. For the 7-ring subgraph, the percolation threshold in p_7 systematically decreases with decreasing α. This is because a decreasing α is reflective of increased large-large ring pairings, thus facilitating the formation of a connected giant component of 7-rings. This process is demonstrated in figures 5.9d-5.9f. The 6-ring shows what appears to be a more complex relationship with α. Initially as α is increased, the percolation threshold in p_6 decreased, before suddenly increasing again at high α. This is a consequence of the fact that the



(a) $p_7 = 0.2$, $\alpha \approx 0.25$ **(b)** $p_7 = 0.3$, $\alpha \approx 0.25$ **(c)** $p_7 = 0.4$, $\alpha \approx 0.25$

(d) $p_7 = 0.32$, $\alpha \approx 0.3$ **(e)** $p_7 = 0.32$, $\alpha \approx 0.1$ **(f)** $p_7 = 0.32$, $\alpha \approx -0.1$

Figure 5.9: The formation of giant components in disordered networks is a function of both the proportion of each ring size (here p_7) and the ring correlations (as measured by α). Panels (a)-(c) show the effect of increasing p_7 at constant α, with a giant component only forming in (c), once a sufficient number of 7-rings are present. Conversely panels (d)-(f) show the effect of decreasing α at constant p_7, with a giant component only forming in (f), once sufficient clustering of 7-rings is achieved.

6-ring is the "middle" ring size. Hence when α is strongly negative, $7-6$ pairings are most favoured and when α is strongly positive $6-5$ pairings are more abundant. It is only when α sits in the intermediate region that the $6-6$ ring correlations are maximised and percolation is most readily facilitated. It is interesting to note that this also around the value of $\alpha \approx 0.25$ that is also common in nature.

The results discussed above can be combined to draw a percolation phase diagram for aG, presented in figure 5.8c. In this diagram there are three phases:

- **Phase I**: exists at low $p_6 \lesssim 0.35$ and preferentially low α, where networks contain a giant component of 7-rings.

- **Phase II**: occupies intermediate values of p_6, where no subgraph contains a giant component.

- **Phase III**: encompasses the largest region of phase space, for $p_6 \gtrsim 0.5$, where networks contain a giant component of 6-rings.

From this phase behaviour it can be seen that relatively low values of p_6 must be achieved before the percolation of 6-rings is broken. In addition it is unlikely that an phase I could be experimentally realised and percolation of the 7-rings achieved. This is because from maximum entropy, the most disordered lattice possible would have $p_5 = p_6 = p_7 = \frac{1}{3}$ which is on the fringe of the percolation threshold for p_7, and would necessitate a value of α much lower than is currently seen experimentally. This could have implications when designing materials, for which there are eigenstates which are localised on specific ring sizes [141, 142].

5.5 Chapter Summary

An innovative method has been presented to generate two-dimensional materials with well defined topology. This targeted Monte Carlo search algorithm allows configurations to be constructed which have precise ring size distributions and ring-ring correlations. The advantage of this approach is that configurations can be produced rapidly with controllable properties; which may lie outside experimentally or physically accessible regions of phase space. These configurations may then be used as starting points for further investigations. For example, the algorithm outlined in this work has already been utilised to study the mechanical properties of vitreous silica under deformation [143–145]. In this chapter the targeted optimisation method was employed to probe the physical meaning of the Aboav-Weaire parameter. The effect of α on the ring structure has been quantified through partial RDFs. In addition the energetic minima for a range of systems has been shown to correspond well with values commonly found in nature. Finally, the method was employed in a study of the ring percolation in amorphous graphene, with the phase behaviour quantified in terms of the ring statistics and the Aboav-Weaire parameter.

6 | Generalisation of Disordered Physical Networks

The properties of a wide range of physical two-dimensional networks are investigated by formulating a generalised network theory. The methods developed are shown to be applicable to a wide range of systems generated from both computation and experiment; incorporating atomistic materials, foams, fullerenes, colloidal monolayers and geopolitical regions. The ring structure in physical networks is described in terms of robust measures from network science: the node degree distribution and the assortativity. These quantities are linked to previous empirical measures such as Lemaître's law and the Aboav-Weaire law. The effect on these network properties is explored by systematically changing the coordination environments, topologies and underlying potential model of the physical system.

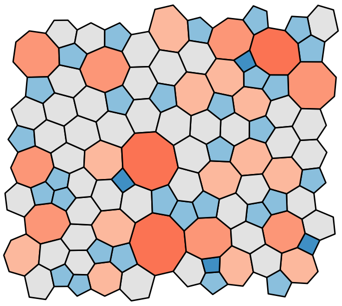
6.1 Two-Dimensional Networks in Nature

So far this thesis has focussed on 3-coordinate atomic networks such as silica and amorphous graphene. These atomic systems can however be considered a subset of a much larger class of two-dimensional networks which occur throughout the natural world (see figure 6.1). Such networks emerge across all disciplines and span many orders of magnitude in size. In physical sciences random tessellations are not restricted to atomic materials, but are observed in foams, crack-patterns (in dessicated films, ceramics *etc.*) as well as in colloidal films through the Voronoi construction - to name a few [32–36, 68, 146]. Similar mosaics can also be seen in the biological world in the form of epithelial cells and polymer networks such as collagen [97, 147–149], as well as in geology in the guise of rock formations and geography in context of geopolitical borders [37, 41, 45]. Whilst this last example may seem to fall into the category of seemingly more esoteric offerings from the literature (including for instance crocodile scales and oil paintings [40, 46]), it provides an interesting insight into the formation of tessellations through random point processes. Although man-made maps are nominally carefully constructed, the influence of random geographical features serve to generate tessellations which are entirely consistent with others found in the natural world.

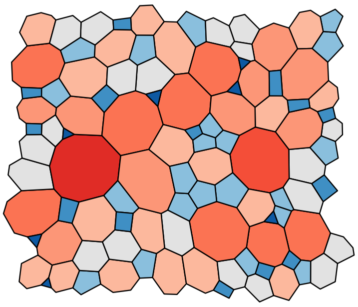
This is to say that the study of atomic networks fits into a wider remit of understanding the behaviours of generic physical networks. Similarly the techniques and theory used to model and characterise atomic networks can be readily deployed to understand a wide range of other complex physical systems. Therefore the focus of this chapter is on extending theory and computational methods to study general two-dimensional networks which are physically motivated (*i.e.* have an underlying physical potential model). To demonstrate the effectiveness and potential of this approach, results will be compared to those from a wide variety of experimental systems.

6.2 Generalised Network Theory

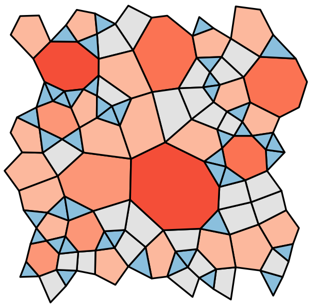
A consequence of the universality of two-dimensional networks is that both the language and the metrics used to describe then varies considerably between fields, as demonstrated in table 6.1. From a nano-materials perspective there are rings formed from a set of bonded atoms, in crystals there are grains separated by boundaries and in biological tissues cells which divide. Further complication may arise from the concept of graph duality, where ring structure emerges only after transforming the physical coordinates. In the context of colloidal monolayers for instance, rings are generated using the Voronoi construction; where the vertices have no real manifestation and the particle positions are the simplices in the dual Delaunay triangulation. In addition as seen in previous chapters, there remains a prevalence of empirical laws to describe their structure.



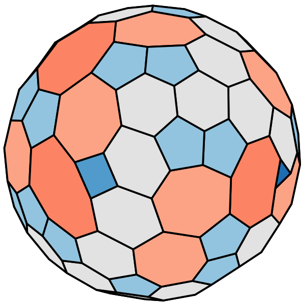
(a) Planar, $c = 3$, length controlled.



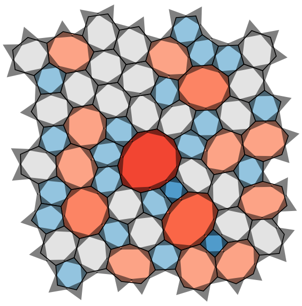
(b) Planer, $c = 4$, angle controlled.



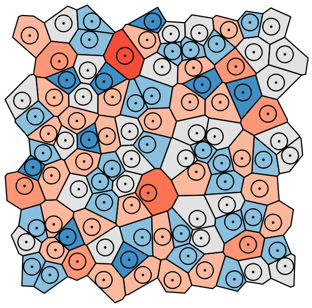
(c) Planer, $c = 4$, angle controlled.



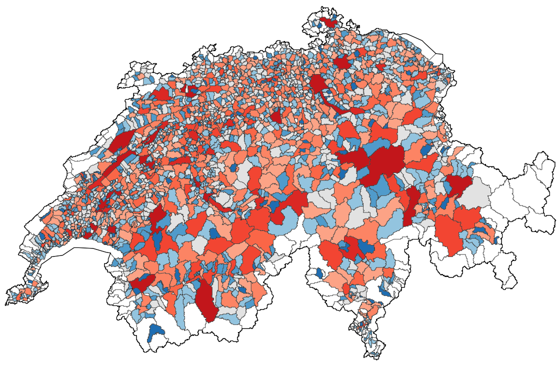
(d) Spherical, $c = 3$.



(e) Triangle raft.



(f) Hard disk Voronoi.



(g) Communes of Switzerland.

Figure 6.1: Two-dimensional networks emerge in diverse physical systems and span a range of length scales, coordination environments, and topologies: (a) 3-coordinate, bond-length controlled network, *e.g.*, glass; (b) 3-coordinate, angle controlled network, *e.g.* foam; (c) 4-coordinate network; (d) 3-coordinate network in spherical geometry, *e.g.* nonclassical fullerene; (e) triangle raft, *e.g.* silica bilayer; (f) hard disk Voronoi *e.g.* colloidal monolayer; and (g) communes of Switzerland. Rings are coloured similarly according to size with blue, gray, and red indicating smaller than, equal to, and greater than the mean ring size, respectively.

Network science offers an opportunity to unite the study of these disparate physical systems through a generalised theory. Much of the groundwork for this has been laid in chapter 2, but there are some important additions, namely the introduction of the assortativity to describe ring-ring correlations. Some of the

Table 6.1: Terminology to describe ring structure in literature reflects the diversity of the underlying physical systems.

Term	Synonyms and Examples
Ring	Face, polygon, cell, grain, pore, Voronoi cell
Network	Graph, tiling, packing, tessellation, partition, arrangement, decomposition, net, mosaic
Link	Edge, bond, boundary, interface
Node	Vertex, point, atom

key aspects which were introduced in chapter 2 will be briefly recapped, before these extensions are highlighted.

Chapters 5 and 4 of this thesis focussed on planar atomic systems which had a fixed coordination of three. The main difference in this chapter is that the scope has increased to include networks with variable coordination and topologies. For generic physical networks the equivalent of the atomic coordination number, c, is not necessarily precisely defined by an atomic species. The consequence of this is that the mean ring size as dictated by Euler's law is no longer always six, but rather determined by equations (2.11),(2.12); so that for example a network of $c = 4$ will have a mean ring size of $\langle k \rangle = 4$. That being said, the majority of naturally occurring networks still have $c = 3$, as higher order sites are unstable with respect to small perturbations, with for example a 4-coordinate site readily splitting into two 3-coordinate sites [58].

The ring statistics, p_k, remain an important measure, and have a clear analogue in network science, being the node degree distribution of the ring network (see section 2.1.2). The node degree distribution is still expected to follow Lemaître's maximum entropy distribution, provided the constraints are appropriately adjusted to reflect the mean node degree. Whilst all natural networks lie on the universal Lemaître curve, it will be seen in section 6.5.1, that the specific location of a given network is dependent on the underlying physics of the system.

The other empirical law heavily discussed in this thesis, the Aboav-Weaire law, is more problematic. Although it has proved useful in materials science, it is largely confined to this area, and is not without flaws. These flaws will be discussed in detail below, but they essentially arise from the empiricism of the law and the resulting difficulty in interpreting its results. However, network science has a well-adopted metric for measuring node degree correlations, termed the assortativity. This chapter therefore provides a good opportunity to replace the empirical Aboav-Weaire law with a concrete measure, and it will be shown in section 6.2.2 that there is a mapping between the Aboav-Weaire parameter and the assortativity.

6.2.1 Deficiencies in the Aboav-Weaire Law

For all its perceived success in characterising amorphous materials, the Aboav-Weaire law suffers from several deficiencies, some of them academic and others practical. To begin with, it remains the case that despite numerous efforts [150–155] there is no satisfactory theoretical justification behind the Aboav-Weaire law; the various attempts and their drawbacks summarised excellently by Mason *et al.* [156]. In fact it seems increasingly likely that the difficulty in finding an adequate theoretical proof for the Aboav-Weaire law simply arises from the fact that there just isn't a strong physical basis behind it.

One may then reasonably question why the linear Aboav-Weaire law holds so well for a range of different systems. The answer again may be the fact that unfortunately it is not as infallible as its widespread usage would suggest. In particular the assumption that the law holds and is indeed linear is often overlooked. This is not in reference to somewhat contrived examples, such as regular crystalline arrangements, as the Aboav-Weaire law is a really a comment on disordered systems [31]. Even "conventional" examples often show deviations [32, 151, 157]. These manifest in two ways. Firstly the data may not be linear over the whole range. This size of this effect can be understated, as such deviations from linearity occur in the tails of the ring distribution at low or high k, where the discrepancy is often attributed to poor sampling statistics. Nevertheless, as Mason *et al.* astutely point out *"a linear model is a good approximation of any smooth function over a small domain, and that the success of the law of Aboav-Weaire does not necessarily indicate that the average excess curvature is actually linear"*. The second issue is that little attention is paid to the exact form of the law and the fact that the intercept should be $\langle k\rangle^2+\mu_2$ is rarely adhered to. Enforcing this condition often leads to a less satisfactory fit.

The consequences of the difficulty in obtaining an accurate Aboav-Weaire fit are naturally that the resulting α parameter has associated with it a degree of "greyness". Yet even in the case where the Aboav-Weaire law seems wholly appropriate, there is still a difficulty in fully interpreting its meaning. It is not intuitive what the actual value of α represents nor its limits. Even for a simple system of $\{5,6,7\}$ rings, equation (5.6) illustrated that the relationship between ring structure and α is non-trivial. More generally, it was shown in section 2.2.3, if rings are arranged purely randomly that $\alpha=-\frac{\mu_2}{\langle k\rangle^2}$, but without a well-defined upper limit for comparison interpretation remains restricted. This equation also highlights that α is dependent on the ring statistics and that its sign is an insufficient classifier for positive or negative correlation. Hence even if a high quality fit is achieved, a combination of these effects make it difficult to draw accurate comparisons between different systems.

This is not to say that the Aboav-Weaire law does not have value, and certainly the general observation is extremely interesting, even if the underlying relationship is more complex than originally suggested. It is more to point out that there is scope to improve the quantification of the effect and that a robust approach which is applicable to diverse systems will be required to study generic two-dimensional networks.

6.2.2 Assortativity as a Measure of Ring Size Correlations

The assortativity was introduced by Newman to measure the preference of low degree nodes to be adjacent to high degree nodes in generic networks [158]. It has proved highly popular in the network science and the study of social and biological networks [159], but has also been applied for example in theoretical studies of hard disk packings [160]. The calculation of the assortativity revolves around the edge joint degree distribution, e_{jk}, which measures the probability of two nodes of degrees j,k sharing a link (*i.e.* two rings of sizes j,k being adjacent). The probability of any link having degree k is distributed according to $q_k = kp_k/\langle k\rangle$, and so if nodes are randomly arranged $e_{jk} = q_j q_k$. Deviation from this random arrangement is the assortativity, and can be measured by Pearson's correlation coefficient:

$$r = \frac{\sum\limits_{jk} jk\left(e_{jk} - q_j q_k\right)}{\sum\limits_{k} k^2 q_k - \left(\sum\limits_{k} k q_k\right)^2} = \frac{\langle k\rangle^2 \sum\limits_{jk} jk e_{jk} - \langle k^2\rangle^2}{\langle k\rangle\langle k^3\rangle - \langle k^2\rangle^2}. \tag{6.1}$$

For this coefficient to be calculable, the second and third moments of the degree distribution must be finite [161]. This condition is satisfied for these physical systems, as the proportion of large rings quickly becomes vanishingly small. Include some plots of matrices here?

The advantages of adopting this measure of assortativity are clear. The correlation coefficient is bounded between $-1 \leq r \leq 1$ and has three well defined limits: $r = 0$ indicating a random network, $r = 1$ a perfectly assortative network and $r = -1$ a perfectly disassortative network. This allows physical networks to be readily compared in a way that the Aboav-Weaire law does not allow. Physical networks can now be fitted in to the wider field of network science, introducing them as important examples alongside more traditionally studied networks. Using the assortativity also provides a natural extension to higher dimensions, which has been difficult to reconcile with the empirical Aboav-Weaire law [156].

For completeness, it will be shown that the assortativity can be related to the Aboav-Weaire parameter. This can be achieved by using the fact that that the mean node degree about a j-degree node is given in equation (5.3) as $q_j m_j = \sum\limits_k k e_{jk}$. Substituting this expression into equation (6.1), and assuming the Aboav-Weaire law (2.20) holds *exactly*, it can be shown that:

$$\alpha = -\frac{r\left(\langle k\rangle\langle k^3\rangle - \langle k^2\rangle^2\right)}{\mu_2\langle k\rangle^2} - \frac{\mu_2}{\langle k\rangle^2}, \tag{6.2}$$

which is consistent with the topological gas, when $r = 0$. In reality, the Aboav-Weaire fit is never perfect, and so equation (6.2) provides an approximation to the value of α. The accuracy of this equation will therefore depend on the applicability of the linear fit. Include that figure somewhere .

The assortativity also provides a natural framework to extend Lemaître's maximum entropy arguments to factor in ring adjacencies. The entropy is first defined in terms of the edge joint degree distribution, as $S = -\sum_{jk} e_{jk} \log e_{jk}$. Considering e_{jk}, the following constraints must hold:

$$\sum_{jk} e_{jk} = 1 \tag{6.3}$$

$$\sum_{jk} k e_{jk} = \frac{\mu_2}{\langle k \rangle} + \langle k \rangle \tag{6.4}$$

$$\sum_{jk} \frac{1}{j} e_{jk} = \frac{1}{\langle k \rangle} \tag{6.5}$$

$$\sum_{jk} jk e_{jk} = c\,(r)\,; \tag{6.6}$$

resulting from the normalisation condition, Weaire's sum rule [19] and Euler's formula and finally a constraint imposing the assortativity from equation (6.1). As for Lemaître's law, Lagrange's method can be used with the constraints above (noting that $e_{jk} = e_{kj}$) to generate a maximum entropy joint distribution which satisfies:

$$e_{jk} = \frac{e^{-\frac{\lambda_1}{2}(j+k) - \frac{\lambda_2}{2}(1/j+1/k) - \lambda_3 jk}}{\sum_{jk} e^{-\frac{\lambda_1}{2}(j+k) - \frac{\lambda_2}{2}(1/j+1/k) - \lambda_3 jk}}, \tag{6.7}$$

and equations (6.4)-(6.6). This can again be solved numerically, and the resulting distribution can be related to a single node degree probability (*e.g.* p_6) and an assortativity value.

6.3 Generalised Bond Switching Algorithm

In order to study generic physical networks, a simulation method is required which can generate configurations across a wide range of coordination environments, topologies and potential models. The bond switching algorithm, introduced in section 3.2, is a good candidate as it has proved effective for studying atomic networks in chapter 5. However, currently it is only adapted to study constant coordination planar systems (in this work 3-coordinate but there is one previous example of study of 4-coordinate systems [162]). Therefore a further natural

Table 6.2: List of starting crystalline lattices for bond switching for a range of coordination environments, and the corresponding mean ring size.

Topology	x_3	x_4	$\langle k \rangle$	Lattice
Planar	1	0	6	Hexagonal
Planar	0	1	4	Square
Planar	2/3	1/3	5	Cairo
Planar	x_3	x_4	$4 \rightarrow 6$	Mixed Hexagonal-Square
Spherical	1	0	~ 6	12-Pentagon Fullerene
Spherical	0	1	~ 4	8-Triangle Fullerene

extension of the bond switching method is presented here, to variable atomic coordination environments and overall system topology.

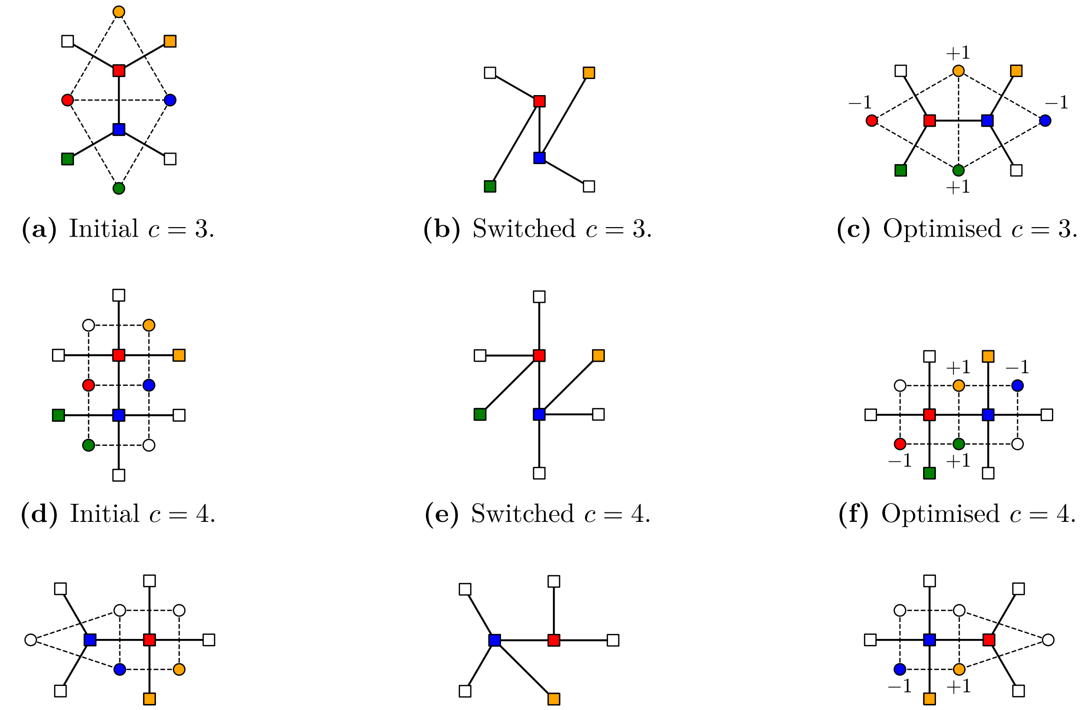
As a review from a networks perspective, the bond switching algorithm is a stochastic sampling method. Starting from an initially well-ordered network, links between neighbouring nodes are switched and the effect on the potential energy of the system (*i.e.* the amount of strain which is introduced or removed) is calculated. The energy of the system is determined by the potential model, which expresses the total energy of the network as a function of all node positions. After the links between nodes are switched, geometry optimisation of the node positions takes place to minimise the total potential energy. By incorporating switches which reduce the potential energy of the network with greater probability, one can bias the search towards networks of lower energy and therefore which occur more commonly in nature. The specificities of the algorithm will however depend on the exact nature of the system in question.

6.3.1 Extension to Variable Coordination

The choice of the starting lattice can be used to determine the system properties *i.e.* the atomic coordination environments and topologies (table 6.2, figure 2.3). This is because in the bond switching algorithm the node degree distribution of the atomic network is constant, and hence from equation (2.10) so is mean node degree of the dual network. Therefore whichever topology, atomic coordinations and mean ring size the system is initialised with will be preserved throughout the simulation.

The bond switching move will then vary depending on the coordination properties, as outlined in figure 6.2. Figures 6.2a-6.2c show the original move, which was designed for purely 3-coordinate atoms, and is in effect introducing a Stone-Wales defect. This move augments the ring size of two rings and decrements two others, preserving both the mean ring size and the coordination number of the individual atoms involved in the transformation. The changes in ring size (equivalent to the changes in node degree of the dual network) are highlighted in the figure as "±1". The extension to 4-coordinate atoms (figures 6.2d-6.2f) is relatively straightforward, simply involving extra spectator atoms, but for mixed coordination it is subtly different (figures 6.2g-6.2i). For the both systems the local ring sizes are again changed by ±1 (as highlighted, and preserving the mean ring size). However, whereas for the pure systems the switch move must be coordination preserving, for mixed coordination systems this prevents true melting. This can be countered by using a move in which the coordinations of neighbouring atoms are exchanged, whilst maintaining a constant mean ring size.

The thermalisation of the initial lattice requires a large number of random moves as described above, the purpose being for the system to "forget" all memory of the original ordered lattice. To ensure the lattice is fully randomised, observables such as the second moment of the ring sizes and assortativity can be monitored. For mixed lattices it is also important that the variously coordinated atoms are adjacent to the number of others as expected from pure chance, namely the binomial expansion of $(3x_3/\langle k\rangle + 4x_4/\langle k\rangle)^2$. For the potential model, as discussed in section 3.2.2, a simplified Keating (SK) potential can be effectively employed, with the option of being augmented with a restricted bending (ReB) potential. To capture the possibility of variable coordination environments, the equilibrium bond length was set equal for all interaction types and the equilibrium angles were set to $2\pi/c$ for c-coordinate atoms.



(a) Initial $c = 3$. (b) Switched $c = 3$. (c) Optimised $c = 3$.

(d) Initial $c = 4$. (e) Switched $c = 4$. (f) Optimised $c = 4$.

(g) Initial $c = 3, 4$. (h) Switched $c = 4, 3$. (i) Optimised $c = 4, 3$.

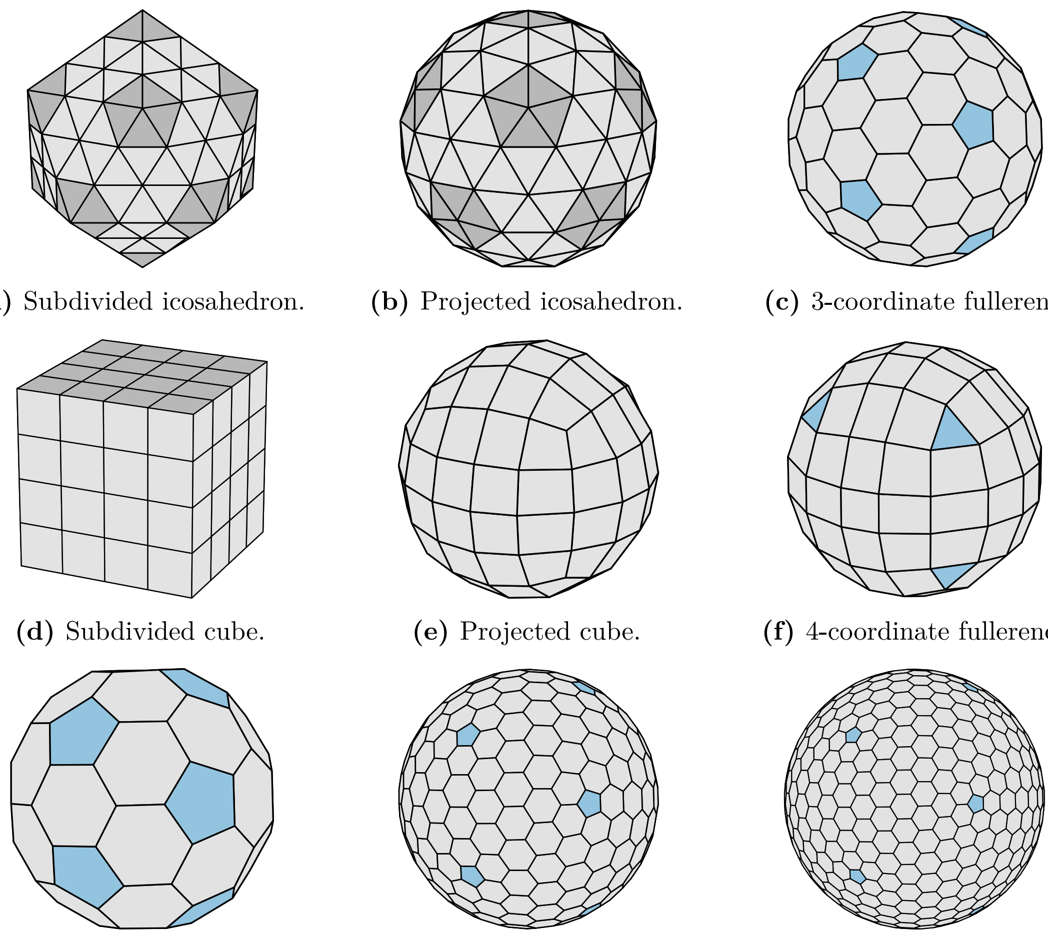
Figure 6.2: Bond switching Monte Carlo moves for different atomic coordination environments: 3-coordinate sites (a)-(c), 4-coordinate sites (d)-(f) and mixed 3/4 coordination (g)-(i). For each coordination type the atomic connectivity is shown for the starting structure (left), the initial switched structure (middle), and a geometry optimised switched structure (right), via the squares and solid lines. The effect on the dual network (circles and dashed lines) is also demonstrated, with the numbers indicating the change in node degree after the move is applied. Colouring is used as a guide for the eye, to track changes between the pre- and post-switch configurations.

6.3.2 Extension to Variable Topology

As a demonstration of the general applicability of the bond switching algorithm, configurations can also be generated in spherical topology. In order to do this an initial crystalline fullerene-like structure must be generated. A practical method to do this is from the dual lattice of a platonic solid, as outlined below and illustrated in figure 6.3.

1. Construct an icosahedron for 3-coordinate networks or a cube for 4-coordinate networks and subdivide the faces into triangles or squares respectively (figures 6.3a,6.3d).

2. Project the lattice onto a sphere (figures 6.3b, 6.3e).



(a) Subdivided icosahedron. (b) Projected icosahedron. (c) 3-coordinate fullerene

(d) Subdivided cube. (e) Projected cube. (f) 4-coordinate fullerene.

(g) 42-ring fullerene. (h) 252-ring fullerene. (i) 492-ring fullerene.

Figure 6.3: Lattices of given coordination in spherical topology can be generated from the dual of a subdivided platonic solid, projected onto a sphere. Figures (a)-(c) show this process for the icosahedron, giving a 3-coordinate lattice and (d)-(f) for the cube, giving a 4-coordinate lattice. Altering the number of subdivisions allows the number of rings in the fullerene to be controlled.

3. Generate atomic network from the dual lattice (figures 6.3c,6.3f).

Once this lattice has been constructed, the bond switching algorithm may proceed as usual.

A further question that must be addressed is how to handle the potential model in spherical topology. One could implement all the potentials using spherical polar coordinates, which would strictly enforce the system topology. However, here a simpler solution was used, whereby a simple harmonic restraining potential was between all atoms and the surface of the sphere, and the standard potential was used in three dimensional Euclidean space. The atoms are therefore approximately maintained on a sphere of a fixed radius. The radius is selected before the bond switching routine commences, corresponding to the minimum energy structure for the initial fullerene.

It is noted here that extensions to other topologies are certainly possible, albeit with varying degrees of difficulty. All that is required is generation of a lattice which satisfies the underlying topological constraints, and an adequate potential model. For example, a relatively easy extension would be to toroidal topology. As previously mentioned, the periodic two-dimensional lattice has the same topology as the torus, and so there is a trivial mapping between the two (which is also the case for a Möbius strip and Klein bottle, although these currently seem less chemically relevant). Application to systems with a larger number of topological holes would however require a different method to generate the initial lattice.

6.4 List of Studied Networks

In this chapter data on networks will be presented from a range of different sources, covering both computation and experiment. These are detailed here as a reference for the remainder of the chapter.

6.4.1 Computational Networks

All networks derived from computation were calculated using methods described in this thesis. They are as follows:

1. **Bond switching**: networks of 1024 rings for $\langle k \rangle = 4, 6$ and 1152 rings for $\langle k \rangle = 5$. In a simulation, the system was thermalised with 2×10^5 random moves, and annealed over a further 4×10^6 moves. A series of different potential models were also used with bond length/angle force constant ratios of $k_r/k_\theta = 16, 4, 1, 1/4$. For each parameter set, 100 simulations were run starting from different random seeds.

2. **Triangle rafts**: networks of 1000 rings across a temperature range of $T = 10^{-4.5} \to 10^{-1.5}$, as outlined in section 4.3, totalling some 27,500 configurations.

3. **Hard disk Monte Carlo**: systems of 1000 disks at packing fractions in the range $\phi = 0.0 \rightarrow 0.77$. Each simulation comprised cycles of 1000 random displacement moves, with 10^5 equilibration cycles, 10^5 production cycles and with sampling every 10 production cycles. For each packing fraction 10 simulations were run using a different random seed. A Voronoi analysis was performed for each configuration to generate a system of tessellating rings, as discussed in section 3.3.

6.4.2 Experimental Networks

Comparison is also made to a variety of experimental networks obtained from a variety of publications. They are as follows:

1. **Colloidal monolayers**: particles with a diameter of $\sim 2.79\ \mu$m dispersed in a water-ethanol mixture and confined by gravity to form a monolayer on the base of a glass sample cell, with packing fractions in the range $\phi = 0.29 \rightarrow 0.66$ [50]. Out-of-plane fluctuations are quantified by the gravitational height of the particles, which is a very small percentage of their diameter, and as such the system is structurally two-dimensional. Each packing fraction has 100 associated frames, with the time between frames around 10s. At the highest packing fractions considered, the area of the system imaged contains around 3000 particles. As the system is aperiodic, after Voronoi analysis the cells on the image boundary are neglected.

2. **Silica bilayers**: configurations of thin films of silica glass grown on graphene [4] and Ru(001) [11]. Three samples were obtained consisting of 291, 444 and 446 rings.

3. **Amorphous graphene**: configurations of graphene amorphised by irradiation with an electron beam [163]. Exposure to increasing doses created 14 samples with varying levels of disorder. For each sample, defects were identified from the presence of under-coordinated atoms, arising largely from

the sample perimeter or from holes in the centre, which were removed. After processing, configurations comprised $\sim 3000 \rightarrow 5000$ rings.

4. **Geopolitical regions**: physicists have previously studied the regions of France and Ireland, and noted the similarity in their properties to materials [41, 114]. This tradition has been continued by analysing five further maps: the communes of Switzerland, the parishes and Westminster constituencies of Great Britain and the socio-economic regions of the European Union (EU) and the European Free Trade Association (EFTA) (including both current and candidate countries at the time of writing) [164–166]. Details of the analysis can be found in appendix appendix on maps , and a summary of the results in table link .

6.5 Investigations into Generic Physical Networks

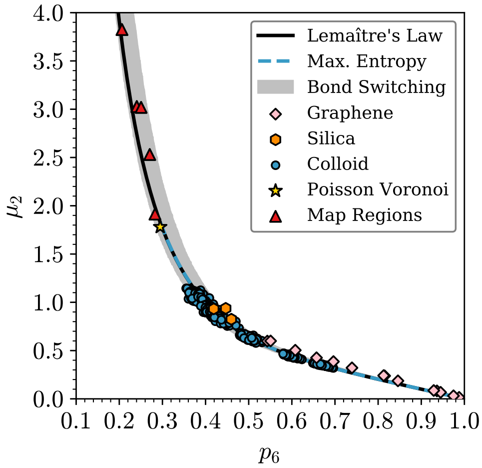
The properties two-dimensional atomic networks are now discussed alongside generic physical systems introduced in this chapter. This primarily focusses on the network properties introduced in chapter 2, but also covers a case study on the energetics of a 92-ring fullerene.

6.5.1 Degree Distributions

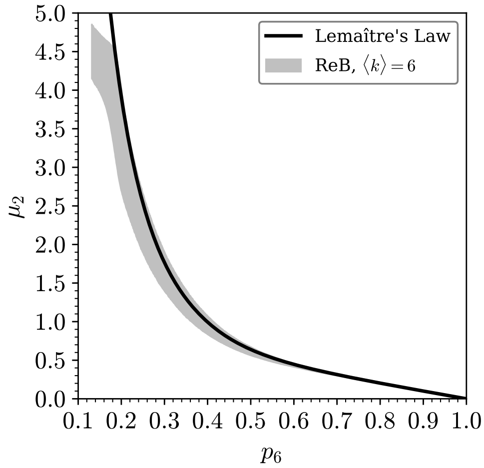
The degree distributions of physical networks are discussed in terms of Lemaitre's law; with the distribution variance, μ_2, plotted against the proportion of rings of mean size. Figure 6.4a presents these data for a range of 3-coordinate systems comprising experimental, computational and theoretical. There are many things to note, but primarily it can be seen that regardless of the nature of the underlying system, all data fit very well with the maximum entropy solution provided by Lemaître's law. Whilst Lemaître's law highlights the similarities between these systems, it is also important to examine some of their differences. For example, what determines where a system sits on the Lemaître curve *i.e.* what controls the number of hexagons? For materials this is based on energetics - the strain associated with bond and angle distortions. For instance, experimentally silica bilayers have more diverse ring statistics than graphene owing to the reduction in ring strain due to the presence of oxygen linkages[11]. Even for the graphene samples which have been modified by an electron beam (pink diamonds), the disorder does not approach that of the silica glasses (orange hexagons). For the colloid systems (blue circles) however, the rings are formed from the Voronoi tessellation, with no intrinsic cost to distortions and instead it is the packing fraction, ϕ, which determines p_6. The limit $\phi \to 0$ achieves the Poisson Voronoi ring distribution (yellow star) [120], with a lower bound of $p_6 \approx 0.295$. For for the administrative geopolitical regions, there is no energy cost for rings, regardless of shape, convexity or separation, and so we find these points (red triangles) in the low p_6, high entropy portion of Lemaître's curve.

On the other hand, using a flexible computational method allows access to the entire range of μ_2 values, where the level of disorder is controlled by the Monte Carlo "temperature" parameter. The results from bond switching highlight the typical dispersion that can be expected within Lemaître's law, with the grey shaded region indicating the bounds of μ_2 within two standard deviations of the mean. Finally it can be seen that using equation (6.7) with the p_6 and r values from hard disk Monte Carlo (blue dashed line) reproduces the results from Lemaître's law without the need for the empirical constraint. The calculation of this line is explained at the end of section 6.5.2.

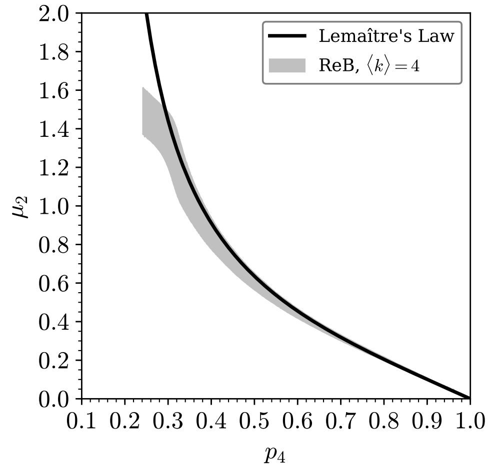
The effect on the maximum entropy solutions can also be explored for different atomic coordination environments and constraints. Figures 6.4b and 6.4c gives two such examples where ring convexity is enforced by using the ReB potential. Figure 6.4b gives results for a purely 3- coordinate system, $x_3 = 1$, whilst figure 6.4c gives results for a purely 4- coordinate system, $x_4 = 1$. The maximum entropy solution each case is again given by equation (2.17), with $\langle k \rangle = 6, 4$ respectively. The value of μ_2 is very similar for $\langle k \rangle = 4$ and $\langle k \rangle = 6$ above $p_{\langle k \rangle} \approx 0.5$. This is because in this region rings of sizes $k = \langle k \rangle$ and $k = \langle k \rangle \pm 1$ dominate the distribution and so $\mu_2 \approx 1 - p_{\langle k \rangle}$. However, as the value of $p_{\langle k \rangle}$ is reduced further, the two maximum entropy curves begin to diverge as the $k = \langle k \rangle - 2$ ring becomes accessible only to the 3- coordinate system, which in turn facilitates the presence of higher order rings. In



(a) SK potential, $c = 3$



(b) ReB potential, $c = 3$



(c) ReB potential, $c = 4$

Figure 6.4: Physical networks are shown to agree well with Lemaître's law. Panel (a): Lemaître's law (black line) is compared to to bond switching simulations of 3-coordinate two-dimensional materials (grey area representing two standard deviations from the mean), amorphous graphene (pink diamonds), silica bilayers (orange hexagons), colloidal monolayers (blue circles), the Poisson Voronoi diagram (yellow star) and maps of geopolitical regions (red triangles). Panels (b) and (c): comparison to bond switching with ring convexity maintained using the ReB potential for 3- and 4-coordinate systems respectively.

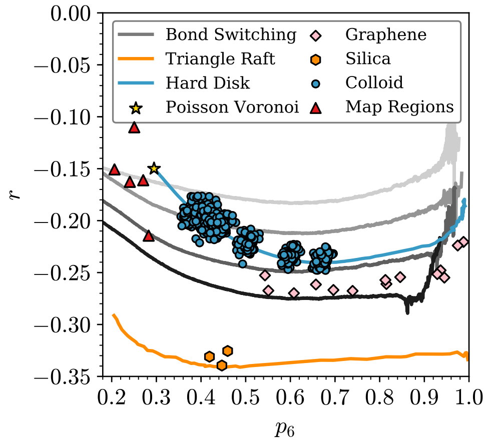
both cases, the results from bond switching both begin to deviate from the analytical results of Lemaître's law at low $p_{\langle k \rangle}$. The origins of this deviation can be traced back to the fact that if ring convexity is strictly maintained, it becomes increasingly difficult to accommodate the very large rings required to achieve large μ_2 values.

6.5.2 Assortativity

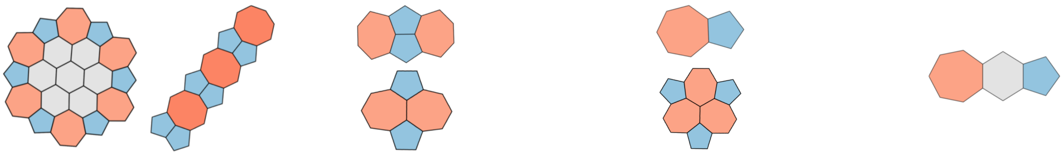
The ring correlations as measured by the assortativity are given for all 3- coordinate systems in figure 6.5a. It is found that all these 3- coordinate networks are disassortative and lie in the region $-0.35 < r < -0.10$ and that curves display a similar characteristic shape. The experimental colloid samples are well described by the hard disk model (blue points and line), with $p_6 \approx 0.84$ corresponding to packing fractions above the freezing transition limit ($\phi \approx 0.70$) [167]. The curves generated from bond switching and triangle rafts display different assortativities which depend on the balance of the length- and angle-drive forces. The driving force for the hard disk model is purely entropic, whereas for the other methods there is also a complex energy landscape, which may favour specific assortativities and which can be "tuned" by altering the balance of the interactions. For example the bond switching results show the effect of varying this balance with $k_r/k_\theta = 16, 4, 1, 1/4$ (black to light grey lines), leading to a shifting in the assortativity curves. This is supported by the experimental results from amorphous graphene (pink diamonds), which lie in the between the two curves with the largest bond length to angle force constant ratio, as would be intuitively expected for atomic systems and from empirical potential models [99].

For all the systems we note that there are different regimes, with the high p_6 limit corresponding to configurations best described as crystalline with defects rather than truly amorphous as in the low p_6 limit - with the two often being linked by a phase transition. The high p_6 limit can be rationalised by considering the frequency of common defect types at infinite dilution in a hexagonal lattice [20, 168] (figure 6.5b-6.5e). These can be calculated by considering the explicit edge joint probability distribution for a specific defect. For example, for the Stone-Wales defect 6.5c, each 5-ring has two 7-ring neighbours, and each 7-ring two 5-ring and one 7-ring neighbours such that:

$$\mathbf{e} = \begin{array}{c} \\ \\ \\ \\ \end{array}\begin{array}{cccc} & 5 & 6 & 7 & \\ & \left[\begin{array}{c} 0 \\ 3\delta \\ 2\delta \end{array}\right. & \begin{array}{c} 3\delta \\ 1-19\delta \\ 4\delta \end{array} & \left.\begin{array}{c} 2\delta \\ 4\delta \\ \delta \end{array}\right] & \begin{array}{c} 5 \\ 6 \\ 7 \end{array} \end{array} \tag{6.8}$$



(a) Assortativity.



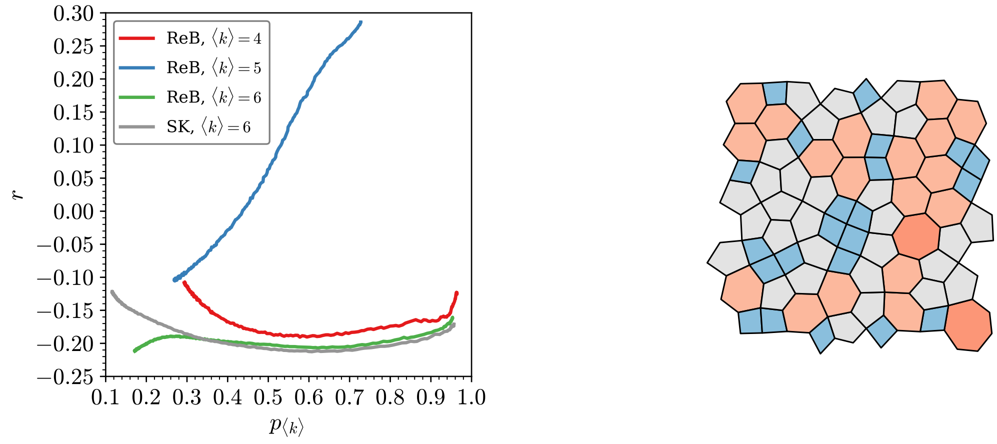
(b) $r = -0.\dot{3}$ **(c)** $r = -0.25$ **(d)** $r = -0.1\dot{6}$ **(e)** $r = 0.0$

Figure 6.5: Panel (a): variation in assortativity, r, against p_6 for a range of 3-coordinate systems comprising experimental and simulation data. For bond switching data, darker grey colouring indicates a greater k_r/k_θ. Panels (b)-(e) show common defects found in crystalline systems, with their limiting assortativity value (b) isolated pair, $r = 0.0$; (c) adjacent pair, cluster, $r = -0.1\dot{6}$; (d) Stone-Wales, mitosis, $r = -0.25$; (e) 5-7 chain (flower defect), 5-8 chain, $r = -0.\dot{3}$.

where $\delta = (1 - p_6)/12$. From here it is straightforward to evaluate the dilute p_6 limit as $\lim_{\delta \to 0} r = \frac{1}{4}$.

This helps to rationalise the high p_6 disassortative behaviour for these 3-coordinate systems. For hard disks as $p_6 \to 1$ the adjacent pair appears to be dominant, whereas for bond switching and triangle rafts the potential model determines the balance of defect types. For bond switching the standard deviation is large as each sample contains a single defect corresponding to one of the low energy forms. By visual inspection, increasing the length relative to the angle driving force preferences chain-like structures over isolated defects. Similarly the rigidity of the triangle units in the triangle raft method leads to a very tight length distribution which encourages the formation of defects such as 6.5e. As p_6 decreases more defects are introduced and the system becomes truly amorphous. Again one can posit that as the hard disk model has no energetic term, it is able to incorporate less correlated defects, and in the low packing fraction the hard disk model provides an estimate for the Poisson Voronoi limit of $r \approx -0.15$.

Again the effects of coordination environment and potential model on assortativity in complex networks can be demonstrated using bond switching. Figure 6.6a shows such a comparison, where the assortativity is plotted against the primary ring size for different coordination environments, averaged by Monte Carlo temperature. The effect of imposing a hard constraint on ring convexity can be seen through the two curves corresponding to $\langle k \rangle = 6$. These curves show very similar behaviour for $p_6 \gtrsim 0.3$, below which there is increasing deviation. This is as expected given the violation of ring convexity will only occur for very large rings at high temperatures, which can undergo deformation to reduce bond angle strain. This allows larger rings to pack next to each other, reducing the disassortativity. The behaviour of the pure 4-coordinate system, $\langle k \rangle = 4$, is qualitatively the same as for the 3-coordinate network, and indeed all the defects in figure 6.5 have analogues in 4-coordinate networks. The network of greater interest is that with mixed 3- and 4-coordinate vertices, corresponding to $\langle k \rangle = 5$. In this case one can see fundamentally different properties as these networks are assortative at high p_5, in contrast to limiting pure coordination cases. This assortative mixing is readily explainable through energetic considerations. The hexagonal and square tilings are strainless and so the disruptive effects of any defect rings is minimised when such rings are adjacent. Unlike the hexagonal and square lattices, the Cairo lattice is not strainless, due to a distortion in one of the edge lengths in the pentagonal tiles. Therefore, any 4- or 6-ring defects experience a driving force to cluster into the low energy regular tilings. In effect the lattice de-mixes into Cairo, square and hexagonal regions (as in figure 6.6b), which can be identified as inherently assortative behaviour. It is for this same reason that the limit of $p_5 \to 1$ cannot be reached, as the minimum energy lattice will be a mixture of the square, hexagonal and Cairo lattices, the exact proportion of which will depend on the potential model.



(a) **(b)**

Figure 6.6: Panel (a) shows the variation in assortativity, with ring statistics for 3-coordinate (green, grey lines), 4-coordinate (red line) and mixed 3/4-coordination systems (blue line) using the simplified Keating (SK) and restricted bending (ReB) potentials as indicated. Panel (b) gives a small configuration of a mixed coordination lattice displaying clustering of rings of similar size.

Finally the accuracy of the extension to Lemaître's maximum entropy method in equation (6.7) is assessed. Calculation of the maximum entropy joint degree distribution requires two parameters, p_6 and r, but the resulting distribution contains all the information required to calculate ring statistics, p_k, and the mean ring size about each ring, m_k. This has been performed using the parameters of p_6 and r from hard disk simulations. As demonstrated in figure 6.4a, the ring statistics calculated in this way regenerate those from Lemaître's law. In addition, plots of the mean ring sizes for selected packing fractions are given in figure 6.7. Whilst the fit is not perfect, this method does provide a close approximation to the hard disk results, particularly in the vicinity of $k = \langle k \rangle$. The results are especially good in the context that only two variables are required in p_6 and r to generate the distributions.

6.5.3 Energetics of Fullerenes

As an illustration of the generalisability of the methods described in this work, results are presented for two-dimensional networks in spherical topology. Such systems are also of experimental interest, as experimentalists now have access to "non-classical" fullerenes[169–172], metal-organic nano-cages[173, 174] as well as

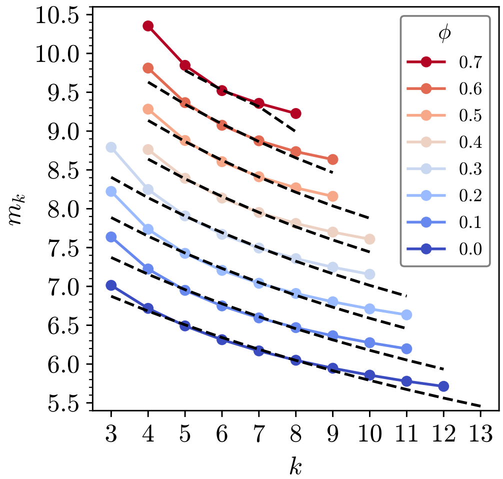
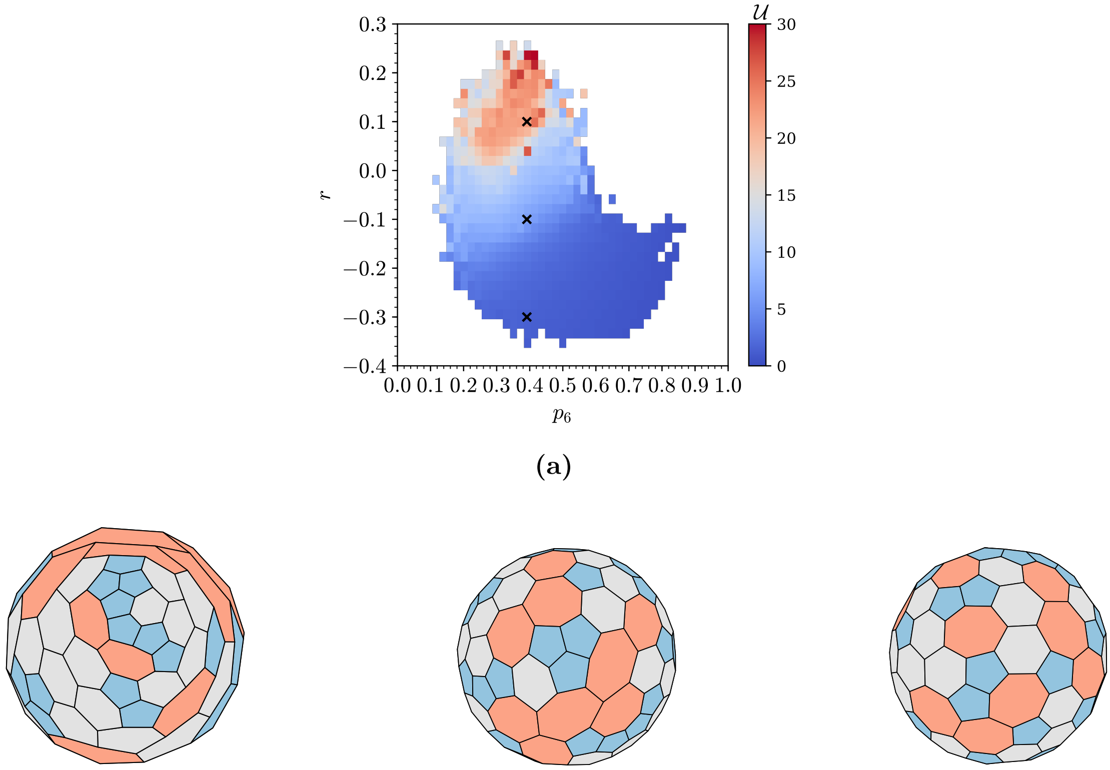


Figure 6.7: Mean ring size of hard disk simulations at different packing fractions (full lines) compared to results from maximum entropy (dashed lines). In both cases only ring sizes with $p_k > 10^{-4}$ are displayed and results are offset by 0.5 along the abscissa for clarity.

curved froths [175]. One such fullerene was investigated here: a 92-ring 3-coordinate fullerene consisting of 5,6 and 7- rings. Possible configurations were again generated via bond switching, starting from the lattice depicted in figure 6.3c. Here 10^6 total configurations were sampled from 100 different simulations, with $k_r/k_\theta = 4$.

Results of the network properties averaged across configurations are given in figure 6.8a, coloured by potential energy. In this plot the value of p_6 is discretised, owing to the the small and well-defined number of rings, and cannot exceed the upper limit imposed by the 12-pentagon rule, whereas the assortativity is averaged. As expected, the energy of the fullerenes increases with the increasing diversity in the ring statistics, as more pentagons and heptagons are accommodated. However, it is also the case that the arrangement of the rings, as measured by the assortativity, is also very important in determining the stability of the networks. To emphasise this, three example configurations are provided in figure 6.8b-6.8d. These amorphous fullerenes have the same p_6 value (and therefore p_5, p_7), but very different strain energies. In 6.8d defects appear which are similar to the common motifs as in figure 6.5 *i.e.* those associated with being low energy. The increased clustering of similar sized rings in 6.8b,6.8c leads to increasingly irregular ring geometries that generate high levels of strain. As previously noted with planar networks, systems which are



(a)

(b) $p_6 = 36/92$, $r = 0.1$

(c) $p_6 = 36/92$, $r = -0.1$

(d) $p_6 = 36/92$, $r = -0.3$

Figure 6.8: Panel (a) gives a map of fullerene stability as a function of ring statistics and assortativity. Potential energy increases as more pentagons and heptagons are accommodated, but is also strongly related to their arrangement as shown by the value of the assortativity, r. Panels (b)-(d) give three example fullerenes with the same $p_6 = 36/92$ but different assortativities of $r = 0.1, -0.1, -0.3$, respectively and as highlighted by the crosses in panel (a).

disassortative are energetically favoured. Although this is a simple consequence of the mechanical properties of the system, neglecting any electronic contributions, such is the difference in stability that we would expect disassortative fullerenes of this type to be more prevalent in nature.

6.6 Chapter Summary

In summary, this chapter has thoroughly examined the network properties of a wide range of naturally occurring two-dimensional systems; spanning varying coordination environments, potential models and topologies. Data has been collected from a range of experimental sources, and have the theoretical bond switching method has been further developed to aid the study of these diverse systems computationally. These data have been analysed with rigorous metrics from network science, with the aim of highlighting the study of real-world physical systems as an important and interesting addition to the wider field. In particular these networks display unique constraints as a result of their underlying physics. It has been shown that their mean node degree is fixed and the node degree distribution is well defined, following Lemaître's law. In addition the concept of network assortativity has been introduced to measure ring correlations, and its preferability over the previous empirical measure known as the Aboav-Weaire law has been argued. Although the assortativity has been shown to be a function of the potential model for a system and the limits of the assortativity linked to the occurrence of well-known physical motifs; most physical networks show a very similar overall level of disassortativity, as experienced in nature. An exception to this rule has also been found, where variable-coordination systems can de-mix to exhibit assortative behaviour.

In this chapter it has demonstrated how network science is applicable to understanding and analysing generic systems in physics, but also how physical systems form a key and under-explored area of network science. Going forward there is lots of potential scope to extend these explorations. For example, there are still questions to be answered from this work such as how network properties such as the assortativity are explicitly related to the physics of the underlying system and whether this information can be utilised experimentally - for example to control and effectively quantify the pore size in materials. This has also set up extensions to investigating more disordered networks still, such as biological networks which have a wider range of coordination environments.

7 | Voronoi Analysis of Quasi-Two-Dimensional Hard Spheres

Experimental investigations have led to the synthesis of colloidal monolayers, where particles are sedimented on a surface, creating effective quasi-two-dimensional hard sphere systems. Voronoi diagrams are constructed for a range of configurations of such systems, generated from both experiment and computation. The evolution in the network properties with packing fraction is explored with packing fraction for mono-, bi- and polydisperse particle systems. A detailed comparison is presented of unweighted and weighted variants of the Voronoi construction in the context of quasi-two-dimensional systems. It is shown that the two-dimensional unweighted Voronoi, favoured in experimental analyses, has a well-defined physical interpretation, corresponding to the basal section of a three-dimensional weighted Voronoi. In addition the stereology of the three-dimensional Voronoi is examined and constrasted with equivalent systems of hard disks.

7.1 Quasi-2D Hard Sphere Systems

Hard particle models are a central tenet of statistical physics, forming the basis of fundamental research dating from the earliest computations to current research [107]. This is because, despite their simplicity, the hard disk and hard sphere models are able to explain many of the behaviours of classical particles in two and three dimensions. In particular, these models effectively complement the study of colloids [176]. The interest in colloids themselves stems from the fact that they occupy a “sweet-spot”, in that the particles are large enough to observe with confocal microscopy, whilst being small enough that their motion is governed by simple fundamental forces on a reasonable time scale. It is therefore possible to track and visualise particle positions in real-time, also making them an good proxy for classical atomic systems.

The hard particle model was introduced in section 3.3.1, in the context of hard disks and hard spheres, which, as mentioned, have been extensively studied in the literature. However, section 6.4.2 introduced a new system of recent experimental interest, that of a monolayer of hard spheres. These systems, where colloidal particles are sedimented on a plane in a single layer, are three-dimensional (3D) systems with effectively two-dimensional (2D) interactions. As such, they are termed quasi-two-dimensional (quasi-2D). In section3.3.1, all monolayers comprised spheres of the same size, but one can equally engineer systems where the sphere radii have a distribution of sizes *i.e.* different size dispersity [177, 178]. In the polydisperse case, the centres of the spheres no longer lie in the same plane, but rather at a height equal to their radius. The result of this dispersity is that the interaction distances between particles of different sizes cannot be trivially projected into 2D. Nevertheless, it is still possible to model monolayers of spheres in 2D by using non-additive distances. For two particles in a quasi-2D arrangement with radii, R_i, the contact distance in 2D, R_{ij}, is related to the geometric mean of the radii:

$$R_{ij} = 2\left(R_i R_j\right)^{1/2} , \tag{7.1}$$

as illustrated for two spheres, A, B, in figure 7.1. Alternatively, this can be expressed in terms the arithmetic mean and a non-additivity parameter, Δ, as is common for asymmetric systems [179]:

$$R_{ij} = \left(R_i + R_j\right)\left(1 + \Delta\right) , \tag{7.2}$$

where

$$\Delta = \frac{2\left(R_i R_j\right)^{1/2}}{R_i + R_j} - 1 \, . \tag{7.3}$$

This allows quasi-2D systems to be modelled purely in 2D.

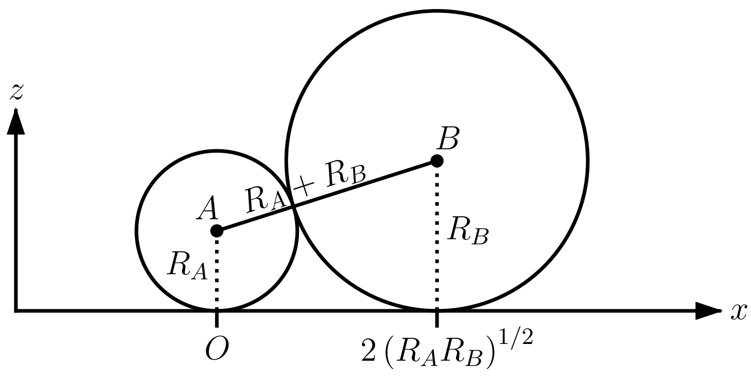


Figure 7.1: Quasi-2D hard spheres can be modelled in 2D by using non-additive interaction distances. Here two spheres sedimented on a plane have a contact distance given by twice the geometric mean of the radii.

Since systems of colloidal monolayers can be treated as a 2D problem, this should enable them to be analysed with a 2D Voronoi diagram. Voronoi analysis allows determination of the coordination environments around each particle, so that network properties such as the neighbour degree distribution and correlations can be calculated [32, 160, 180, 181]. This information can be used, for example, to give insights as to the phase behaviour in these systems [50, 182–184]. However, it is not initially clear how the non-additivity will affect the calculation of the Voronoi diagram. For instance, unlike a system of additive hard disks, it is not obvious if a unweighted or weighted variant of the Voronoi construction is most appropriate, and in the latter case what the weightings should be. It will be demonstrated in the second half of this chapter that in fact the unweighted construction retains a precise physical meaning for quasi-2D systems, and is the natural choice for partitioning space in these monolayers.

7.1.1 Experimental Analysis

Raw experimental coordinates were kindly provided for a range of mono- and bidisperse colloidal systems by Thorneywork and Dullens [50, 177, 185]. The colloidal monolayers were prepared by dispersing particles of specific radii in a water-ethanol mixture and sedimenting them on the base of a glass sample cell. The samples were then imaged using an inverted bright-field microscope and particle coordinates obtained using standard particle tracking routines. Allowing a time

Table 7.1: Summary of the experimental parameters which can be controlled in colloidal monolayers.

	Mono	Binary		
		Large	Small	Total
Radius	R	R_l	R_s	-
Radius ratio	1	-	-	$\gamma = R_l/R_s$
Composition	1	c_l	c_s	1
Number density	ρ	$\rho_l = c_l\rho$	$\rho_s = c_s\rho$	$\rho = \rho_l + \rho_s$
Packing fraction	$\phi = \rho\pi R^2$	$\phi_l = \rho_l\pi R_l^2$	$\phi_s = \rho_s\pi R_s^2$	$\phi = \phi_l + \phi_s$

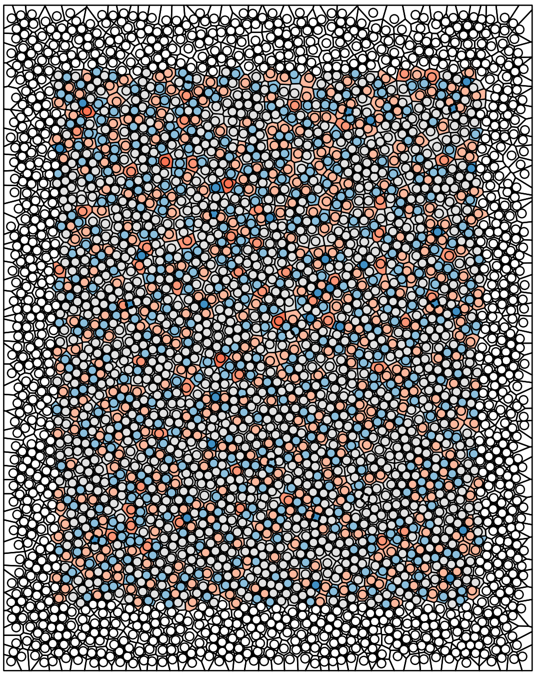
of around $10s$ between frames ensured that the particle positions are sufficiently decorrelated to be statistically independent.

Data was made available across a range of experimental conditions. For the monodisperse systems, samples contained particles with radii of $R = 1.395\mu$m, across a range of 2D packing fractions, ϕ, as defined in equation ref . For bidisperse systems, which consist of two particle sizes, one "large" and the other "small", there are more free parameters to control, which are summarised in table 7.1. Firstly there is the radius ratio of the particles, $\gamma = R_l/R_s$ (the subscripts "l" and "s" will be used to denote variables corresponding to large and small particles respectively), which relates the relative particle sizes. Secondly, the composition, c_l, c_s, determines the proportion of each particle type, where $c_l + c_s = 1$. Finally there is also the packing fraction, calculated with reference to one of the components ϕ_l, ϕ_s, or using all particles to give a total packing fraction $\phi = \phi_l + \phi_s$. For completeness, it is noted that the composition can also be represented by ϕ_l/ϕ, but for the purposes of this work the definition above was found to be more intuitive. The available experimental data contained systems at two radius ratios and a variety of compositions and packing fractions. More specifically, these had either $R_s = 1.395\mu$m and $R_l = 3.05, 2.02\mu$m, corresponding to size ratios of $\gamma = 2.19, 1.45$ respectively. In addition for each set of experimental conditions there were 100 configurations, which the exact compositions and packing fractions for each were determined and presented in table 7.2

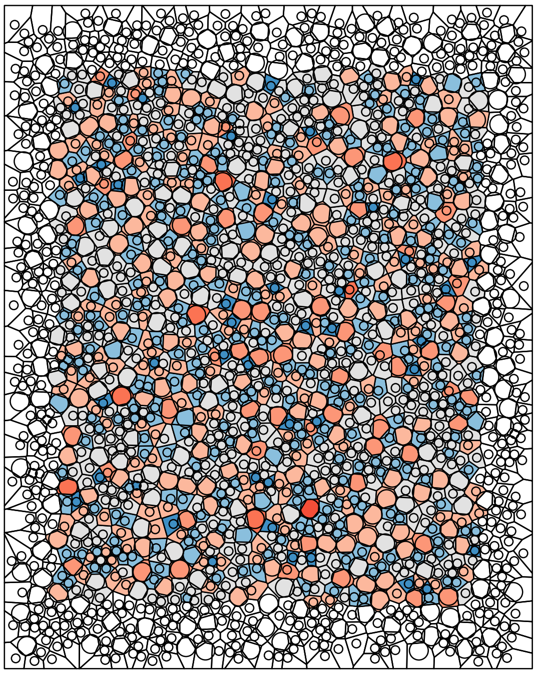
Table 7.2: Details of experimental colloidal samples. Samples are defined in terms of particle size dispersity, particle radius ratio, composition and total packing fraction. The mean value is supplied for each, with the standard deviation given in brackets.

Mono	Binary, $\gamma = 2.19$		Binary, $\gamma = 1.45$		Binary, $\gamma = 1.45$	
ϕ	c_l	ϕ	c_l	ϕ	c_l	ϕ
0.655(1)	0.201(1)	0.764(2)	0.532(4)	0.571(1)	0.233(1)	0.607(2)
0.616(1)	0.182(1)	0.639(2)	0.522(4)	0.477(4)	0.208(1)	0.406(1)
0.509(1)	0.192(2)	0.507(3)	0.543(5)	0.267(1)	0.226(3)	0.309(2)
0.427(1)	0.189(2)	0.339(3)	0.332(1)	0.629(1)	0.091(1)	0.663(1)
0.341(1)	0.187(7)	0.150(2)	0.314(3)	0.499(3)	0.107(1)	0.500(2)
0.289(0)			0.345(2)	0.337(2)	0.087(1)	0.257(1)

The 2D experimental coordinates can be analysed using a Voronoi construction. An example for both a mono- and bidisperse configurations can be found in figure 7.2. With the raw coordinates, the experimental samples can be treated analogously to configurations generated from computation. The only difference is that the experimental images are by nature aperiodic, and so when analysing the network properties of the Voronoi diagram, the cells close to the boundary must be neglected, to remove any edge effects. However, careful examination of figure 7.2a reveals some small imperfections in the data, as there are instances where the monodisperse particles appear to overlap. This is not a result of differences in particle size, but rather the presence of “phantom” particles deriving from the imaging process. Extracting particle positions from the experimental data is highly non-trivial, in particular removing points arising from interstitial sites between densely packed particles [185]. Whilst this becomes more problematic at higher packing fractions, they are however still few and far between. It should also be point out here that the tracking routines are able to detect the difference between large and small particles in bidisperse systems, allowing Voronoi cells to be associated to particles of a specific size. The actual results of the analysis of these experimental systems will be discussed alongside the results from computation in section ref .



(a) Mono, $\phi = 0.509$



(b) Binary, $\gamma = 2.19$, $c_l = 0.192$, $\phi = 0.507$

Figure 7.2: Voronoi analysis of two example experimental snapshots, of a monodisperse and bidisperse quasi-2D colloidal monolayer (system parameters in captions). Circles indicate particles of a given radius whilst Voronoi cells are coloured according to size. Voronoi cells without shading indicates those that were neglected from network analysis due to proximity to the image boundary.

7.1.2 Non-Additive Hard Disk Monte Carlo

Experimental data can be compared and contrasted with configurations generated from simulation. Hard particle Monte Carlo was introduced in section 3.3 as a method to generate such configurations computationally. One could set up a 3D system of hard spheres constrained to a plane, but as discussed above it is preferable to employ a non-additive hard disk model. In this modification, if two particles are separated by a distance, r_ij, the pair potential is:

$$\mathcal{U}_{ij} = \begin{cases} \infty & r_{ij} < 2\left(R_i R_j\right)^{1/2} \\ 0 & \text{otherwise} \end{cases} . \tag{7.4}$$

The remainder of the algorithm proceeds the same as in the standard methods, as outlined in section 3.3.

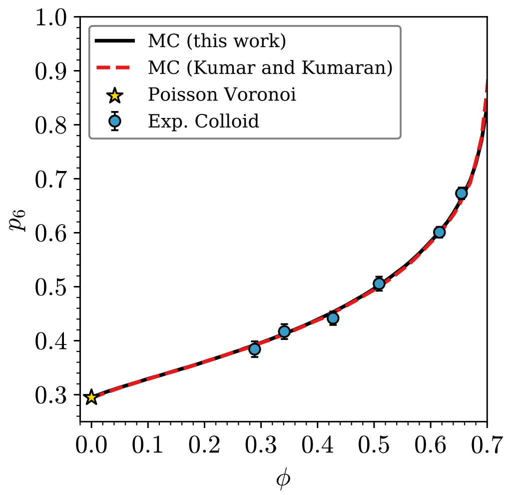
For all simulations in this chapter, unless otherwise stated, $\mathcal{N} = 1000$ particles were placed in a 2D periodic box, then equilibrated with 10^5 Monte Carlo cycles with each cycle consisting of $\mathcal{N}$ random moves. After equilibration, 10^5 Monte Carlo cycles are performed with sampling every 100 cycles. For each set of simulation parameters, averaging was carried out over 10 different random seeds.

7.2 Monodisperse Spheres

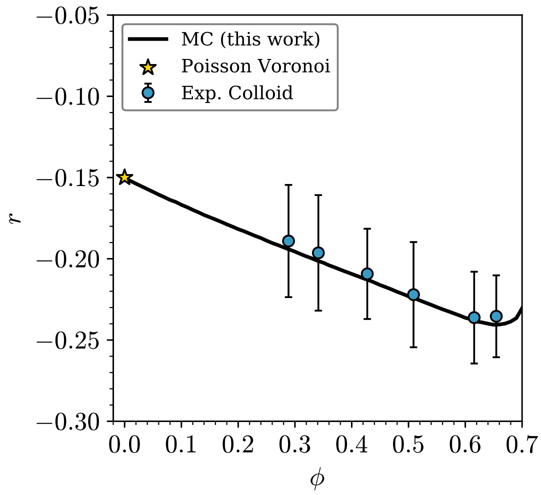
The simplest quasi-2D system to study is that of a monolayer of monodisperse spheres. This is because as all sphere radii are the same, the non-additive model reduces here to a simple additive model of hard disks. The case of monodisperse spheres has therefore already been partially explored the previous chapter. For example, sections 6.5.1 and **??** discussed the degree-distributions and assortativity in contrast with a range of other physical networks. To avoid repetition, this section will therefore focus on the network properties in terms of a measure specific to colloids, the packing fraction.

7.2.1 Network Properties with Packing Fraction

The network properties of the Voronoi diagrams in this section are again summarised by the proportion of hexagons, p_6, and the assortativity, r. As was shown in section 6.5.1, as the ring statistics in the Voronoi diagrams of monodisperse quasi-2D colloids follow Lemaître's maximum entropy law, p_6 can also be considered a descriptor for the width of the ring distribution. The assortativity once again measures the local ring correlations. Figure 7.3a shows the results of the evolution of p_6 with ϕ, for both experiment and simulation. The first thing to note is that there is excellent agreement between the experimental system and the hard particle model, indicating that this model is indeed suitable for exploring the behaviour of colloidal monolayers. In addition, comparison is also provided to a previous computational study [181], to which again there is perfect agreement. More generally, the value of p_6 decreases with packing fraction. The hexatic phase, which exists at high packing fraction is characterised by having particles in 6-coordinate environments. Under melting, as the available volume increases, 5- and 7-ring defects are initially



(a)



(b)

Figure 7.3: Network properties of monodisperse systems of quasi-2D hard spheres, in terms of the proportion of hexagons, panel (a), and the assortativity, panel (b). Data is presented both from simulation and experiment, with the experimental points indicating the mean value and one standard deviation. A comparison is also made in panel (a) to a previous study [181], to which there is excellent agreement.

introduced, depreciating p_6, before more extreme ring sizes manifest. Finally, in the limit of the ideal gas (as $\phi \to 0$), the Poisson Voronoi tessellation is obtained.

The behaviour of the assortativity with packing fraction is given in figure 7.3b. Once again, there is very good agreement between configurations from experiment and computation. The fluctuations in the values of r are larger than p_6, for the experimental systems. This is because......Local measure? Mark? Strikingly, the assortativity seems to display behaviour that is linear in packing fraction, at least for intermediate values. It is not immediately clear why this should be the case, and is an interesting result that warrants more research. In addition, it can be seen that this linearity comes to an end around $\phi \approx 0.66$, with r displaying a sudden upturn. This point is close to the limit of the liquid phase for hard disks, before it undergoes transition to a hexatic phase [50]. This phase transition seems to be captured clearly through the assortativity.

Check all this with Mark

7.3 Bidisperse Spheres

Monolayers of bidisperse spheres present a natural extension to the monodisperse case, and come with the advantage of having available experiment information for comparison. In this section the ring statistics of bidisperse systems will be explored, with a focus on systems matching the experimental parameters, as outlined in table 7.2. Owing to the fact that there are two types of particle present (which experimental imaging is able to differentiate), metrics in this section will be considered in reference to both large and small particles. To this point, the ring statistics can then be divided into partial ring distributions for each particle type, denoted p_k^l and p_k^s. A subtle but important point is that the mean ring sizes for these partial distributions ($\langle k \rangle_l$ and $\langle k \rangle_s$) are no longer constrained to be six. Instead it is their weighted sum which is constrained by Euler's formula such that:

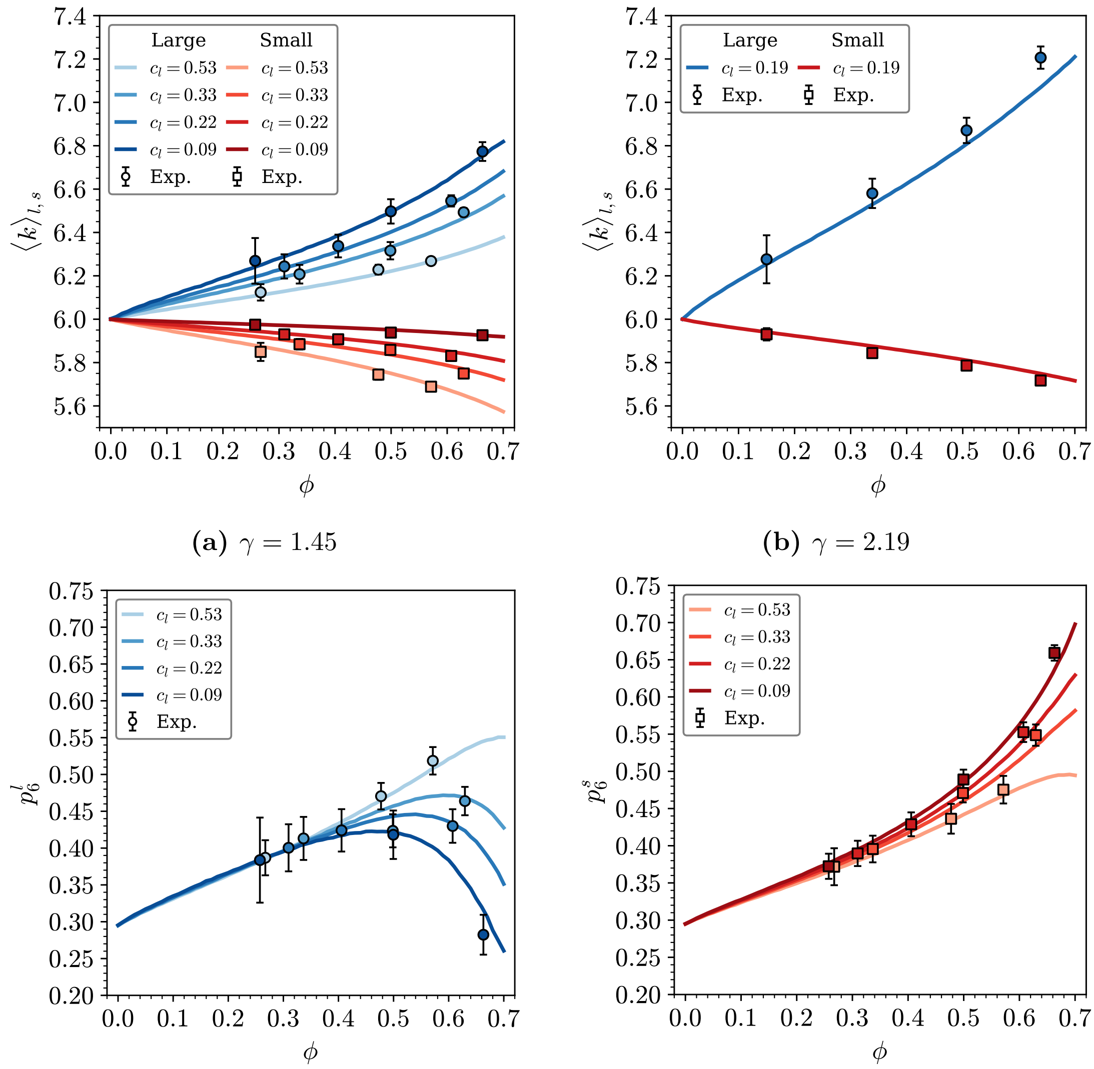
$$c_l \langle k \rangle_l + c_s \langle k \rangle_s = \langle k \rangle = 6\,. \tag{7.5}$$

The mean ring sizes can then take values either side of six, quantifying the relative coordination numbers of the particles of different sizes.

In this section systems of two radius ratios will be discussed: the first with $\gamma = 1.45$, termed the small size ratio; the second with $\gamma = 2.09$, termed the large size ratio. The primary focus will however be on the small size ratio, as this is complemented by the most experimental data. The overarching properties of the ring structure will first be explored, before the explicit ring statistics are again examined in relation to a maximum entropy solution.

7.3.1 Ring Statistics with Packing Fraction

The mean values of the partial ring distributions for large and small spheres are given for the bidisperse systems of both radius ratios in figures 7.4a, 7.4b. Beginning with the small size ratio ($\gamma = 1.45$, figure 7.4a), it can be seen that there is once again good agreement with the experimental data. With the aid of the Monte Carlo simulations, it is clear that these experimental points actually lie on a series of curves. To rationalise the broad behaviour of these curves, one can begin by noting



(a) $\gamma = 1.45$

(b) $\gamma = 2.19$

(c) $\gamma = 1.45$, large spheres

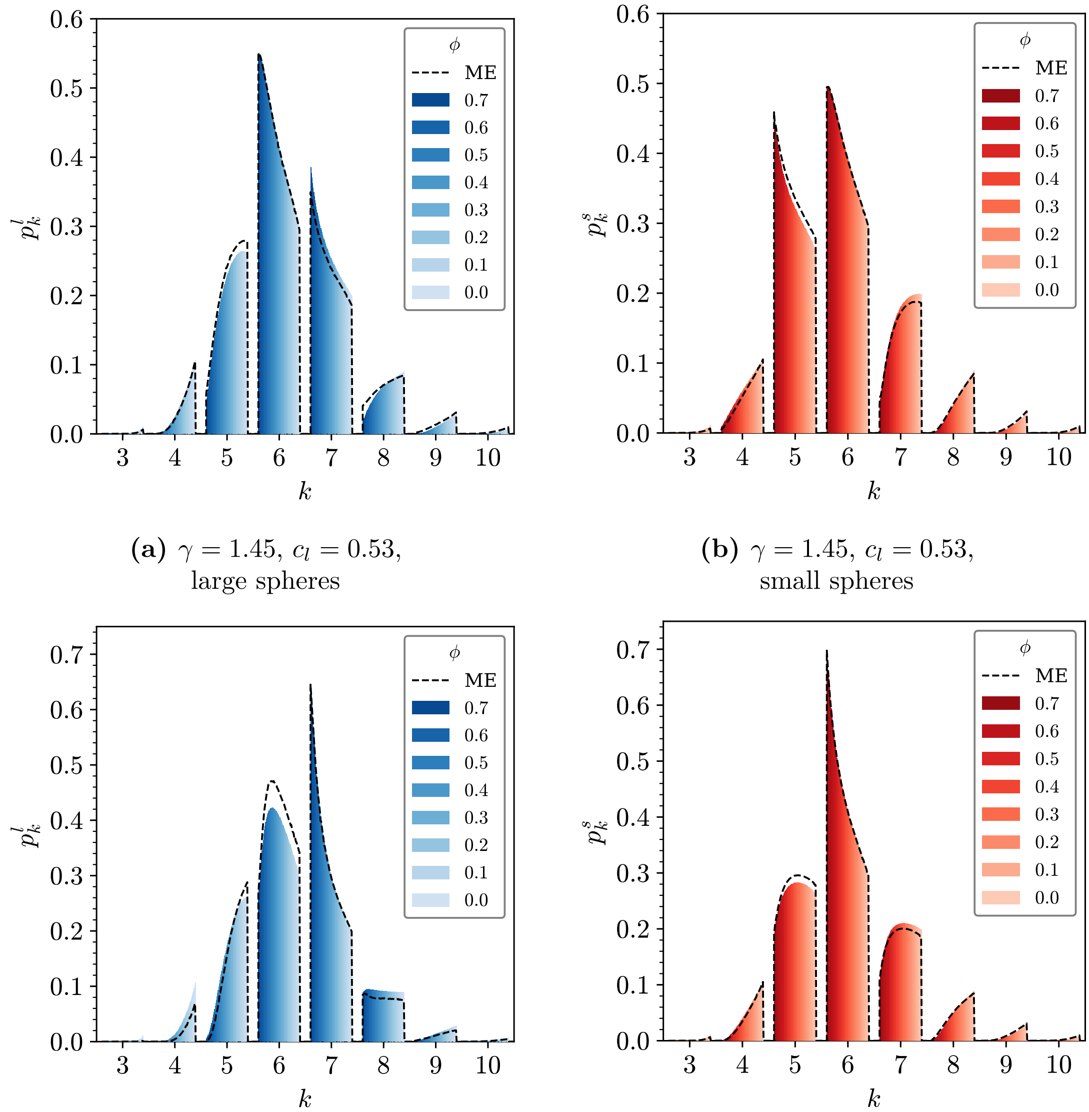
(d) $\gamma = 1.45$, small spheres

Figure 7.4: Overview of the evolution in the partial ring statistics in bidisperse colloidal monolayers. Panels (a),(b) show the partial mean sizes for the two radius ratios (as indicated in captions). Panels (c),(d) show the proportion of hexagons in the partial distributions for the $\gamma = 1.45$ systems. The blue curves correspond to quantities pertaining to large particles, red curves to small particles. The system composition is indicated by the depth of shading, as in the legends. Experimental points are coloured using the same scale as for simulation curves, and error bars indicate one standard deviation of the mean.

that $\langle k \rangle_l > 6$ and $\langle k \rangle_s < 6$ for all values of $\phi > 0$. This is simply an expression of the fact that larger spheres will always have on average more nearest neighbours than smaller spheres. Following this observation, it can be seen that $\langle k \rangle_l$ increases as the fraction of large particles decreases, for a fixed packing fraction. This is a due to large spheres having statistically fewer large sphere neighbours the lower the concentration c_l. Each large particle can accommodate more particles around it, the greater the dilution of small particles. As a further illustration, one can consider a large sphere in a "sea" of small spheres ($c_l \to 0$). Here the large particle nearest neighbours must be maximised and equally the small particles only interact with other small particles. As such the mean coordination for the small particles must approach the average of six. This reciprocal behaviour of small spheres also explains the trend in $\langle k \rangle_s$, which approaches the horizontal limit of $\langle k \rangle_s = 6$ as $c_l \to 0$. Once again the effect of decreasing packing fraction is to diminish the effects of relative sphere size, as the excluded volume is lower, until eventually the random limit is reached. Come up with a satisfactory relative model for this? Perhaps I have a semi-decent one somewhere...Nope, don't think so.

The analogous data is also presented for the large size ratio system in figure 7.4b. Here it can be seen that similar trends hold for $\langle k \rangle_l$ and $\langle k \rangle_s$ as discussed for the small size ratio above. However, the increased size differential leads to even more extreme differences in the partial mean ring sizes. For example, for the large size ratio with $c_l = 0.19$, $\langle k \rangle_l \approx 7.21$ but for the small size ratio with $c_l = 0.22$, $\langle k \rangle_l \approx 6.68$. It should be noted that the fit with experiment is less accurate for the large size ratio spheres, a problem that is particularly pronounced at higher packign fractions. This suggests two possible inadequacies. Either the hard particle model is insufficient, and there are some additional short range interactions which become non-negligible at higher packing fractions, or the difficulties in directly imaging the particles become more significant.

The proportion of hexagons for each subsystem is also plotted in terms of the packing fraction in figures 7.4c, 7.4d, for the spheres with small size ratio. Again, even at this more fine-grained level, the correlation with experimental data remains strong. The results for the small spheres in figure 7.4d reflect a general trend regularly seen in this thesis, where the number of hexagons is maximised as ordering increases with higher packing fraction, and again the p_6 increases as the number of disrupting large spheres decreases. The results for the large spheres in figure 7.4c are somewhat more dramatic, with p_6 displaying a maximum with changing packing.



(a) $\gamma = 1.45$, $c_l = 0.53$, large spheres

(b) $\gamma = 1.45$, $c_l = 0.53$, small spheres

(c) $\gamma = 1.45$, $c_l = 0.09$, large spheres

(d) $\gamma = 1.45$, $c_l = 0.09$, small spheres

Figure 7.5: Variation of partial ring distributions with packing fraction for two different compositions of a $\gamma = 1.45$ bidisperse monolayer, contrasted with maximum entropy solutions. Left panels show the statistics for rings associated with larger particles, and right panels with smaller particles. The maximum entropy solutions are overlaid as dashed lines for comparison.

This will be discussed to greater extent in the following section, but essentially results from the hexagon no longer being the dominant ring size, begin replaced by the heptagon. Similar trends were in fact seen for x-rings in triangle rafts in section ref .

7.3.2 Maximum Entropy Solutions

Is figure 7.5 clear or confusing?

To further investigate the ring statistics in bidisperse colloidal monolayers, the numerical partial ring size distributions are compared to maximum entropy solutions. These maximum entropy solutions to are assumed to follow the same form as Lemaître's law, *i.e.* satisfy the constraints:

$$\sum_k p_k = 1\,, \tag{7.6}$$

$$\sum_k k p_k = \langle k \rangle_{l,s}\,, \tag{7.7}$$

$$\sum_k \frac{p_k}{k} = \text{constant}\,. \tag{7.8}$$

Unlike Lemaître's law, which requires only the value of a single ring, such as p_6, to obtain the entire distribution, the additional freedom afforded to the mean ring size means that here the mean ring size of the partial distributions must also be provided.

The comparison of the partial ring statistics from simulation and maximum entropy is given in figure 7.5. The partial distributions are provided for the large and small spheres, with a small size ratio, at the two limiting compositions of $c_l = 0.53, 0.09$. The continuous evolution in the ring statistics with packing fraction is detailed by the shading of the bars (note for instance that the $k = 6$ bars in figure 7.5 are the reverse of the lines of $p_6^{l,s}$ in figures 7.4c, 7.4d). In these plots, the left-most points in each bar represent the system at the highest packing fraction, and the right-most points the limit of zero packing fraction. In all cases, the maximum entropy solutions fit the observed numerical results quite well. The difference in the ring statistics for each component is also stark. For the $c_l = 0.53$ system, which is roughly equal in large and small spheres, although at high packing fraction p_6 is comparable in magnitude, the distribution for large spheres is skewed heavily towards large rings (particularly $k = 7$), whereas that for the small spheres is skewed towards small rings (particularly $k = 5$). For the $c_l = 0.09$ system, the distribution for the large spheres is even more dramatic, being dominated at high packing fraction by 7-rings. This causes the value of p_6 to exhibit a maximum in ϕ, as the proportion initially rises as 7-rings decay, before heading towards the PV limit. One might think that this necessarily must also enforce the small sphere distribution to be more "extreme", but in fact the reverse is true. As the small spheres interact primarily with other small spheres, due to their high relative abundance, the ring distribution actually appears more similar to that of the monodisperse case. Hence, in this way, the two effects offset each other, so that by extremising the distribution of one component, the other must become more regular.

Again, check with Mark

7.4 Voronoi Construction in Quasi-2D

The case of quasi-2D polydisperse spheres is a further generalisation of the mono- and bidisperse systems discussed already in this chapter. Owing to this generality, a different approach will be taken here, with the meaning of the Voronoi diagram carefully examined in the context of quasi-2D hard sphere systems, before numerical results are presented for polydisperse spheres. The motivation for this is that up to this point, it has merely been stated that the unweighted 2D Voronoi is an appropriate construction (as opposed to a weighted variant - see section 3.3.3), even for collections of particles of varying radii. The opportunity will now be taken to explore this claim more fully.

Various experimental approaches exist to create quasi-2D systems, including confining particles between two planes [168, 186], adsorbed at an interface [187, 188] as well as sedimentation on a surface [178, 189]. In fact, it is the nature of the system which will determine which Voronoi analysis is most appropriate. Taking first the simplest quasi-2D case, of the monolayer of monodisperse spheres, the particle centres will occupy the same plane. As demonstrated in figure 7.6a, the 3D Voronoi polyhedra for such system are prismatic, such that any horizontal cut through the construction will yield the same 2D tessellation. In addition, it is relatively straightforward to see that this 2D tessellation is topologically equivalent to the 2D Voronoi diagram, calculated in the plane of the particle centres. This is to say

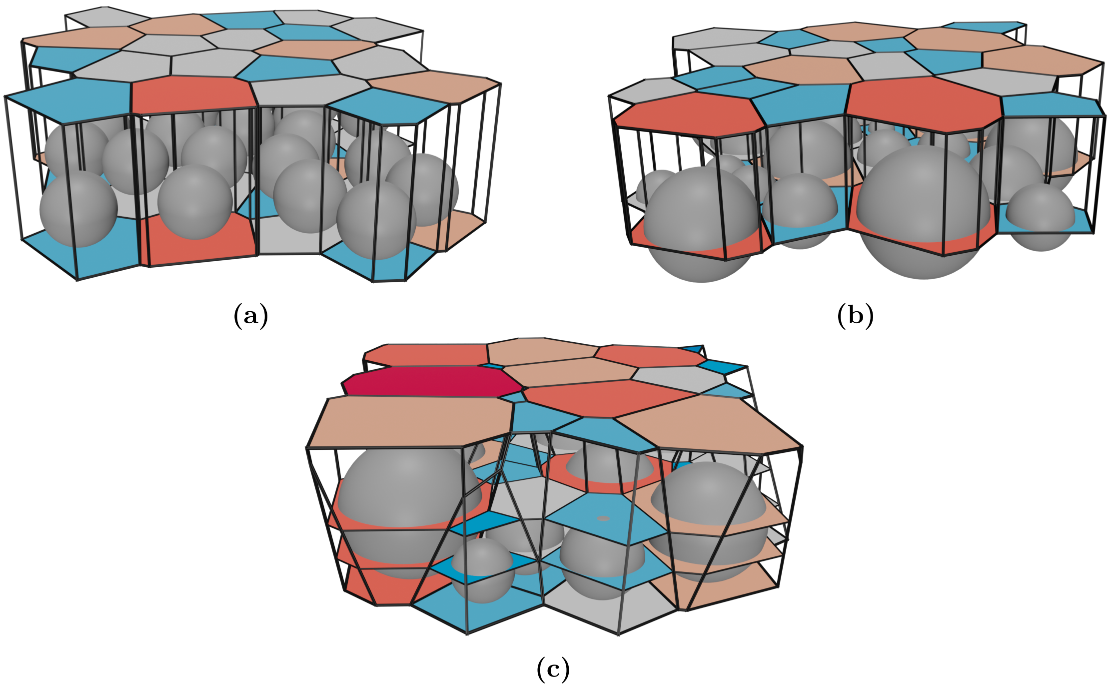


Figure 7.6: Voronoi diagrams in three different quasi-2D systems. The 3D Voronoi polyhedra are demarked by the thick black lines, whilst selected 2D Voronoi polygons are shaded. Panel (a) shows a monodisperse system with spheres sharing a basal plane and particle centres also occupying the same horizontal plane. Panel (b) shows a polydisperse system where particle centres lie in the same horizontal plane. A 3D Voronoi diagram weighted by the sphere radii is overlaid. Panel (c) shows a polydisperse system where particles share a basal plane. A 3D Voronoi diagram weighted by the sphere radii is overlaid.

that the 3D Voronoi can be trivially constructed from the 2D Voronoi, and so such systems can be analysed with an unweighted 2D Voronoi, with no loss of information.

A simple extension is then to take polydisperse particles in which the particle centres occupy the same plane (*e.g.* particles suspended at an interface). As previously alluded to, a weighted Voronoi construction is required in this case, to avoid the possibility of the polyhedra cutting through the larger particles. Calculating the 3D Voronoi diagram, weighted by the sphere radii, once again leads to prismatic polyhedra, as shown in figure 7.6b. Now the 2D tessellations formed from horizontal cuts are topologically equivalent to the 2D Voronoi diagram, weighted by the sphere radii, calculated in the plane of the particle centres. Once again the partitioning of space can be fully described in two dimensions, and so such systems can be analysed with a weighted 2D Voronoi.

The more complex case is to consider a monolayer of polydisperse particles sedimented on a surface. Under these conditions the spheres share a basal tangent plane, and the centres no longer occupy a common plane. Although there is a trivial projection of the particle centres into 2D, the interaction distances between particles are non-additive, reflecting the essential 3D nature of the problem. As can be seen in figure 7.6c, the polyhedra in the 3D Voronoi diagram, weighted by sphere radii, are no longer prismatic. The 2D tessellations formed from horizontal cuts through the system therefore have a non-trivial relationship with cut height. Rather, these 2D tessellations show an evolution in structure as the height above the basal plane increases - both in terms of the number of polygons and their properties. Under these quasi-2D conditions, it is not necessarily clear what is the best approach for conducting a Voronoi analysis, or even how to define even basic properties such as the packing fraction.

It is these final quasi-2D systems which are the focus of this chapter. In this section, it will be demonstrated that for these systems, the basal tessellation in the weighted 3D Voronoi is topologically equivalent to an unweighted 2D Voronoi diagram. This result will be extended to explore the more general stereology problem, showing that tessellation at an arbitrary cut height above the basal 2D plane can be related to a 2D weighted Voronoi diagram. These tessellations will also be compared to those formed from equivalent arrangements of hard disks.

Revisit Voronoi section in methods. Introduce weighted Voronoi there or here?

7.4.1 Polydisperse Hard Sphere Model

As a generalisation, this section now considers a system of polydisperse hard spheres sedimented onto a surface, such that all spheres share a common basal tangent plane. For numerical simulations, it will further be constrained that the radii, R_i, of the particles follow a lognormal distribution:

$$f(R) = \frac{1}{R\sqrt{2\pi\sigma^2}} \exp\left[-\left(\frac{\ln R - \mu}{\sqrt{2\sigma^2}}\right)^2\right], \tag{7.9}$$

where as usual μ and σ are respectively the mean and standard deviation of the logarithm of the radii. This distribution is chosen to ensure the radii of randomly generated particles are always positive. Whilst the choice of distribution does not affect the fundamental conclusions of the Voronoi analysis below, it will help quantify properties of the system such as the packing fraction.

7.4.2 Stereological Relationships

The unweighted and weighted Voronoi constructions were introduced in section 3.3.3. To review, the unweighed Voronoi requires only the positions of the particle centres, placing dividing planes halfway between neighbouring points. The weighted radical variant requires knowledge of both positions and radii of the particles, repositioning the dividing planes further from the larger particles, hence allocating more them more space. When studying quasi-2D colloids, it is possible to construct a 3D weighted Voronoi using the coordinates of the particle centres, $\mathbf{r}_i = (x_i, y_i, R_i)$. However, in most cases researchers prefer to analyse configurations using a 2D Voronoi using the projected particle centres $\mathbf{r}'_i = (x_i, y_i)$. The rational behind this is clear: 2D analysis is easier to visualise and rationalise. In addition, as in figure 7.6a, for monodisperse systems the 3D representation contains no more information on neighbouring interactions and assigned volumes than the 2D analogue. However, as in figure 7.6c, for polydisperse particles whilst this may be a good first approximation, this is not necessarily accurate, and so it is important to examine the relationship between Voronoi diagrams in two and three dimensions to ensure correct physical meaning is attributed to the Voronoi analysis.

7.4.2.1 Limiting Case with Basal Plane

To begin with, the stereological relationship between the 3D weighted Voronoi and the 2D tessellation formed from the intersection of the radical planes with the basal tangent plane will be considered. In figure 7.6c, this corresponds to the bottommost 2D tessellation. To do this, take the arrangement in figure 7.7a, with two spheres A,B separated by the dividing radical plane, V, and sharing the tangent plane, T, with equation $z = 0$. The distance between the sphere centres is r_{AB}, whilst the distance between the projected centres is denoted r'_{AB}. The dividing plane V, has a normal given by $\mathbf{n} = \overrightarrow{AB} = r'_{AB}\hat{\mathbf{\imath}} + (R_B - R_A)\,\hat{\mathbf{k}}$. In addition the point D which lies on V is located at $\overrightarrow{OD} = r_A\hat{\mathbf{n}} + R_A\hat{\mathbf{k}}$ where $r_A = \frac{R_A^2 - R_B^2 + r_{AB}^2}{2r_{AB}}$ and $|\mathbf{n}| = r_{AB}$. Defining the direction of x as as the projection of $d_A\hat{\mathbf{n}}$ on the basal plane, the plane equation for V can be deduced as:

$$r'_{AB}x + (R_B - R_A)\,z = \frac{(r'_{AB})^2}{2}\,. \tag{7.10}$$

The intersection of V with T occurs at $z = 0$, and so the line of intersection is given simply by $x = r'_{AB}/2$. However there is something significant about this result: this dividing line in 2D is the same as that which would be obtained from constructing the 2D unweighted Voronoi using the projected particle centres $\mathbf{r}'_i$. As this must be true for all pairs of neighbouring particles, it follows that the entire 2D unweighted Voronoi will be constructed. This leads to a key relationship: the unweighted 2D Voronoi diagram constructed from the projected particle centres is topologically equivalent to the tessellation formed by taking the basal faces of the polyhedra in the weighted 3D Voronoi diagram.

7.4.2.2 General Case with Arbitrary Horizontal Plane

Taking the basal faces of the 3D Voronoi polyhedra can equally be thought of as taking a horizontal cut through the tessellation at $z = 0$. The result above is essentially a special case of the fact that a cut through a 3D Voronoi must yield a tessellation which is equivalent to some weighted 2D Voronoi [190, 191]. Following from the result above, one may ask what is the analogous relationship between a 2D and 3D Voronoi when taking a cut at arbitrary z, as in figure 7.6c. Revisiting the simple arrangement in figure 7.7a, for a given horizontal cut height, z, the projected particle centres are now defined as $\mathbf{r}'_i = (x_i, y_i, z)$. Regardless of the the value of z, we note that the distances between the projected centres remain the same at r'_{AB}. To obtain the same distance between the projected particle centres and the dividing plane using both a 2D and 3D Voronoi, it can be seen from combining equations (3.27) and (7.10) that:

$$w_A^2 - 2R_Az = w_B^2 - 2R_Bz\,. \tag{7.11}$$

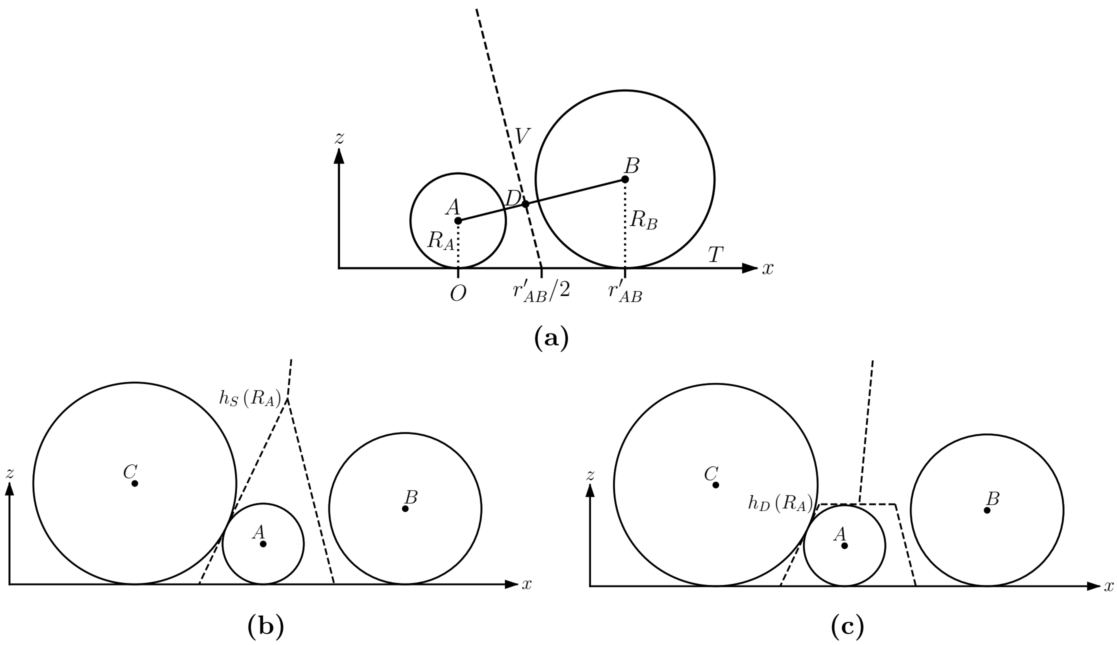


Figure 7.7: Panel (a) shows two spheres and the dividing radical plane between them. The radical plane intersects the tangent plane at half the horizontal distance between the particles. This intersection generates the same dividing line in the tangent plane as would the unweighted Voronoi using the projected particle positions. Panel (b) shows three spheres and the 3D weighted Voronoi polyhedron formed around sphere A. The height of the cell is given by the topmost point, denoted $h_S(R_A)$. Panel (c) shows the same three spheres and the polyhedron that would be formed from stacking 2D Voronoi tessellations with disk radii as weights. As the weight for A is not defined above the sphere diameter, the polyhedron is truncated in comparison with in panel (b). The height of the cell is then given by $h_D(R_A) = 2R_A$.

The simplest solution for the weights that satisfy this equation is therefore:

$$w_{S,i} = (2R_i z)^{1/2} \ . \tag{7.12}$$

Again as this will naturally extend to a collection of many particles, more general relationship is obtained: the tessellation formed from a horizontal cut through a 3D weighted Voronoi diagram is topologically equivalent to the 2D Voronoi diagram calculated from the projected particle centres, weighted according to equation (7.12). Indeed, the that the first result is simply a limiting case of this second result, where $z = 0$ and the weights are trivially zero. This 2D weighted Voronoi diagram shall be referred to as the sphere-weighted Voronoi and variables associated with it denoted with a subscript S.

7.4.2.3 Connection to System of Hard Disks

A horizontal cut through a 3D system of hard spheres will produce a 2D system of hard disks. The radii of these disks will be related to the radii of the spheres, and if they are used as weights in a 2D Voronoi analysis they will also satisfy equation (7.11):

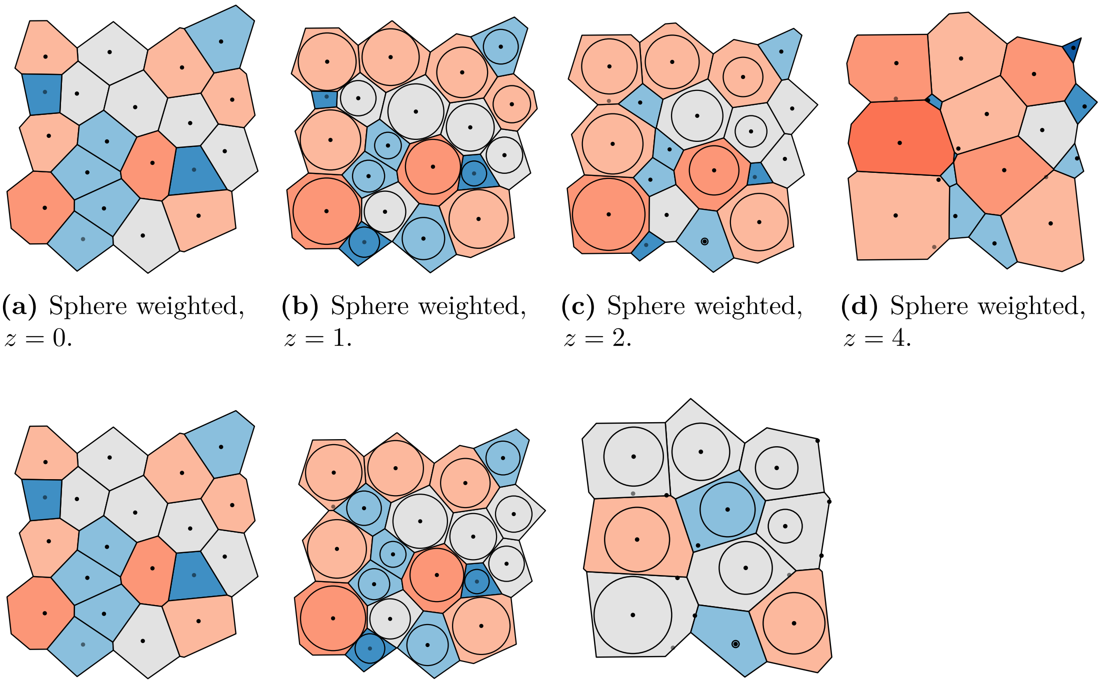
$$w_{D,i} = \left(2R_i z - z^2\right)^{1/2} . \tag{7.13}$$

This suggests that the tessellation formed from the horizontal cut through the 3D Voronoi should also be equivalent to the 2D Voronoi using these disk radii.

However, there is a caveat to this result. It is clear that equation (7.13) is only well defined for $z \leq 2R_i$ *i.e.* the cut height is less than the sphere diameter. There is no such constraint for the polyhedral cells in the 3D weighted Voronoi, which may readily extend above the associated particle. Therefore, the result above is only strictly true when the cut height is less than the smallest sphere diameter. A modification is made where points which have diameters less than the cut height are excluded from the 2D calculation. These excluded particles will therefore not have cells in the Voronoi tessellation. In this case some 3D cells will become truncated in the 2D representation, as shown in figures 7.7b and 7.7c. This will lead to an increasing difference between the two types of partitions as the cut height is increased, as demonstrated in figure 7.8. This 2D weighted diagram shall be referred to as the disk-weighted Voronoi and variables associated with it denoted with a subscript D.

7.4.3 Properties of Voronoi Tessellations

The important properties of the sphere and disk-weighted Voronoi diagrams are outlined in this section, which will aid with their analyses. These include the expected number of points at a given cut height, the weight distribution of the included points and the nearest-neighbour distances between them. These are the quantities that appear in equation (3.27) and therefore govern the wider properties of the Voronoi diagram. The analytic results will be demonstrated for the model of lognormally distributed sphere radii will, but the results extend to any configuration where the sphere radii are known.



(a) Sphere weighted, $z = 0$. **(b)** Sphere weighted, $z = 1$. **(c)** Sphere weighted, $z = 2$. **(d)** Sphere weighted, $z = 4$.

(e) Disk weighted, $z = 0$. **(f)** Disk weighted, $z = 1$. **(g)** Disk weighted, $z = 2$.

Figure 7.8: Comparison of the tessellations formed from a horizontal cut through a 3D Voronoi diagram weighted by sphere radii (a)-(d) and a 2D Voronoi diagram weighted by disk radii (e)-(g), at increasing z values. The same configuration is used as in figure 7.6c, with the horizontal cuts corresponding to the left hand diagrams. Black circles indicate the particle radii at the cut height, which also correspond to the weights in the disk-weighted Voronoi. Black points indicate the projected particle centres, whilst grey points indicate particles without corresponding cells in the 2D tessellation.

7.4.3.1 Average Nearest-Neighbour Distances

A useful metric is the expected distance between points in a Voronoi diagram at different cut heights, as this can be used to rationalise much of the observed behaviour. In order to do this, it is necessary to assume that particles are uniformly distributed (the ideal gas), neglecting the short range liquid structure. Whilst this might seem an over-simplification, not only does the approximation get better at lower packing fractions but also as the cut height is increased, particles are lost from the tessellation and their positions become effectively decorrelated.

Firstly, the average distance between all points in neighbouring Voronoi cells at a given cut height, $\langle r' \rangle_z$, can be found as follows. The number density at a given cut height, denoted ρ_z, is introduced as the number of points at height z, divided by the total area of the 2D Voronoi diagram. It can therefore be expressed:

$$\rho_z = \rho_0 N(z) \ , \tag{7.14}$$

where ρ_0 is the number density considering all points (as must be the case at $z = 0$), and $N(z)$ is the proportion of particles included at a given cut height. By dividing the area equally between all particles, we expect the average distance between neighbouring points in the 2D Voronoi diagram to follow:

$$\langle r' \rangle_z = 2 \left(\frac{1}{\pi \rho_z} \right)^{1/2} . \tag{7.15}$$

Alternatively, one can opt to find the average distance to the n^{th} nearest neighbour, $\langle r'_n \rangle_z$. If again a uniform distribution of spheres is assumed (*i.e.* the dilute limit), Bhattacharyya and Chakrabarti [192], showed that in two dimensions, this is given by the equation:

$$\langle r'_n \rangle_z = \left(\frac{1}{\pi \rho_z} \right)^{1/2} \frac{\Gamma(n + 1/2)}{\Gamma(n)} . \tag{7.16}$$

7.4.3.2 Properties of Disk-Weighted Voronoi

The properties of the disk-weighted Voronoi are somewhat easier to see and so it is logical to begin with these. The height below which a Voronoi cell is defined for a particle of a certain radius is given exactly by $h_D(R) = 2R$, as in figure 7.7c. Hence, for a given cut height, the particles with associated cells in the tessellation will be all those with radii in excess of the inverse function $h_D^{-1}(z) = z/2$. The proportion of particles with associated cells above a given cut height, denoted $N_D(z)$, is therefore related to the cumulative distribution function of the sphere radii:

$$N_D(z) = \int_{z/2}^{\infty} f(R)\, \mathrm{d}R = \frac{1}{2} \mathrm{erfc} \left[\frac{\ln(z/2) - \mu}{\sqrt{2\sigma^2}} \right] , \tag{7.17}$$

as all particles with $R < z/2$ are neglected.

The moments of the weight distribution function at a given cut height can then be calculated by integrating over the range of the remaining particle radii,

$$\langle w_D^n \rangle_z = \frac{1}{N_D(z)} \int_{z/2}^{\infty} \left(2Rz - z^2 \right)^{n/2} f(R)\, \mathrm{d}R . \tag{7.18}$$

as required. For example, the analytic solution for the second moment can be found in section 7.5.

7.4.3.3 Properties of Sphere-Weighted Voronoi

For the sphere-weighted Voronoi, the maximum height of the Voronoi cell for a particle of a certain radius, $h_S(R)$, is not precisely defined with R. Instead an approximate form will be derived below. Regardless of the functional form of $h_S(R)$, in a similar manner to above an equation can be written for the proportion of particles with associated cells above a given cut height, denoted $N_S(z)$, as:

$$N_S(z) = \int_{h_S^{-1}(z)}^{\infty} f(R)\,\mathrm{d}R = \frac{1}{2}\mathrm{erfc}\left[\frac{\ln\left(h_S^{-1}(z)\right) - \mu}{\sqrt{2\sigma^2}}\right], \tag{7.19}$$

with the moments of the weights at a given cut height provided by:

$$\langle w_S^n \rangle_z = \frac{1}{N_S(z)} \int_{h_S^{-1}}^{\infty} (2Rz)^{n/2} f(R)\,\mathrm{d}R, \tag{7.20}$$

in analogy with the disk-weighted case.

The functional form of $h_S(R)$ can now be considered. Rather than an exact relationship, an expression for the expected Voronoi cell height for a given particle radius is proposed. The height of a Voronoi cell is a complex function that must depend on both the radius of the associated particle and the distances to neighbouring particles. In the following arguments it will be useful to refer to figures 7.6, 7.7 and 7.8. Consider a reference particle of a given radius, R'. Crucially, the dividing planes forming the Voronoi polyhedron around it will only converge to a point when the neighbouring particles are larger than the reference particle. For the purposes of this analysis, one can therefore think of the particle operating in a reduced density system containing only the particles with radii greater of equal to the reference particle. The density is therefore,

$$\rho' = \rho_0 \int_{R'}^{\infty} f(R)\,\mathrm{d}R, \tag{7.21}$$

and mean average particle radius,

$$\langle R' \rangle = \frac{\int_{R'}^{\infty} R f(R)\,\mathrm{d}R}{\int_{R'}^{\infty} f(R)\,\mathrm{d}R}. \tag{7.22}$$

In addition, the average n^{th} nearest neighbour distance will be given by equation (7.16), substituting the density from equation (7.21) above.

Therefore the average environment around the reference particle can be considered to consist of particles with radii $\langle R' \rangle$ at located at successive distances $\langle d'_n \rangle$. These particles will generate dividing Voronoi planes according to equation (7.10). If it is then assumed that, on average, the highest point in the Voronoi polyhedron will occur above the particle centre (*i.e.* $x = 0$, see for example the cells in figure 7.6), the expected maximum cell height can be approximated by:

$$h_{S,n}(R') \sim \frac{\langle r'_n \rangle^2}{2\left(\langle R' \rangle - R'\right)} \,. \tag{7.23}$$

These can be averaged over m nearest neighbours to give the result:

$$h_S(R') = \frac{1}{2m\left(\langle R' \rangle - R'\right)} \sum_{n=1}^{m} \langle r'_n \rangle^2 \,. \tag{7.24}$$

It will be shown in section 7.6.1 and figure 7.10, that averaging over $m = 3$ nearest neighbour distances gave a good fit to the observed numerical distribution. This can be rationalised on the basis that the top vertex is formed from the intersection of three neighbouring planes, which originate from the closest three particles.

The important thing about this functional form is that the expected height of the Voronoi cell scales quickly with particle radius, unlike in the disk-weighted case. This is a result of the density of larger particles decreasing quickly with particle size, such that the average interaction distances become increasingly long and the radical planes approach vertical.

7.5 Quasi-2D Packing Fraction

One final complication with quasi-2D hard spheres is that the definition of packing fraction is ambiguous. For a quasi-2D system, if the 2D packing fraction,

$$\phi_{2D} = \pi \rho \langle R^2 \rangle \,. \tag{7.25}$$

is used, there can be adverse consequences, as this definition can lead to a packing fraction greater than unity. As an example, consider the binary crystalline lattice in figure 7.9, consisting of a two particle types with radius ratio γ. At $\gamma = 0$ the lattice is the hexagonal close packed structure which has the well known maximal monodisperse packing limit of $\frac{\pi}{2\sqrt{3}} \approx 0.907$. As γ increases, the smaller particles swell in the tetrahedral holes without affecting the large particle positions, and so the naïve 2D packing fraction rises, peaking at $\frac{11\pi}{18\sqrt{3}} \approx 1.108$ when $\gamma = \frac{1}{3}$ (the contact distance). Beyond this radius ratio the small particles can no longer be accommodated in the tetrahedral holes, and so the larger particles are forced apart which leads to a an initial decrease in packing fraction before the hexagonal close packed limit is approached once more. This is problematic, as over this range the packing fraction exceeds unity and therefore obscures the physical meaning.

Moving to 3D does not necessarily solve the issue, as it is now not clear what volume (and therefore number density) is most appropriate. However, considering quasi-2D packings as a series of stacked sections through the 3D system allows an alternative definition of packing fraction:

$$\phi\left(z\right) = \pi \rho_z \langle w_D^2 \rangle_z, \tag{7.26}$$

as a function of z, the horizontal cut height, where ρ_z and $\langle w_D^2 \rangle_z$ are provided by equations (7.14) and (7.18) respectively. The rationale for this equation is as follows. As the value, $\langle w_D^2 \rangle_z$, measures the average square radius of the circular sections in a cut of given height; here $\phi\left(z\right)$ quantifies the proportion of the total area occupied by the particle sections in the cut plane.

The overall packing fraction could thereby by quantified in these systems by $\max\left[\phi\left(z\right)\right]$. Referring again to figure 7.9, it can be seen that this definition gives values for the packing fraction which are now consistent with intuition. For the region where $\gamma \leq \frac{1}{3}$ and the large particle lattice is unperturbed by the small particles, the maximal monodisperse packing limit is found. Again beyond this the packing fraction then drops as the smaller particles swell above the contact distance before the hexagonal close packed lattice is once again recovered. In addition, if this definition were applied to the simpler systems in figures 7.6a, 7.6b, it would match with packing fraction calculated using the disks of the same radii as the spheres.

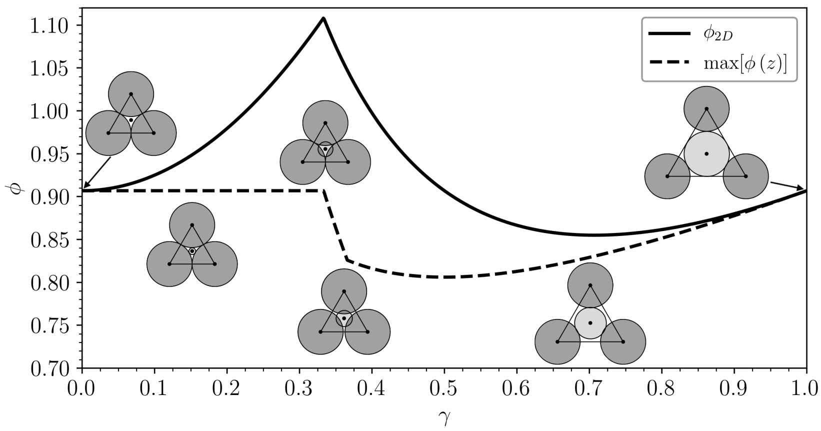


Figure 7.9: Variation in packing fraction for a binary crystal with varying radius ratio, γ. The packing fraction using the 2D definition is contrasted against that using the maximum packing with cut height.

For a polydisperse system, the explicit form of the second moment for the lognormal distribution can be calculated as,

$$\begin{aligned}\langle w_D^2\rangle_z &= \frac{1}{N_D(z)}\int_{z/2}^{\infty}\left(2Rz - z^2\right) f(R)\,\mathrm{d}R \\ &= 2\langle R\rangle z\frac{N_D\left(ze^{-\sigma^2}\right)}{N_D(z)} - z^2\,. \end{aligned} \tag{7.27}$$

There is a pleasing similarity between this result and equation (7.13), and there there will be parity when $\sigma = 0$ *i.e.* the monodisperse limit. Intuitively, the form of the packing fraction with z and its maximal value is independent of the number density and is instead solely a property of the sphere radii distribution.

7.6 Polydisperse Spheres

The previous section explored theoretically properties of Voronoi diagrams in quasi-2D hard sphere systems. The conclusions of that section are now tested numerically with monolayers of polydisperse spheres. As with the mono- and bidisperse cases, the network properties are also analysed. Systems at five different number densities in the range $\rho = 0.00 \to 0.20$, are initialised with radii drawn from the lognormal distribution in equation (7.9), with $\mu = 0.0, \sigma = 0.3$. Configurations are then generated using the Monte Carlo technique outlined in section 7.1.2. Tessellations for each sample configuration can be made on-the-fly using both a weighted 2D

Voronoi and the full 3D weighted Voronoi cut at a given height. The periodicity in the system ensures that the Voronoi diagrams generated have no boundary, removing any potential edge effects. This maintains the mean polygon size of $\langle k \rangle = 6$, and ensures all node degrees are satisfied in the calculation of the assortativity.

It is important to note that because there are differing number of particles in the tessellations at a given cut height, care must be taken to ensure comparable statistical sampling for each state point. As discussed previously, the greater the cut height, the fewer particles there are with associated cells in the tessellation. Therefore to ensure an equal number of cells are averaged for each value of z, simulations must be repeated with a different number of random seeds, proportional to $1/N(z)$. This sampling also has an impact on the optimal number of particles to include. If this number is too low the statistics will be affected at high z, regardless of number of starting seeds (*e.g.* in the case where only a handful of particles are included in a 2D Voronoi, it becomes impossible to achieve a polygon with a large number of sides). One must therefore balance these statistical considerations with the limits of computational efficiency. In this section simulations included $\mathcal{N} = 1000 \rightarrow 5000$ particles, depending on cut height, so that all tessellations contained at least 50 cells, with a total of $\sim 10^7$ cells sampled for each state point.

To reiterate, the following different Voronoi methods will be compared:

1. The section of a 3D Voronoi weighted with sphere radii, R_i, cut at height z.

2. The sphere-weighted 2D Voronoi weighted according to equation (7.12).

3. The disk-weighted 2D Voronoi weighted according to equation (7.13).

As explained in section 7.4.2 methods 1 and 2 are equivalent and produce identical results. Method 3 on the other hand is expected to approximate to the other methods in the region near $z = 0$.

7.6.1 Cell Frequency and Packing Fraction

To begin with, the number of cells that can be found in each tessellation is evaluated, as the cut height is increased; displayed in figure 7.10a-7.10c. This is measured by the quantity $N(z)$, the proportion of the total particles that have associated cells in the Voronoi construction. For the sphere-weighted Voronoi, figure 7.10a, the number of cells at a given cut height scales with the system density, such that a plot of $N_S(z)$ against $z \times \rho_0$ produces a near universal curve. This curve is fit well by equation (7.19) when $m = 3$ *i.e.* averaging over three nearest neighbours. The curves for $m = 2, 4$ are also presented, lying either side of the $m = 3$ curve, systematically under- and over-estimating $N_S(z)$. The main small discrepancy between systems of different density comes at the low cut heights, where the effects of short range ordering are still observable, as in 7.10b. After the large initial drop in the number of cells at small cut heights, we see a long tail in the function $N_S(z)$. This is indicative of the fact that the cells for the largest particles are very persistent. As these particles are in low density and likely spaced far apart, the dividing planes are near-vertical, leading to a slow convergence of the cell walls.

For the disk-weighted Voronoi, the behaviour in $N_D(z)$ is exactly described by equation (7.17). In this case, the functional form is dependent only on the underlying particle distribution, and not the system density. In contrast to the sphere-weighted case, the number of cells featured in these tessellations decays rapidly to zero beyond $\langle R \rangle$, simply reflecting that the maximum cell height is defined by the particle diameter, and so the polyhedra are truncated instead of extending above the particle. Figure 7.10d explores the change in packing fraction as a function of the cut height, as explained in section **??**. As can be seen the fundamental shape is invariant to the number density, with ρ_0 merely acting as a scaling factor. Once again there is excellent agreement between the numerics and the theoretical results in equations (**??**). Table 7.3 shows how the maximum packing fraction differs from the naive 2D packing fraction that can be calculated from the sphere radii and the cell area as in equation (7.25). Calculating packing fraction in this way always acts to reduce the overall number, and prevents the packing ever exceeding unity.

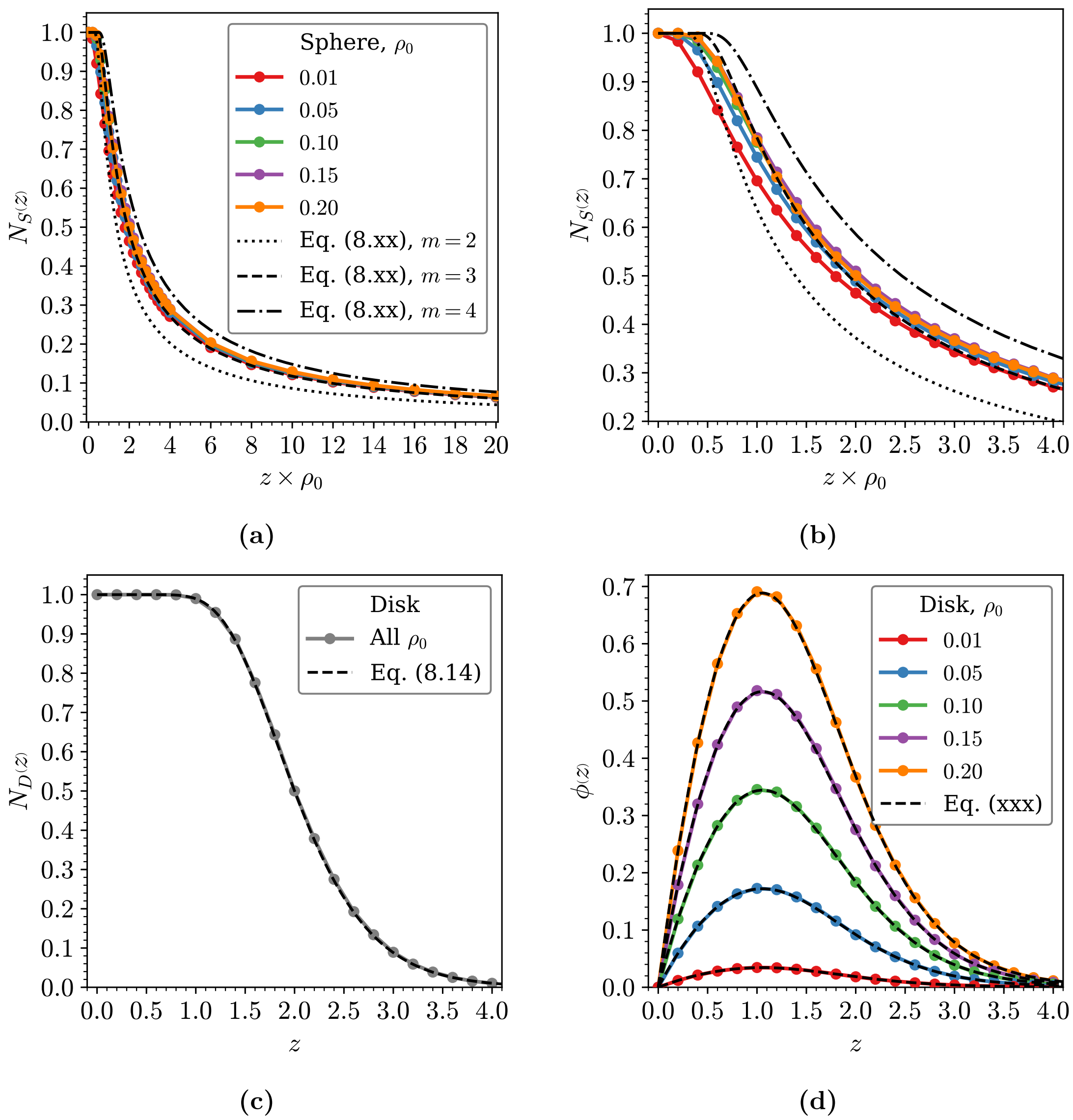


Figure 7.10: Cell frequencies and packing fraction for Voronoi variants. Panels (a) and (c) compare the proportion of particles with cells in the sphere- and disk- weighted Voronoi tessellations respectively. Panel (b) gives a closer inspection of the data in panel (a) over the low z range, using the same legend. Panel (d) gives the change in packing fraction as a function of the horizontal cut height.

Table 7.3: Difference between the maximum packing fraction as a function of height above the surface, and the 2D packing fraction using sphere radii for different number densities.

ρ	$\max\left[\phi\left(z\right)\right]$	ϕ_{2D}
0.05	0.172	0.188
0.10	0.344	0.376
0.15	0.515	0.564
0.20	0.687	0.752

7.6.2 Network Properties

The effects of the different methods of dividing space on the network properties of the system are also a subject of interest. Previous studies which have compared the tessellations formed from cuts through 3D packings of mono- and bidisperse spheres have found differences when comparing the results of the different methods [193, 194]. For the different Voronoi methods discussed in this work, it is found that there is good agreement at low cut heights, but fundamental differences in the asymptotic limit as the cut height is increased. The two network metrics common in this thesis are again compared for each system: the proportion of hexagons, p_6, and the assortativity, r, displayed in figure 7.11. As expected there is good agreement between methods in both metrics at low cut height, with convergence in the limit of $z \to 0$. However, at larger z key differences emerge. The disk-weighted Voronoi approaches the 2D Poisson-Voronoi (PV) limit, $p_6 \approx 0.295$ and $r \approx -0.15$. This is the limit which is obtained from having unweighted points randomly located in 2D space. In contrast the sphere-weighted Voronoi approaches a different, less well understood, limit with more diverse ring distribution and reduced ring-ring correlations. We will discuss the nature of this limit in a following section.

To explain these observations two further measures are examined, the expected weight at a given cut height, $\langle w\rangle_z$, and the average nearest neighbour distance at a given cut height, $\langle r'\rangle_z$, given in figure 7.12. These can be referred to in reference to equation (3.27), which defines the position of the radical plane between two particles. For the disk-weighted Voronoi, it was showed above that the system approached the random PV limit. In this limit the particles are randomly positioned

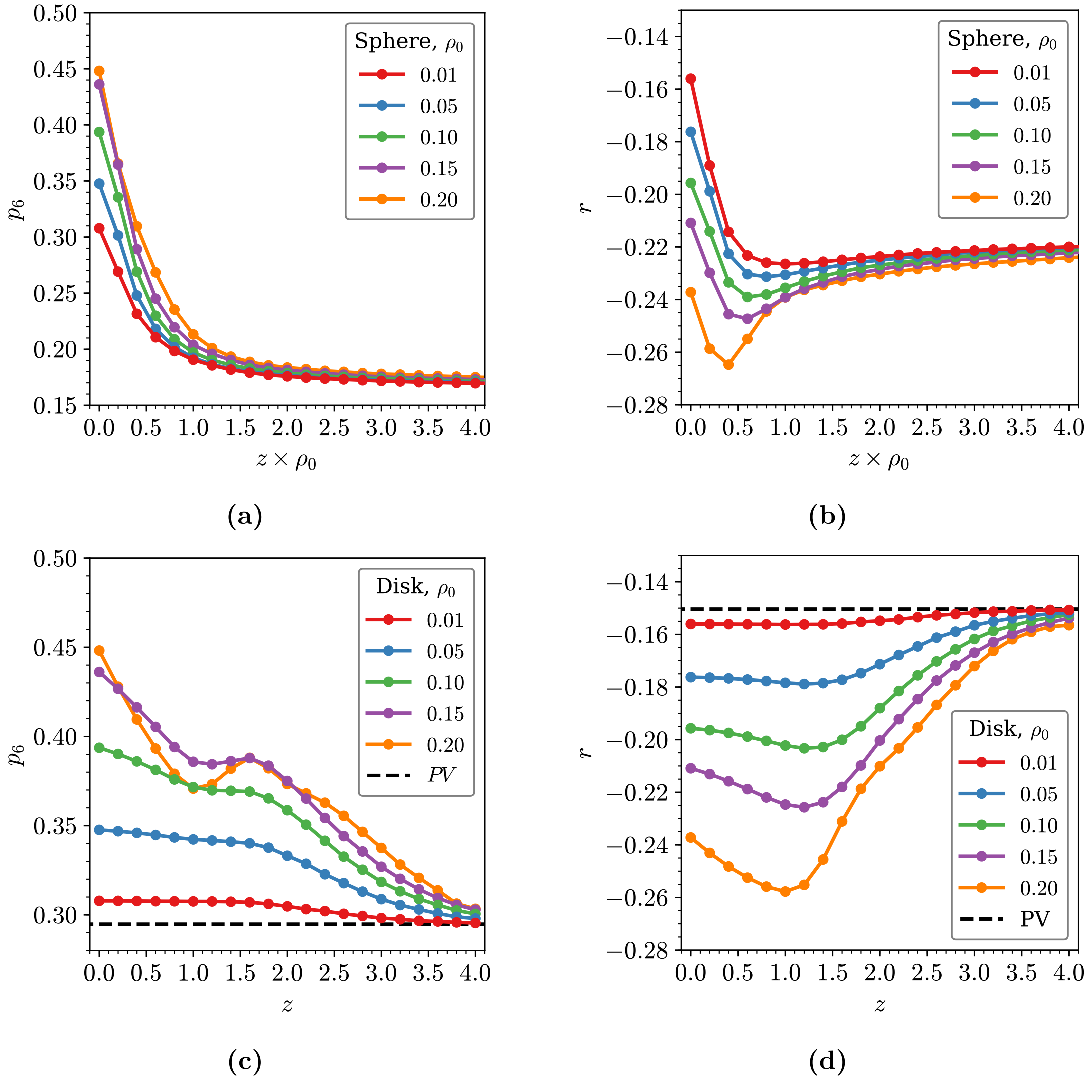


Figure 7.11: Comparison of the network properties in the sphere-weighted and disk-weighted Voronoi diagrams. The proportion of hexagons, p_6, and assortativity, r, with cut height is given in panels (a),(b) and (c),(d) for the sphere- and disk-weighted diagrams respectively.

and effectively unweighted. The only way this is achievable is if the inter-particle distances scale in excess of the weightings with increasing cut height. As seen from figures 7.12c,7.12d, this is indeed the case, with the distances increasing exponentially as the cut height exceeds an increasing number of particle diameters whilst the average weighting remains relatively constant. Two further observations arise from these plots: equation (7.15) describes the average distances very well (with the greatest deviation for higher density systems reflecting the presence of short range liquid structure), and the form of $\langle w_D \rangle_z$ mirrors the oscillations in the network properties. These oscillations, which are particularly pronounced in p_6 (figure 7.11c), are a result of the balance of weightings with cut height which become increasingly visible at higher densities where the inter-particle distances are smaller.

Applying the same reasoning to the sphere-weighted Voronoi, it is initially difficult to see how the network properties can tend to a limit which is *not* PV. However, this type of behaviour is not unknown, being similar to the observation that taking random sections through 3D PV polyhedra leads to 2D polygons which follow a distribution other than PV [195]. As stated previously, as the cut height is increased particle cells are removed at random from the tessellation, so the system must be approaching some random limit. Examining equation (3.27) again, the only way this cannot be PV is if the particle weightings scale in the same way as the inter-particle distances. Figure 7.12a and 7.12b show that this is indeed the case, with both metrics scaling as $z^{1/2}$. As such, the decreasing density of points is directly offset by the increase in weighting and a random limit other than PV is reached.

7.7 Experimental Interpretation

Having examined the numerical results for qtd hard sphere monolayers, the practical conclusions for experimentation can be briefly summarised. As mentioned previously, for those studying quasi-2D systems experimentally, it is often preferable to analyse configurations using particle positions projected into the $x-y$ plane, using a 2D Voronoi method. Not only is this more practicable, it is also leads to straightforward visualisation and analysis. The most common approach is therefore to use a unweighted 2D Voronoi. This is understandable, as it requires no knowledge of the particle radii, which maybe prove difficult to accurately determine. With bi- or polydisperse particle systems, it was previously unclear that the unweighted 2D Voronoi analysis was in fact wholly appropriate, given that for true 2D systems (*e.g.* polydisperse disks) it is known to allocate space disproportionately.

However, in the second half of this chapter it has been shown that possibly contrary to expectation, the use of an unweighted 2D Voronoi analysis, for polydisperse particles sedimented on a plane, is completely valid and well defined.

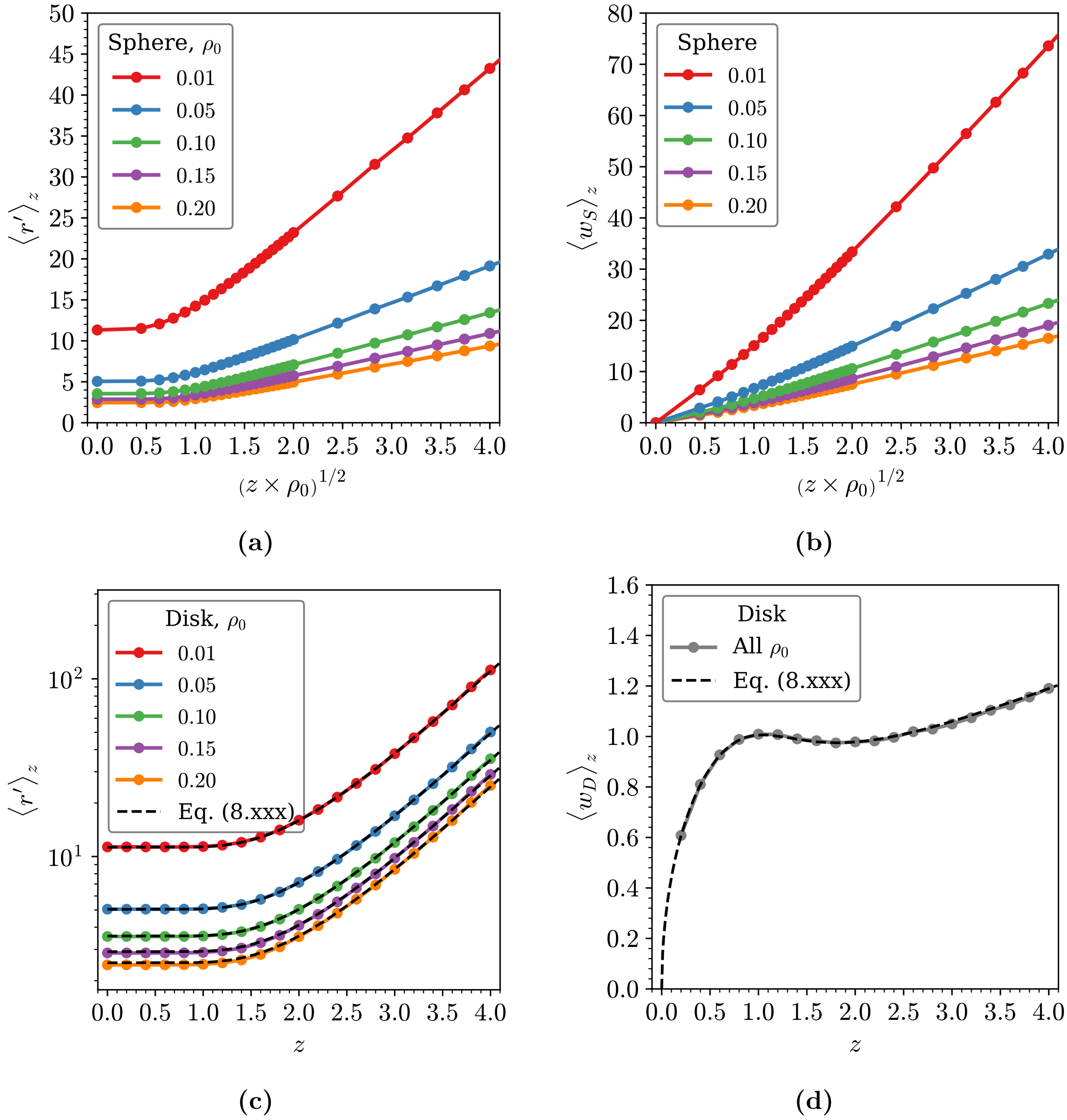


Figure 7.12: Comparison of the inter-particle distances (panels(a),(c)) and particle weights (panels (b),(d)) for the sphere- and disk-weighted Voronoi at increasing cut heights. In the sphere-weighted case both measures scale as $z^{1/2}$ in the high cut height limit, whereas in the disk-weighted case the inter-particle distance increases exponentially, making the weighting effect negligible.

The unweighted 2D Voronoi is topologically equivalent to the tessellation formed by taking the basal faces of the polyhedra formed from the fully *weighted* 3D Voronoi diagram. Therefore, in the absence of the information about particle radii, the unweighted 2D Voronoi can still be mapped to the basal section through the weighted 3D diagram. In addition, for such quasi-2D systems, the unweighted 2D Voronoi diagram may be considered the best construction possible. This is because it corresponds to the only section through the 3D weighted Voronoi which is guaranteed to include a cell for each particle; notwithstanding the fact that it is also the simplest Voronoi method to implement and evaluate. As an aside, this section also seems to maximise the proportion of hexagons rings in the tessellation.

Furthermore, the more general stereology problem has been considered for quasi-2D systems, which is a subject of interest in similar studies of polycrystalline materials [196, 197]. Here it has been demonstrated that the horizontal 2D sections through the 3D weighted Voronoi diagram can be calculated using a 2D weighted Voronoi, with weightings given in equation (7.12). Therefore, if the particle radii are known, one can easily calculate the 2D sections at a given cut height. This is convenient as such computational methods are more readily available than cutting the 3D weighted Voronoi directly, which is a non-standard technique.

Finally an alternative definition of the packing fraction has been introduced by considering the quasi-2D system as a series of hard disk tessellations. This packing fraction is well defined in that it cannot exceed unity and does not require an arbitrary definition of the sample volume.

7.8 Chapter Summary

This chapter has extensively explored the role of the Voronoi construction in the analysis of quasi-2D hard sphere systems. A substantial part of this has been a theoretical study of the relationships between various 2D and 3D tessellations for these systems. Although some of the problems investigated are more esoteric, others are of direct relevance to ongoing experimental research. Most significantly, a link has been drawn between the application of an unweighted 2D Voronoi construction and a 3D construction in which the division of space is weighted in terms of particle size, for the case in which the spheres sit on a surface. As a result, a clear geometrical meaning has been provided to the commonly used unweighted 2D Voronoi diagram; showing it to be equivalent to the tessellation formed from taking the basal polygons in the 3D Voronoi diagram weighted by the sphere radii.

As an example of this, experimental configurations of quasi-2D colloidal monolayers of varying parameters were analysed using Voronoi techniques. The results of these analyses were also compared to analogous configurations generated by non-additive hard disk Monte Carlo, to which there was good agreement. Monodisperse systems were found to have ring statistics concordant with Lemaître's law and network assortativity which was linear in packing fraction. Bidisperse systems were considered in terms of the partial properties of the two sphere components. The partial ring size distributions were shown to be fit well by maximum entropy distributions and can be tuned through varying the system parameters of packing fraction, composition and radius ratio.

8 | Partial Ordering in Procrystalline Lattices

Recent work has introduced the term "procrystalline" to define systems which lack translational symmetry but have an underlying high symmetry lattice. This behaviour arises owing to a difference between the coordination numbers of the molecular units and natural coordination of the underlying lattice. These materials are expected to exhibit structural properties in between crystalline and amorphous phases. The network properties of a range of these procrystals are investigated, encompassing a range of coordination environments. Configurations are generated using a zero-temperature Monte Carlo method, whilst simpler lattices are also considered analytically. Procrystals are shown to be rare examples of systems with violate Lemaître's law, whilst also displaying assortativities different to those calculated for amorphous materials. Procrystalline lattices are therefore shown to have fundamentally different behaviour to traditional disordered and crystalline systems, indicative of the partial ordering of the underlying lattices.

8.1 The Procrystalline State

Investigations into inorganic network-forming materials have led to the introduction of the term "procrystalline" to refer to systems in which molecular building blocks lie on a regular array of lattice points, but directional interactions lead to overall correlated disorder [51]. As an introductory example, consider the procrystal in figure 8.1a. In this configuration the nodes form a square net, but each lattice site is occupied by a "T" shaped unit. If the ends of these units are mutually attractive, they will orient to maximise favourable interactions. The consequence of this is to introduce disorder into the ring structure. This can be detected in the dual

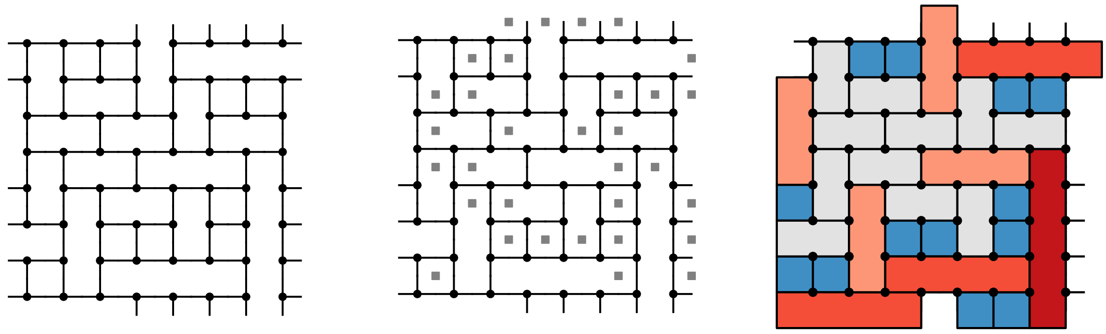


Figure 8.1: Example procrystalline lattice based on the square net. Panel (a) shows the lattice with each node representing a 3-coordinate molecular unit. Panel (b) adds the nodes of the dual network, which form a defective square lattice. Panel (c) highlights the corresponding ring structure, coloured by ring size.

network, as in figure 8.1b, which in this case can be viewed as a defective square net. More strikingly, a system of percolating rings once again emerges, highlighted in figure 8.1c, in analogue with networks from previous chapters.

The local environment around each node in a procrystal is therefore identical, leading them to appear crystalline in their atomic RDFs and structure factors. However, considering the network in its entirety with both nodes and links, it is clear an infinite procrystalline lattice has no unit cell. As such procrystals can be considered to sit somewhere in between traditional crystals and the amorphous materials discussed in previous chapters. This partial ordering is expected to be reflected in their structural and electronic properties.

Experimentally there are several systems which can be thought of as realisations of procrystals. These include self-assembled molecular monolayers, classical bond valence solids, mixed-anion perovskites, and order/disorder ferroelectrics [198–201]. Whilst this list is not extremely extensive (particularly for two-dimensional examples), it demonstrates the diversity in the range of potential structures that can form procrystals. Again although the future focus of the field may be lie more in three-dimensional structures, the constraints and simplifications that arise from reduced dimensionality make two-dimensional structures the natural starting place for investigations into the properties of these procrystalline materials.

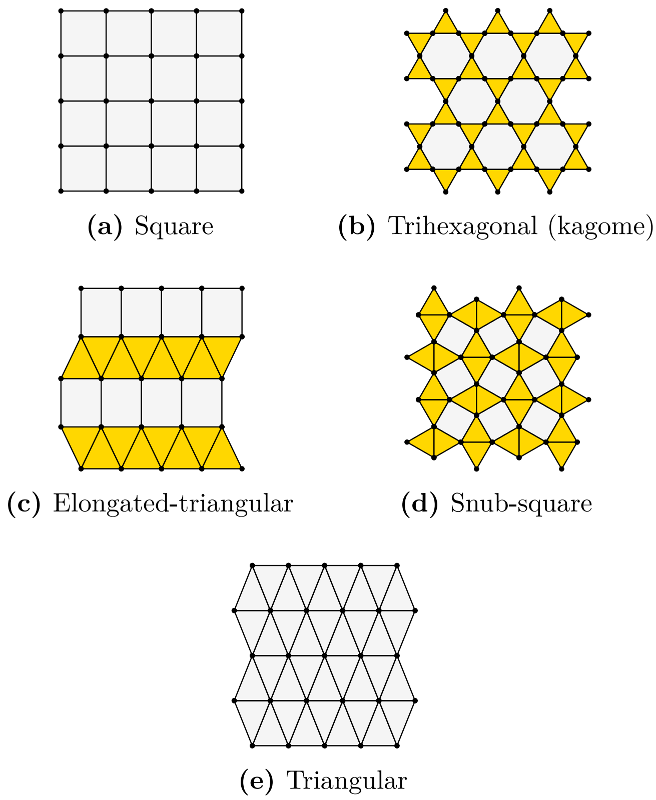


Figure 8.2: High symmetry lattices that form the basis of two-dimensional procrystalline lattices in this work. The square and trihexagonal lattices (panels (a), (b)) have a natural coordination number of 4; the elongated triangular and snub square lattices (panels (c), (d)) of 5 and the triangular lattice (panel (e)) of 6.

8.2 Two-Dimensional Procrystalline Lattices

In this chapter a range of two-dimensional procrystalline systems will be investigated based on a selection of underlying high symmetry lattices and node coordination numbers. Figure 8.2 details the high symmetry lattices that form the basis of these procrystals. These regular and semi-regular tilings have been chosen to provide a series of underlying coordination numbers in the range $4 \rightarrow 6$, whilst also occurring across various theoretical and experimental studies on two-dimensional materials. The 4-coordinate tilings considered are the square and trihexagonal (also known as kagome) nets [202–207], the 5-coordinate tilings are the elongated-triangular and snub-square nets [208–212], and the 6-coordinate tiling is the triangular net.

The disorder in procrystals arises from the discrepancy between the natural coordination numbers of the regular lattice and the actual coordination of the nodes which occupy them. If the coordination of these high symmetry "parent" lattices is denoted c', it follows that each is is able to generate procrystalline lattices with node coordinations, c, in the range $c = 3 \rightarrow (c' - 1)$ (strictly 2-coordinate procrystals are also obtainable, but these form "spaghetti" like structures with ill defined ring structure [213]). In this thesis, for simplicity these procrystals will be referred to with the notation $c', c-$lattices. In addition, whilst for $c', (c' - 1) -$lattices there is only one possible way of arranging the $c' - 1$ links around each node, for the other lattices this is not the case. As a simplification it will be assumed that all arrangements are possible and equally likely.

8.2.1 Network Measures

Procrystals can be considered as additional examples of the tessellating ring structures seen in previous chapters, the difference being that the ring geometries are constrained by the underlying lattice. As such the network measures discussed previously are also applicable to procrystalline lattices. For instance, procrystals are subject to Euler's formula such that the mean ring size is constrained by equation (2.10).

When discussing the maximum entropy ring size distributions, a similar approach can be taken as for Lemaître's law (see section 2.2.2), save constraint (2.16) no longer applies. To get the expected maximum entropy ring statistics, p_k, one can remove this constraint to give a simple modification of equation (2.17):

$$p_k = \frac{e^{-\lambda k}}{\sum_k e^{-\lambda k}} \, . \tag{8.1}$$

As with Lemaître's law, an important additional constraint arises implicitly through the k-range in the summation. Owing to the fact that there are many subtly different systems here, these will be further discussed in the relevant sections link .

Finally, the assortativity will again be used to measure ring-ring correlations.

8.3 Computational Generation of Procrystals

Although experimental procrystalline lattices are aperiodic with no unit cell, computational studies are naturally restricted in scope and an effective way of generating lattices with fully satisfied valence is to reintroduce periodicity. There are then two possible ways to proceed in order to map the configurational space of procrystals. Firstly, one can attempt to find all possible arrangements for a lattice of given size, which here is termed *exact tiling*. Whilst exact tiling gives a complete view of the procrystalline landscape, it will be seen that even with optimisations it quickly becomes computationally intractable.

The second approach is to sample configurations using a stochastic method such as Monte Carlo sampling. This allows procrystalline configurations to be generated which are representative of the wider landscape, for far larger lattice dimensions. This has the advantage of mitigating any effects of enforcing periodicity, and will be the primary method used to generate procrystals in this work.

8.3.1 Exact Tiling Algorithm

The exact tiling algorithm finds all procrystalline lattices for a given lattice dimension using a divide-and-conquer approach. It has been used here to investigate the $4,3-$square lattice. The method starts from the observation that there are seven possibilities for each square in the procrystalline lattice (some of which are symmetry related) forming the 1×1 tiles in figures 8.3a-8.3g (these are colour coded by a central circle). These 1×1 tiles can be stacked to produce 1×2 tiles, of which there are 22 possibilities that satisfy internal coordination requirements (but are not necessarily periodic), the first 5 of which are given in figures 8.3h-8.3l. Again these can in turn be stacked to yield 84, 2×2 building blocks, a selection of which are given in figures 8.3m-8.3t.

It should be clear that this process of combining smaller tiles to form larger tiles can be continued *ad infinitum*, producing tiles of arbitrary dimension $m \times n$. This method is vastly more efficient than a brute force search to obtain the same structures, where notionally each "T" can adopt 4 positions leading to $4^{m \times n}$ possible

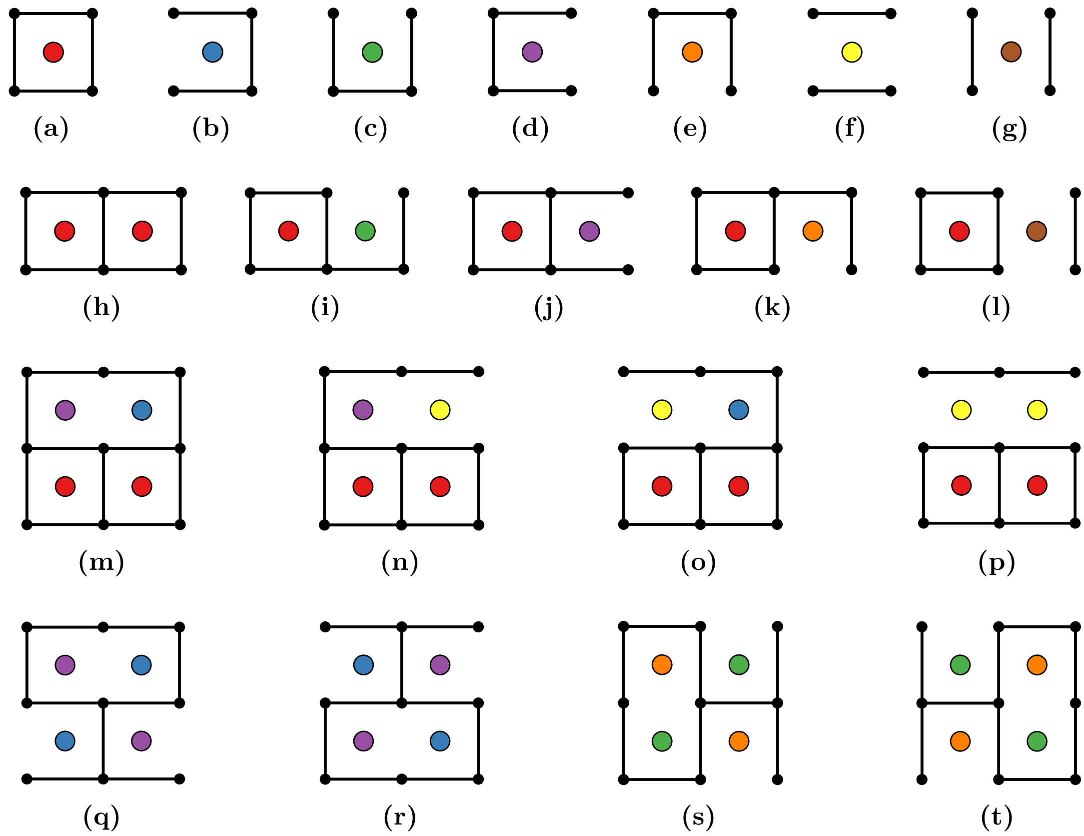


Figure 8.3: Illustration of the process of exact tiling. Panels (a)-(g) give the 7 fundamental 1×1 environments for a square in the $4,3-$square lattice (individually coloured with a central circle, merely intended a guide for the eye). These can be combined to give 2×1 tiles, the 5 acceptable tiles using (a) shown in panels (h)-(l). The process can be repeated to build up larger tiles such as the 2×2 tiles in (m)-(p). Finally the periodic tiles can be identified, the 4 only 2×2 cases given in (q)-(t).

configurations - only a fraction of which satisfy the bonding requirements. The key to the exact tiling method is only to retain units in which the bulk nodes (*i.e.* those not on the perimeter) are all 3-coordinate, thus dramatically reducing the search space. In order to find the periodic procrystals for a $m \times n$ lattice, one then only has to check for units which can tessellate with 3-coordination on the perimeter nodes. For the 2×2 lattice, only 4/84 of the tiles conform to this rule, shown in figures 8.3q-8.3t. It is further evident that these 4 configurations are all in fact symmetry related, and that the only unique solution is in fact the hexagonal tiling.

The ability to leverage symmetry to reduce the search space further is important when looking at larger lattices. This can be achieved by identifying and discarding tiles that are symmetrically equivalent. Care must be taken however to still form

Table 8.1: Performance of the exact tiling algorithm. For each lattice dimension the number of aperiodic, periodic and symmetrically unique periodic tiles are listed. The search space can be found by squaring the number of aperiodic tiles of the previous tile size.

Lattice	Aperiodic	Periodic	Unique
2×2	84	4	1
2×4	1,536	16	4
2×6	27,572	64	8
4×4	87,264	204	9
4×6	4,9147,56	2,368	70
6×6	–	81,736	440

larger tiles by adding the original degenerate set to the reduced set, to avoid losing solutions. A final improvement can be to check for "half-periodicity" when forming $2m \times n$ tiles from $m \times n$ tiles. In this case any units which are not periodic in the fixed dimension can also be discarded before combination takes place.

The application of all the optimisations discussed above serve to make the exact tiling algorithm tractable for a small lattices. Table 8.1 details the performance of the algorithm. Taking even the 4×4 lattice as an example, a naïve search algorithm would require $4^{16} \sim 4 \times 10^9$ iterations compared to the $1536^2 \sim 2 \times 10^6$ for the exact tiling algorithm. The difference for the 6×6 lattice becomes even more marked, spanning some 8 orders of magnitude. However, it is still fighting against the forces of exponential scaling and despite all these improvements in performance, the exact tiling algorithm remains severely limited. Table 8.1 highlights the enormity of the full configurational space and how small a proportion the procrystal solutions are of the total. In order to find all the solutions for larger lattices, a more sophisticated algorithm would be required, although it is hard to see how the hurdle of scaling could be easily overcome.

8.3.2 Monte Carlo Algorithm

To generate procrystals for larger lattice dimensions, a method is required that can quickly search configurational space and find representative samples. As with much of the work in this thesis, this is achieved by utilising a Monte Carlo algorithm.

The algorithm in question has been developed to produce $c', c-$lattices of arbitrary size. It is a zero-temperature Monte Carlo algorithm which proceeds as follows:

1. Initialise the required c'-coordinate periodic lattice from figure 8.2 and assign each node c bonds in a random orientation. This will introduce a number of dangling bonds into the configuration.

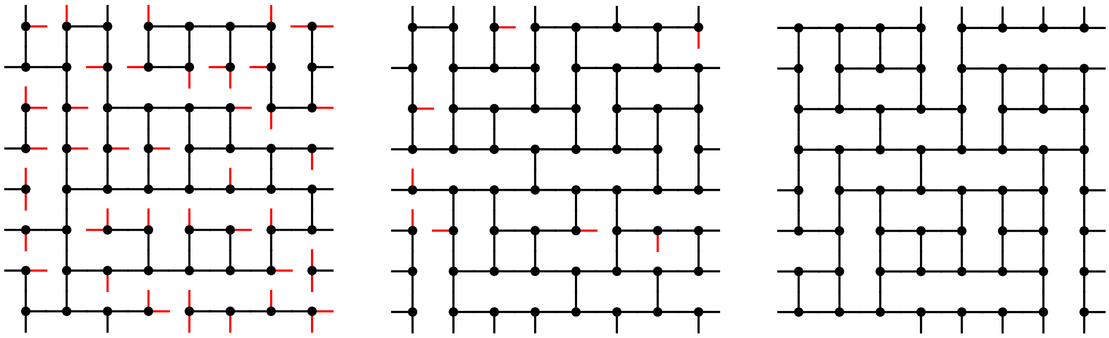
2. Select a node at random and change the orientation of the bonds.

3. If the number of dangling bonds is less than or equal to the number in the previous configuration update the configuration; otherwise revert to the previous structure.

4. Repeat steps 2 and 3 until all dangling bonds have been removed and all node coordinations are satisfied. The final lattice is then in the procrystalline state.

This process is demonstrated for an 8×8, $4, 3-$square lattice in figure 8.4. One aspect of note is that as removing the dangling bonds often requires a correlated motion, it becomes increasingly difficult to remove defects as they reduce in number. Furthermore the structure obtained with a small number of dangling bonds can be quite different to the final procrystalline network as a consequence of the required reorganisation.

This method can be thought of as a simplified version of a site adsorption model, where molecules adsorb to specific sites on an underlying lattice and interact with varying directional potentials [214–216]. The difference is that here the potential model is binary and the aim of the method is to generate a fully coordinate, defect-free "ground state" procrystalline lattice. One could in principle introduce a Metropolis type criterion into step 3 (with the the energy difference reflecting the change in number of dangling bonds) and hence accepting a proportion of "uphill" moves. However, the zero-temperature version is found to converge very well, as there is sufficient flexibility through moves which merely conserve the number of dangling bonds for a global minimum to be reached. In addition, the temperature parameter was not found to appreciably affect the overall properties of the resulting realisations.



(a) Initial random lattice. (b) Partial convergence. (c) Procrystalline lattice.

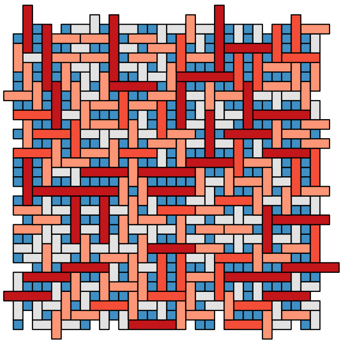
Figure 8.4: Stages in the Monte Carlo search for $4,3-$square procrystalline lattices. Panel (a) gives the initial lattice where each node has 3 bonds in random orientations, panel (b) a snapshot during the search where dangling bonds are being removed and panel (c) the final lattice. Satisfied bonds are coloured black and dangling bonds red.

8.4 Structure of 3-Coordinate Procrystals

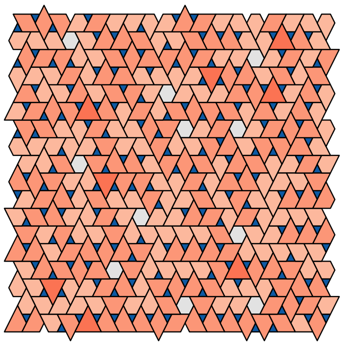
The bulk of the investigation in this chapter will focus on $c', 3-$procrystalline lattices. As stated before, this is because such systems are most prevalent in nature and draw parallels with previous work. To study the ring structure of these procrystals, the Monte Carlo method detailed in 8.3.2 was used to generate configurations for each of the five underlying lattice types, with number of nodes, V, in the lattice scaled to explore system size effects. For each set of parameters some $100,000$ periodic procrystalline lattices were generated. A visualisation of an example configuration based on each parent lattice type is given in figure 8.5 for reference. These configurations highlight some important features of the specific procrystalline lattices which will will be a useful for the coming discussions.

- **$4,3-$square**: only contains even-membered rings in the set $k \in \{4, 6, 8 \cdots\}$. The results from the lack of "cross" bonds (acting between opposite corners of a square). Rings must be linear as any "L"-shapes would require stabilisation of a 2-coordinate site.

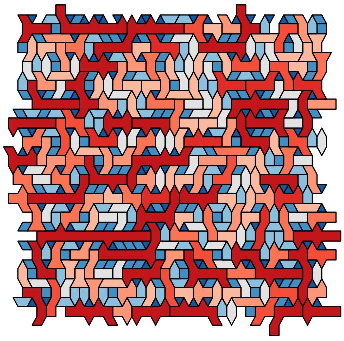
- **$4,3-$trihexagonal**: is yet more constrained, containing only rings in the set $k = \{3, 6, 7, 8, 9\}$. Each "large" ring ($k > 3$) is surrounded by $k - 6$ triangles.



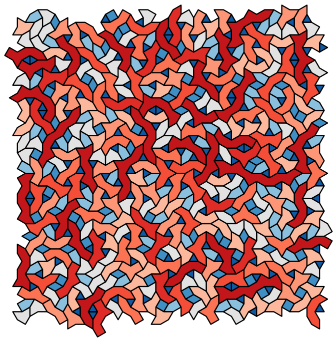
(a) $4, 3-$square



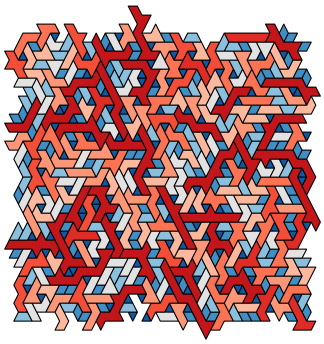
(b) $4, 3-$trihexagonal



(c) $5, 3-$elongated-triangular



(d) $5, 3-$snub-square



(e) $6, 3-$triangular

Figure 8.5: Visualisations of 3-coordinate procrystals based on the 5 different parent lattices (as indicated in panel captions).

- **5, 3−elongated-triangular**: difference between underlying and procrystalline lattice is now 2 and the full ring size range is accessible $k \in \{3, 4, 5 \ldots\}$.
- **5, 3−snub-square**: as above.
- **6, 3−triangular**: difference between underlying and procrystalline lattice is now 3 and again $k \in \{3, 4, 5 \ldots\}$.

Importantly, these 3-coordinate procrystals will be compared and contrasted with two other states. The first are networks generated from bond switching at infinite temperature (see section 3.2, which is in effect a method for producing continuous random networks (CRNs) *i.e.* a fully amorphous state. The second are series of crystalline motifs of the form $8-6^i-5^2$ and $8-6^i-4$ for $0 \leq i \leq 3$ (the nomenclature indicating the number of each ring size in the unit cell) taken from Altman *et al.* [21].

8.4.1 Ring Size Distributions

The first structural measure that will be investigated is the ring size distribution. These will be compared to the maximum entropy (ME) distributions given by equation 8.1, with the accessible k-range for each procrystal outlined in section 8.4 above. For instance, given the $4,3-$square lattice one might expect it to follow a maximum entropy ring distribution of the form:

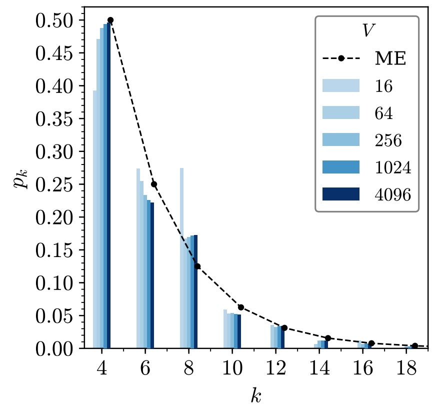
$$p_k = \left(\frac{1}{2}\right)^{k/2-1}, \qquad k \in \{4, 6, 8 \cdots\}\ . \tag{8.2}$$

Similarly for the lattices which can accommodate any ring size, the comparative maximum entropy distribution is:

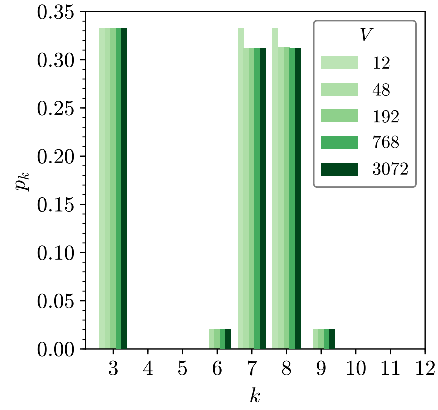
$$p_k = \left(\frac{1}{4}\right)\left(\frac{3}{4}\right)^{k-3}, \qquad k \in \{3, 4, 5 \cdots\}\ . \tag{8.3}$$

These purpose of these maximum entropy distributions is to highlight any discrepancies between the procrystalline systems and an equivalent random lattice.

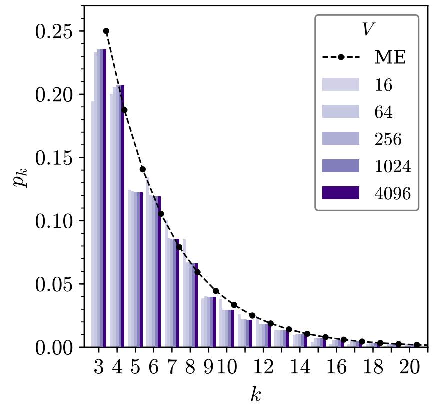
Figure 8.6 shows the ring distributions generated for the five 3-coordinate procrystalline networks shown in figure **??**, as well as from bond switching. These distributions act to highlight how the different tilings impose additional, varying constraints. As discussed above, for the $4,3-$square lattice only even-membered rings are allowed. In addition, the $V = 16$ lattice is small enough to explicitly calculate all the possible lattices using the exact tiling method. These are depicted in figure 8.7 and the corresponding ring statistics are provided in table 8.2. The overall ring statistics of the $V = 16$ case can then be shown to be a simple weighted average of these tilings, namely: $p_4 = \frac{80}{204}$, $p_6 = \frac{56}{204}$, $p_8 = \frac{56}{204}$, $p_{10} = \frac{12}{204}$. This gives



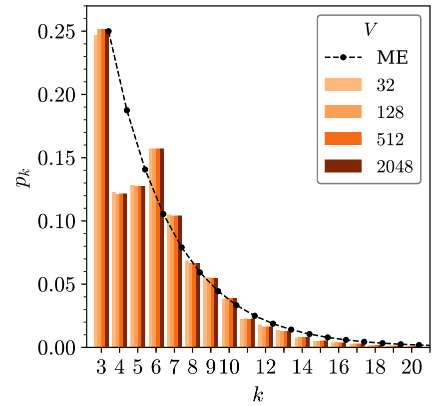
(a) 4, 3−square



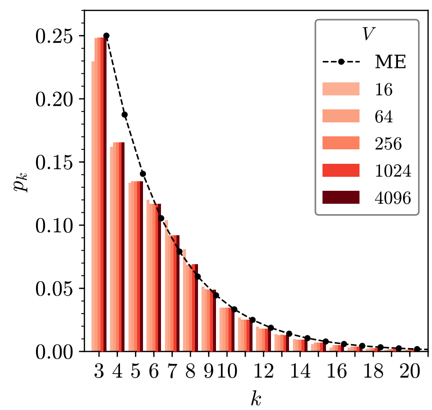
(b) 4, 3−trihexagonal



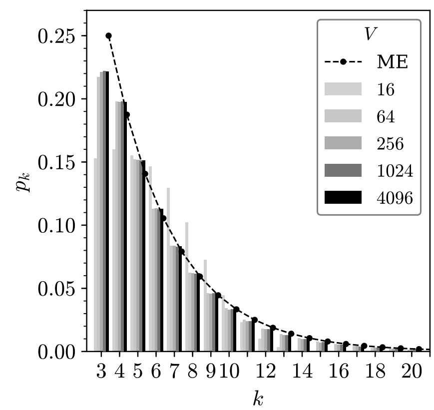
(c) 5, 3−elongated-triangular



(d) 5, 3−snub-square



(e) 6, 3−triangular



(f) CRNs

Figure 8.6: Ring statistics for selected 3-coordinate procrystalline lattices (as highlighted in subfigure captions) and from bond switching (panel (f)). In all panels the points and dashed lines show the respective maximum entropy (ME) solutions. Each panel also highlights potential system size effects by showing the ring size distributions for different numbers of nodes, V, as highlighted in the legends.

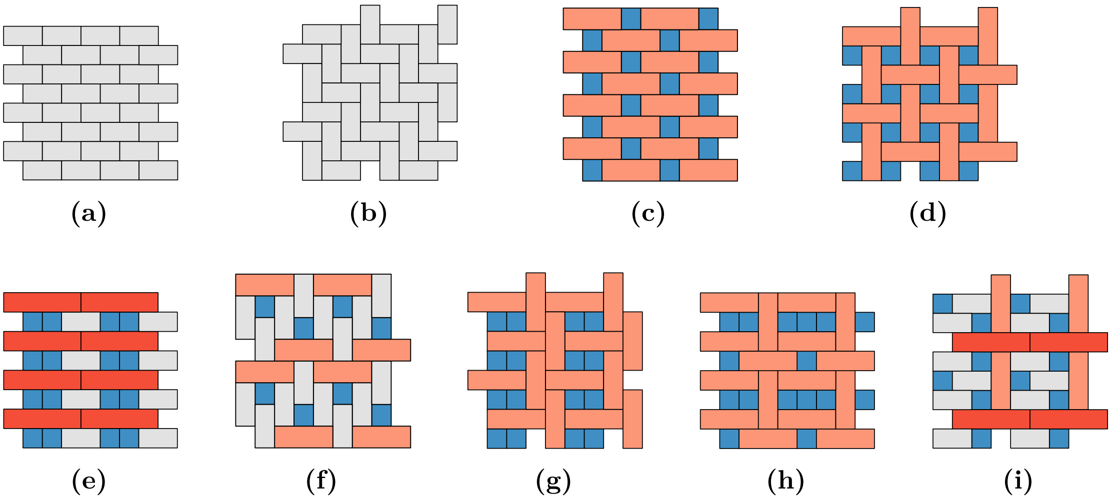


Figure 8.7: The nine unique 4, 3−square lattices for $V = 16$, calculated through exact tiling. Four tessellating units are shown for each solution for clarity.

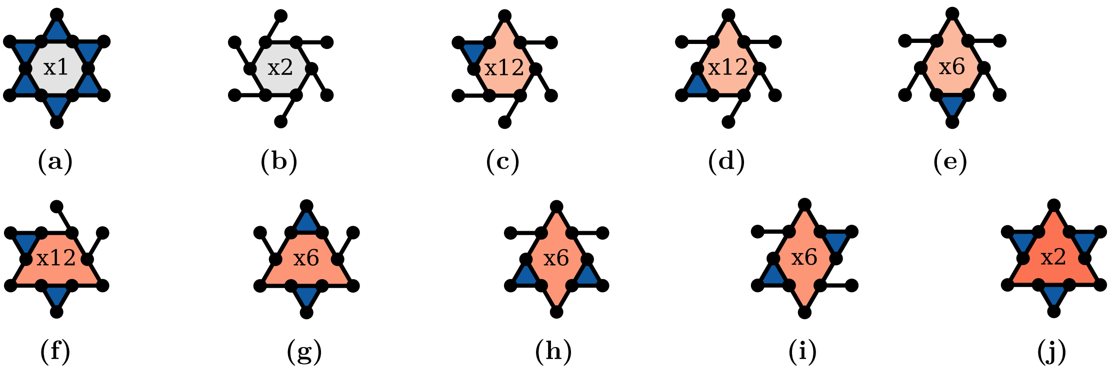


Figure 8.8: Units which can be used to rationalise the ring statistics of the 4, 3−trihexagonal lattice. Panel (a) shows a small $V = 12$ unit of kagome whilst panels (b)-(j) show all possible manifestations of a 6- (b), 7- (c)-(e), 8- (f)-(i) and 9- (j) ring. The numbers in the ring centres indicate the relative degeneracies for each structure.

confidence that the configurations generated via Monte Carlo are appropriately sampling the phase space. As the lattice size increases, the ring statistics initially change markedly, before settling on values close to the ME distribution. This demonstrates that there are important system size effects, the most obvious of which is the largest ring size which can be supported. As only linear rings are allowed, the largest ring possible can have $k = 2\left(V^{1/2} + 1\right)$, which evidently limits the ring statistics. As the lattice dimensions increase and large rings become rarer the statistics naturally converge.

The 4, 3−trihexagonal lattice is clearly the most constrained system, and is

Table 8.2: The fraction of rings of sizes $k = \{4, 6, 8, 10\}$ for the square lattice with $V = 16$ for the configurations (labelled a-i) shown in figure 8.7. W represents the degeneracy of each configuration and hence its weighting in any summation.

Config.	a	b	c	d	e	f	g	h	i
p_4	0	0	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{3}{8}$
p_6	1	1	0	0	$\frac{1}{4}$	$\frac{1}{2}$	0	0	$\frac{3}{8}$
p_8	0	0	$\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{8}$
p_{10}	0	0	0	0	$\frac{1}{4}$	0	0	0	$\frac{1}{8}$
W	4	8	8	8	16	32	32	32	64

simple enough to fully explain analytically. Consider a trihexagonal lattice with V nodes. This parent lattice must have $2V$ edges and V faces by Euler's formula, equation (2.9), $\frac{2V}{3}$ of which are triangles. Generating the procrystal requires removal $\frac{V}{2}$ edges leaving $\frac{V}{2}$ faces. Each edge removed must necessarily remove exactly one triangle so that the final number of triangles is $\frac{V}{6}$. Hence the $4,3-$trihexagonal lattice *must* have $p_3 = \frac{1}{3}$. In addition, this process only allows for rings of size $k = \{3, 6, 7, 8, 9\}$ in the final procrystal. The remainder of the ring statistics can be deduced as follows. Figure 8.8a shows a small unit of the kagome lattice, and figures 8.8b-8.8j the possible resulting procrystalline motifs, with the central number indicating the number of symmetry related species for each motif. This analysis predicts a ratio of $1 : 15 : 15 : 1$ for $p_6 \rightarrow p_9$, leading to ring statistics of $p_3 = \frac{1}{3}$, $p_6 = \frac{1}{48}$, $p_7 = \frac{15}{48}$, $p_8 = \frac{15}{48}$, $p_9 = \frac{1}{48}$. As can be seen in figure 8.6b these are indeed the ring statistics observed for all but the smallest lattice. Again in the $V = 12$ case the 6- and 9-rings cannot be supported and so a uniform distribution results (as must be the case for p_3 and $\langle k \rangle = 6$).

For procrystals with higher parent lattice coordinations, in general the ring statistics become more like the ME solutions on moving from an underlying 4- to 5- to 6-coordinate lattice, reflecting the decrease in constraints along that pathway. Furthermore, the subtleties in the ring statistics for these become more difficult to rationalise, reflective of the increased degrees of freedom. link to later? There are some similarities between the different ring distributions. For example, the snub-square and triangular lattices (figures 8.6d and 8.6e) both show fewer 4- and 5-membered rings, and more 6- and 7-membered rings, when compared to the ME solutions. It is interesting to note that distributions of this general form (*i.e.* dominated by 3-rings and large rings) have been observed previously, for example, for a model using a core-softened potential and long-range repulsions [217], and for models of BN nanotubes encased in amorphous material [218]. Finally, figure 8.6f shows the ring statistics from the bond switching algorithm (and hence corresponding to a high temperature CRN). It is clear that the configurations generated with the procrystalline constraints are fundamentally different purely in terms of the underlying ring statistics.

8.4.2 Lemaître's Law and Assortativity

In addition to the explicit ring statistics, 3-coordinate procrystals can be compared to crystals and CRNs through the second moment of the ring statistics and the assortativity. These are given in figures 8.9a and 8.9b respectively. In addition average values are given in figures 8.9c 8.9d as a function of system size. Before examining the procrystals, the comparative systems can be discussed. The crystalline lattices have well defined ring statistics that are not governed by Lemaître's law but rather by $\mu_2 = 2\,(1 - p_6)$ and $\mu_2 = 4\,(1 - \mu_2)$, for $8 - 6^i - 5^2$ and $8 - 6^i - 4$ respectively. Measuring the ring-ring correlations through the assortativity is slightly contrived for crystals but can still be done for illustrative purposes and one generally finds highly negative values, indicative of the structural ordering (calculations in appendix). In the other extreme, the CRNs generated at infinite temperature lie on the Lemaître curve around the point of highest entropy *i.e.* with $p_6 \approx 0.105$, due to the effective removal of the enthalpic consideration. The assortativity for these systems is likewise closer to the random limit of $r = 0$.

For four procrystalline lattices (excepting the $4, 3-$trihexagonal case) the width of the ring size distribution, as characterised by the second moment, increases as p_6 decreases. The four cases lie towards the high-μ_2 limit of the Lemaître curve similar to CRNs, as the formation of arbitrarily large rings is not precluded on enthaplic

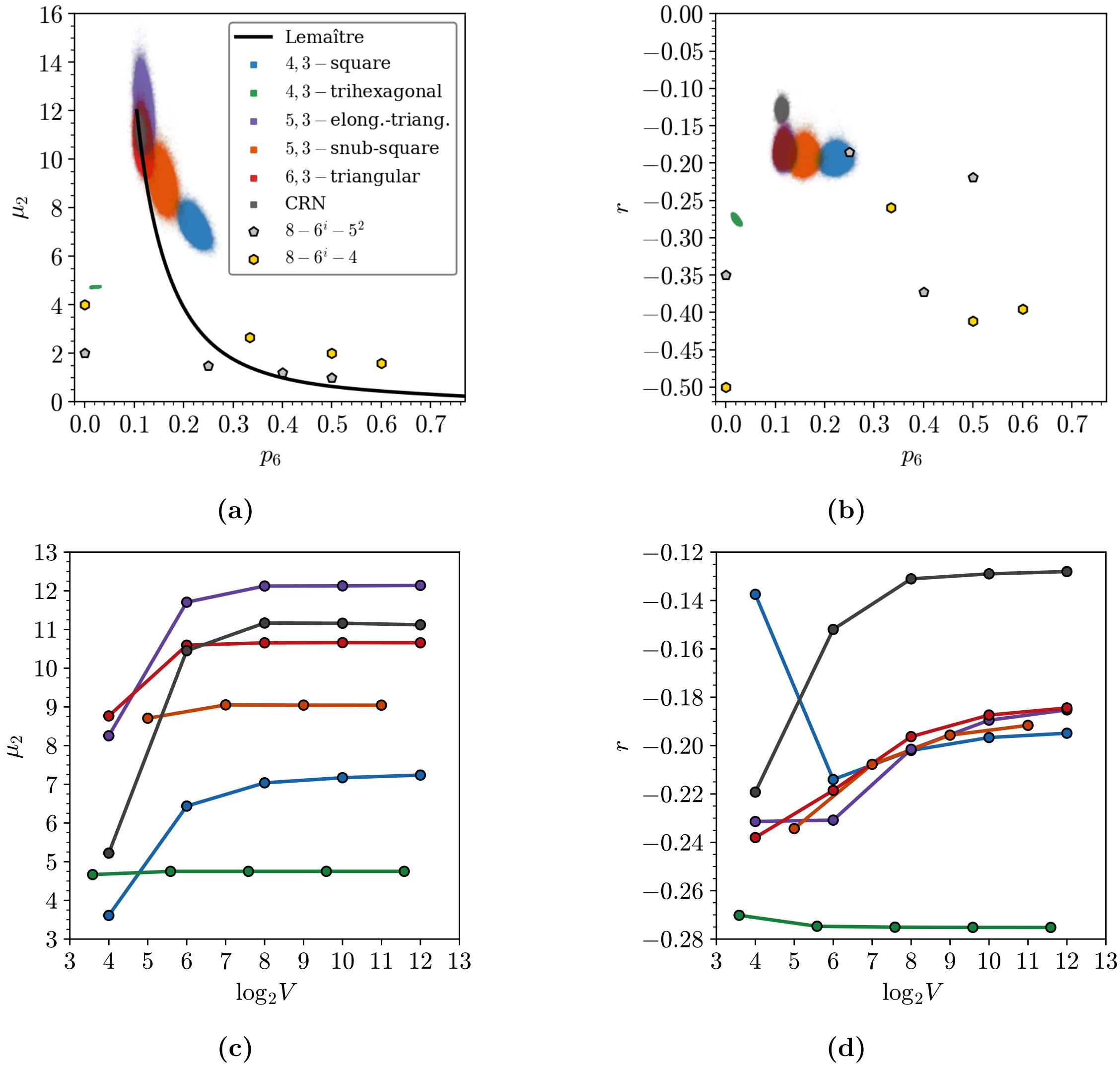


Figure 8.9: Panels (a) and (b) plot the second moment of the ring statistics, μ_2, against the assortativity, r, respectively for the largest lattice dimension investigated for each procrystal, as well as for selected crystals and CRNs. Panels (c) and (d) plot these average values for procrystals of increasing lattice dimensions to highlight any system size effects.

grounds. The $4,3-$square lattice configurations show μ_2 values systematically higher than those predicted from the Lemaître curve. For the five-coordinate lattices the $5,3-$snub-square lattice can be found at μ_2 values significantly higher than those associated with the Lemaître curve (although less removed than those associated with the $4,3-$square lattice), whilst the $5,3-$elongated-triangular lattice lies at high μ_2, again above the Lemaître curve. The second moments generated from the triangular lattice lie closest to the low p_6 ME limit of $p_6 \approx 0.105$, occupied by the CRNs. The exceptional case is once again the $4,3-$trihexagonal procrystal, which has a very well defined μ_2 as a consequence of the constraints on the underlying ring statistics. The configurations generated on the trihexagonal lattice are unique here in lying at both a low p_6 and a relatively low μ_2, and much more in-keeping with systems constrained so as to preclude the formation of large rings (for example, the two-dimensional crystal constructed purely from 4- and 8-membered rings).

The deviation of the second moments from the Lemaître curve is therefore correlated with the strength of the constraints imposed by the underlying crystalline lattice (which decrease from 4- to 5- to 6-coordinate). To reiterate, whilst crystalline lattices are free to locate around the Lemaître curve (their formation usually driven by the energetic landscape), and disordered CRNs constrained to lie upon it; procrystals occupy a region in between these extremes, with the degree of deviation related to the difference in the coordination number of the procrystal and the underlying lattice. This contrasts with previous chapters where it was demonstrated how a very wide range of systems (including atomistic networks, colloidal packings, geopolitical maps *etc*) generated datasets which did fit on the Lemaître curve. The configurations generated here are relatively rare examples of systems which do not.

For the assortativities shown in figure **??**(b), again four of the lattices show similar mean values ($\langle r \rangle \approx -0.19$) corresponding to favouring disassortative configurations. Once again it can be noted that these procrystals occupy the space between crystalline and amorphous systems. Similar to previous observations, the $4,3-$trihexagonal lattice is unique in displaying highly disassortative and well-defined behaviour, with $\langle r \rangle \approx -0.275$. Careful observation of figure 8.8 allows the joint degree distribution to be explicitly written as need appendix on this? :

$$\mathbf{e} = \frac{1}{96} \begin{array}{cc} \begin{array}{ccccc} 3 & 6 & 7 & 8 & 9 \end{array} & \\ \begin{bmatrix} 0 & 0 & 5 & 10 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 5 & 1 & 14 & 14 & 1 \\ 10 & 1 & 14 & 14 & 1 \\ 1 & 0 & 1 & 1 & 0 \end{bmatrix} & \begin{array}{c} 3 \\ 6 \\ 7 \\ 8 \\ 9 \end{array} \end{array} . \tag{8.4}$$

This corresponds to $r = -101/367 \approx -0.275$, as measured from simulation.

8.4.3 System Size Effects

Figure 8.6 highlights potential system size effects in the five procrystalline lattices studied. Most dramatically, the smallest $4,3-$trihexagonal system (containing 12 vertices as shown in figure 8.8) is unable to sustain a 6- or 9-membered ring whilst maintaining full $3-$coordination, in contrast to larger lattice dimensions. This can be rationalised again by examining the configurations from figure 8.8. If a unit containing a 6- or 9-membered ring is taken, then repeating any of these units will automatically generate a 4-coordinate site. For any larger lattice dimension, all the rings in the range $k = \{3, 6, 7, 8, 9\}$ become accessible and the maximum ring size ceases to evolve. This behaviour is in contrast to the other procrystals, in which the maximum ring size always scales with the lattice size (*e.g.* as discussed for the square lattice in section 8.4.1). This means that the system size effects will never completely disappear. What is evident from figure 8.6 though, is that these larger ring sizes become increasingly improbable, and so their effect on the structural metrics should abate with increased system size.

To investigate this, figures 8.9c,8.9d shows the evolution of μ_2 and r as a function of system size (as characterised by the number of vertices in the lattice, V). It is clear that the different systems display structural properties which converge with system size over different length-scales. As expected, both metrics converge quickly for the $4,3-$trihexagonal lattice, owing to the constraints on the ring sizes discussed above. For the remaining procrystals, μ_2 converges at approximately the same rates for the different lattices, and the second moments must increase with system size variationally. In addition, the assortativities appear to converge more slowly than μ_2, reflecting the higher sensitivity of this metric to the potential existence of very large rings.

The $4,3-$square procrystal in particular appears to converge relatively slowly with system size. This can be ascribed to the restriction of forming linear rings only being a stronger constraint on the maximum ring size than if non-linear rings were allowed. Put more simply, a given square area can contain a larger non-linear ring than linear. Similar arguments apply to the two 5-coordinate procrystals. The $5,3-$elongated-triangular procrystal contains alternating chains of percolating squares and triangles the the former again preferentially promoting the formation of more linear rings (evident 'by eye' in figure 8.5c). On the other hand the $5,3-$snub-square lattice can be considered to be more closely aligned to the $6,3-$triangular tiling. In general, it is clear that the effect of the constraints imposed by the underlying high symmetry lattices is to promote a slower convergence with length-scale.

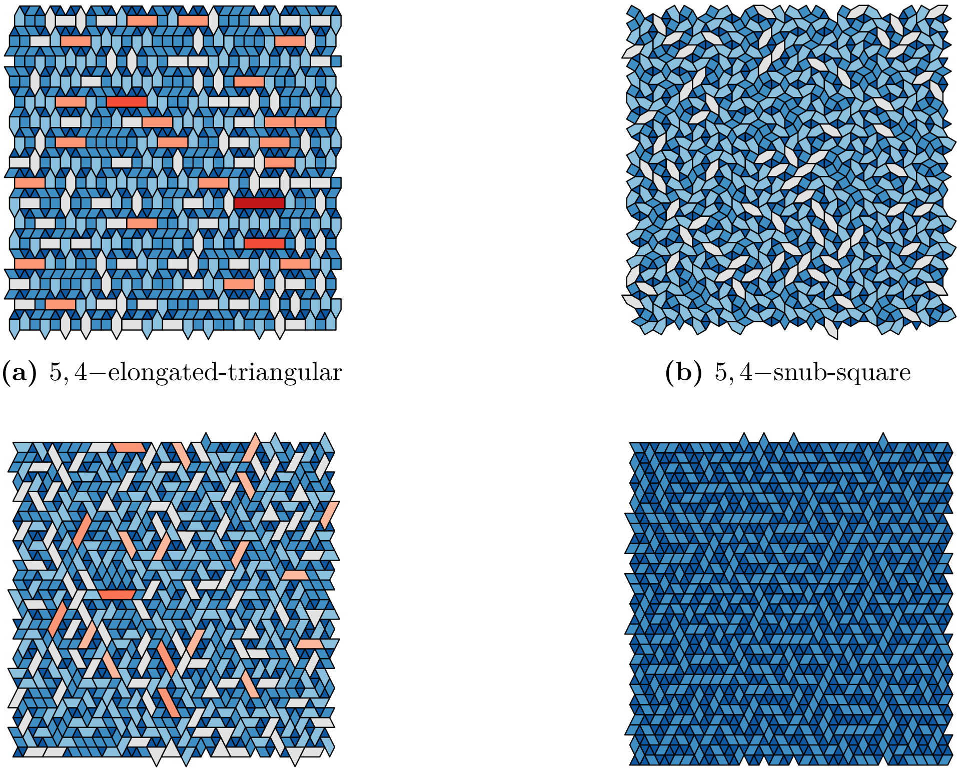
8.5 Structure of Higher-Coordinate Procrystals

To extend the analysis of procrystalline lattices further, similar investigations can be carried out on 4- and 5-coordinate procrystals; based on a subset of the same underlying lattices. From equation (2.10), these systems will have mean ring sizes of $\langle k\rangle = 4$ and $\langle k\rangle = \frac{4}{3}$ respectively. Again the Monte Carlo method in section 8.3.2 was used to generate 100,000 periodic samples of these procrystals, across a range of system sizes. Example visualisations of each lattice type are given in figure 8.10. Unlike networks elsewhere in this thesis, the rings in these visualisations are not coloured relative to the mean ring size. Instead they are coloured in the same way as the configurations in figure 8.5, in order to aid comparison and show evolution of the structure. Again these configurations highlight some important features in the specific procrystalline lattices:

- $\mathbf{5,4-}$**elongated-triangular**: supports all even ring sizes but only the two smallest odd ring sizes i.e. $k \in \{3,4,5,6,8,10\dots\}$. This is due to rings with $k > 6$ only being formed from linear rings in analogy with the $4,3-$square lattice, which must have an even number of nodes.

- $\mathbf{5,4-}$**snub-square**: contains only rings in the set $k \in \{3,4,5,6\}$.

- $\mathbf{6,4-}$**triangular**: difference between underlying and procrystalline lattice is 2, and all ring sizes are accessible $k \in \{3,4,5\dots\}$.



(a) $5,4-$elongated-triangular

(b) $5,4-$snub-square

(c) $6,4-$triangular

(d) $6,5-$triangular

Figure 8.10: Visualisations of 4- and 5-coordinate procrystals based on 3 different parent lattices (as indicated in panel captions). Rings are coloured as for 3-coordinate networks *i.e.*grey $k = 6$; blue $k < 6$; red $k > 6$.

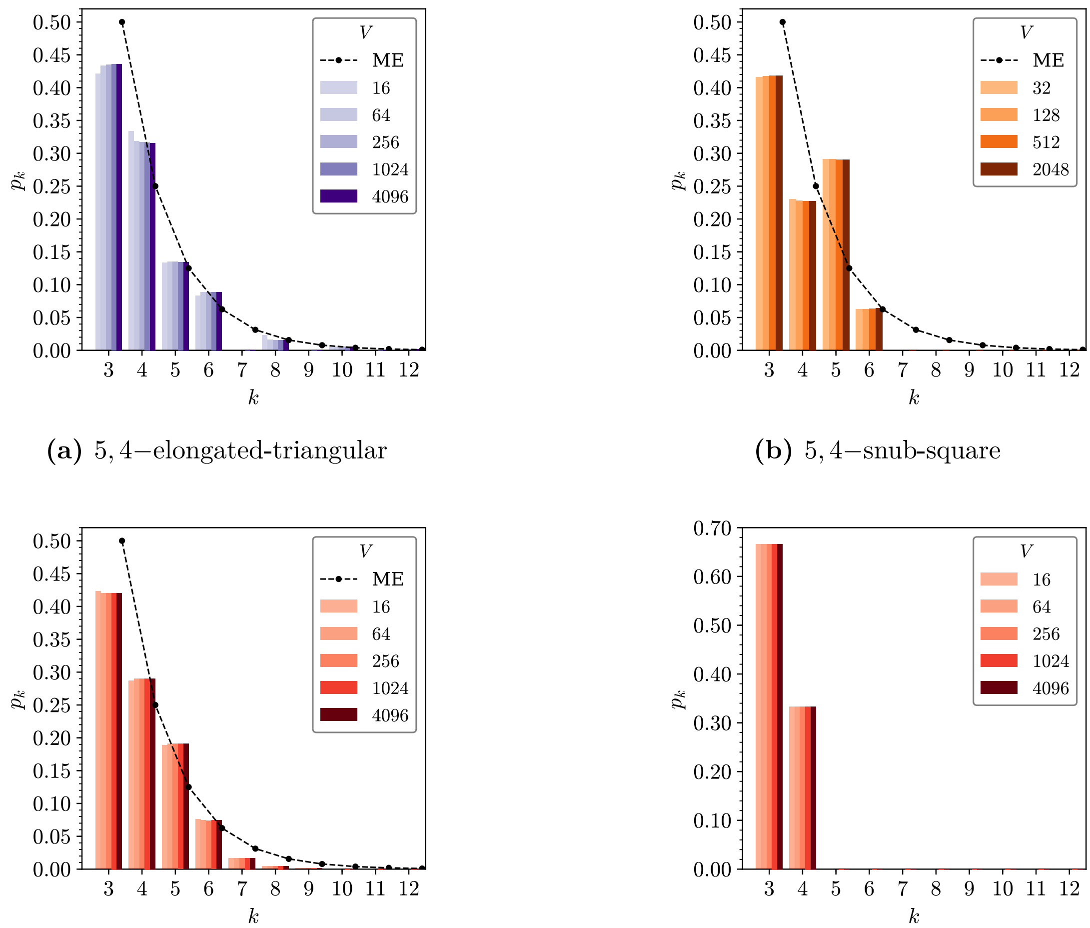
- **$\mathbf{6,5-}$triangular**: only two ring sizes are possible, namely $k \in \{3,4\}$. This is because the procrystal can only be formed by removing a bond between a pair of edge-sharing triangles.

8.5.1 Ring Size Distributions

The starting point for analysing higher coordinate procrystals is again with the ring size distributions. For the 4-coordinate procrystals, the ring statistics will be again compared to the maximum entropy (ME) solution. Assuming the full k-range in accessible, where $\langle k \rangle = 4$ one might expect the maximum entropy distribution:

$$p_k = \left(\frac{1}{2}\right)^{k-2}, \qquad k \in \{3,4,5\dots\}\,. \tag{8.5}$$

The purpose of the ME ring distribution is to highlight the differences in the observed distributions as a result of the specific lattice constraints.



(a) 5, 4−elongated-triangular

(b) 5, 4−snub-square

(c) 6, 4−triangular

(d) 6, 5−triangular

Figure 8.11: Ring statistics for selected 4- and 5-coordinate procrystalline lattices (as highlighted in subfigure captions). Each panel also highlights potential system size effects by showing the ring size distributions for different numbers of nodes, V, as highlighted in the legends.

The ring statistics for the 4 procrystals discussed in this section can be found in figure 8.11. It is also an interesting exercise to compare these distributions to those found for the 3-coordinate procrystals in figure 8.6, to see the evolution in structure. Each lattice will now briefly be examined in turn. As previously mentioned, the 5, 4−elongated-triangular lattice cannot support any odd ring sizes greater than $k = 5$. This is due to the large rings being only formed from the rows of square units, which necessarily makes them even in size. Odd rings can only be manufactured by merging “across” square and triangular rows, which grants access to the 5-ring but no higher. This behaviour percolates through to the 5, 3−elongated-triangular lattice, which can be seen to have slightly depressed values of the odd ring sizes when compared to the ME solution.

The $5,4-$snub-square lattice bears some comparison to the $4,3-$trihexagonal lattice, in that the lattice imposes strong constraints on the obtainable ring sizes. Here it can be seen that the 5-ring is particularly favoured, being formed from the merger of a 3- and 4-ring in the parent lattice. Once more, the memory of this structure can be seen in the $5,3-$snub-square lattice, which has a distinctive spike in the proportion of hexagons. On transitioning to the lower node coordination, the $5,4-$snub-square lattice can generate a hexagon by combination of either of the abundant 5-3 or 4-4 pairings. Conversely, there are few initial hexagons to be lost.

For the triangular-based procrystals, it makes sense to begin with the $6,5-$triangular lattice. This can be simply rationalised, having only two ring sizes, with statistics $p_3 = \frac{2}{3}$ and $p_4 = \frac{1}{3}$, required by Euler's formula. On removal of another degree of freedom, to form the $6,4-$triangular lattice, as with the other lattices all ring sizes immediately become accessible. In the case of the triangular lattice this is facilitated by in the introduction of "diagonal" rings. Whilst some history of the preceding lattice can be detected in the $6,4-$triangular lattice (with an over abundance of small rings), this quickly washes out when obtaining the final $6,3-$triangular procrystal.

8.5.2 Lemaître's Law and Assortativity

Higher-coordinate procrystals can also be viewed in the context of Lemaître's law and their assortativity. These are explored in figures 8.12a and 8.12b respectively, with p_4 plotted against μ_2 and r. Similar behaviour is seen as for the 3-coordinate cases discussed previously, and so this analysis will be covered relatively briefly.

In terms of Lemaître's law, the discrepancy between the random limit and the procrystals is increased as the number of constraints increases. Hence the $6,5-$triangular lattice appears almost crystalline in behaviour, but the $6,4-$triangular lattice is much closer to a CRN. The spread of the data is also correlated with the number of available ring sizes. Hence the $5,4-$snub-square lattice has quite a tight distribution

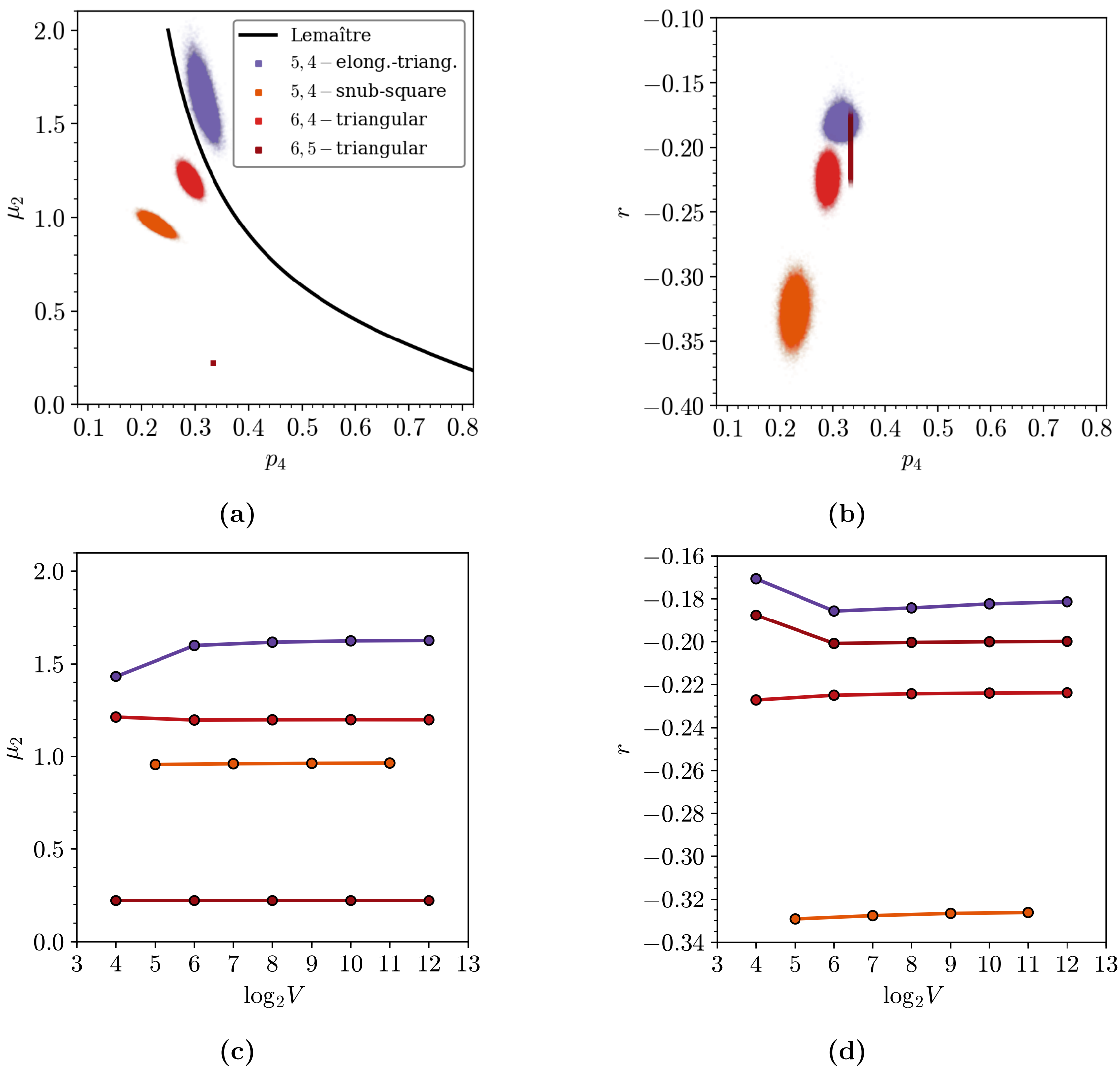


Figure 8.12: Panels (a) and (b) plot the second moment of the ring statistics, μ_2, against the assortativity, r, respectively for the largest lattice dimension investigated for each procrystal. Panels (c) and (d) plot the average values for procrystals of increasing lattice dimensions to highlight any system size effects.

(similar to the 4, 3−trihexagonal case), with the 5, 4−elongated triangular procrystal points being much more diffuse.

The assortativities are also in keeping with the information learnt from the 3-coordinate procrystals. The 5, 4−snub-square lattice has an assortativity similar to the 4, 3−trihexagonal case, likely a reflection of both these systems having a highly constrained set of ring sizes. The remaining procrystals have assortativity values in the now-familiar range of $r \approx -0.2$. Interestingly, despite its precisely defined ring statistics, the 6, 5−triangular lattice still exhibits a range of assortativities comparable to the other procrystals. This illustrates again that the ring statistics are insufficient to fully describe the structure of a given ring system.

8.5.3 System Size Effects

The effects of system size on the network properties of higher-coordinate procrystals are shown in figure 8.12c and 8.12d. All these procrystals show fast convergence of μ_2 and r with the number of vertices in the system, V. This is as expected, as there are no dramatic changes in the accessible ring sizes as the underlying lattice dimensions increase. In particular, the ring statistics are largely constant, even for procrystals in which the maximum ring size continues to grow with the lattice dimension *i.e.*$5,4-$elongated-triangular and $6,4-$triangular. This is because very large rings are in vanishingly small abundance for these procrystals.

As found for the 3-coordinate procrystals, the assortativity proves slower to converge. This holds true even for the $6,5-$triangular lattice, which has well-determined ring statistics. Whilst the proportion of squares and triangles must be constant, the smallest lattice dimensions can still not support certain motifs (*e.g.* pyramids of triangles), which naturally restricts the obtainable assortativity. In general, all the higher-coordinate procrystals still show rapid convergence, reflecting the relatively rigid underlying constraints.

8.6 Chapter Summary

This chapter has applied the tools of network theory to analyse two-dimensional examples of recently-defined procrystalline lattices. These procrystalline configurations, generated by Monte Carlo methods, have been shown to have fundamentally different structural properties to both crystalline and amorphous arrangements. This has been demonstrated through the violation of Lemaître's law, and measured assortativities atypical of more well-understood systems.

These two-dimensional systems provided a good starting place for investigations into the procrystalline state, because they are simpler to understand; for instance by having well-defined ring structure. Extensions to this work pose exciting possibilities. If these results are mirrored by equally anomalous ring statistics in three-dimensional procrystalline networks, one might expect a variety of physical properties that depend on correlation to be affected in otherwise unexpected ways. For example, the disordered pore networks of Prussian blue analogues possess topological characteristics that differ meaningfully from those of random or ordered porous media, in turn influencing their transport properties [219]. On a different lengthscale, photonic procrystals should exhibit photonic band structures different to those of both ordered and amorphous phase [220, 221].

Appendices

References

[1] W H Zachariasen. “The Atomic Arrangement in Glass”. In: *J. Am. Chem. Soc.* 54.10 (1932), pp. 3841–3851.

[2] J. Kotakoski et al. “From point defects in graphene to two-dimensional amorphous carbon”. In: *Phys. Rev. Lett.* 106 (2011), p. 105505.

[3] Alex W. Robertson et al. “Spatial control of defect creation in graphene at the nanoscale”. In: *Nat. Commun.* 3 (2012), p. 1144.

[4] Pinshane Y Huang et al. “Direct Imaging of the a Two-Dimensional Silica Glass on Graphene”. In: *Nano Lett.* 12 (2012), pp. 1081–1086.

[5] Leonid Lichtenstein, Markus Heyde, and Hans Joachim Freund. “Atomic arrangement in two-dimensional silica: From crystalline to vitreous structures”. In: *J. Phys. Chem. C* 116 (2012), pp. 20426–20432.

[6] Shamil Shaikhutdinov and Hans-joachim Freund. “Metal-Supported Aluminosilicate Ultrathin Films as a Versatile Tool for Studying the Surface Chemistry of Zeolites”. In: *ChemPhysChem* 14 (2013), pp. 71–77.

[7] Adrián Leandro Lewandowski et al. “Atomic structure of a metal-supported two-dimensional germania film”. In: *Phys. Rev. B* 97 (2018), p. 115406.

[8] L Lewandowski et al. “From Crystalline to Amorphous Germania Bilayer Films at the Atomic Scale: Preparation and Characterization”. In: *Angew. Chem. Int. Ed.* 58 (2019), pp. 10903–10908.

[9] Panagiotis Trogadas, Thomas F Fuller, and Peter Strasser. “Carbon as catalyst and support for electrochemical energy conversion”. In: *Carbon N. Y.* 75 (2014), pp. 5–42.

[10] Yongfu Sun et al. “Ultrathin Two-Dimensional Inorganic Materials : New Opportunities for Solid State Nanochemistry”. In: *Acc. Chem. Res.* 48 (2015), pp. 3–12.

[11] Christin Büchner and Markus Heyde. “Two-dimensional silica opens new perspectives”. In: *Prog. Surf. Sci.* 92 (2017), pp. 341–374.

[12] Paul A Beck. “Annealing of cold worked metals”. In: *Adv. Phys.* 3.11 (1954), pp. 245–324.

[13] C G Dunn and E F Koch. “Comparison of Dislocation Densities of Primary and Secondary Recrystallization Grains of Si-Fe”. In: *Acta Metall.* 5 (1957), p. 548.

[14] A J Stone and D J Wales. “Theoretical Studies of Icosahedra C60 and Some Related Species”. In: *Chem. Phys. Lett.* 128.5,6 (1986), pp. 501–503.

[15] J. Shackelford and B. D. Brown. “The Lognormal Distribution in the Random Network Structure”. In: *J. Non. Cryst. Solids* 44 (1981), pp. 379–382.

[16] J Lemaitre et al. “Arrangement of cells in Voronoi tesselations of monosize packing of discs”. In: *Philos. Mag. B* 67.3 (1993), pp. 347–362.

[17] Leonid Lichtenstein et al. “The atomic structure of a metal-supported vitreous thin silica film”. In: *Angew. Chemie - Int. Ed.* 51 (2012), pp. 404–407.

[18] D A Aboav. “Arrangement of grains in a polycrystal”. In: *Metallography* 3 (1970), pp. 383–390.

[19] D. Weaire. “Some remarks on the arrangement of grains in a polycrystal”. In: *Metallography* 7 (1974), pp. 157–160.

[20] Torbjörn Björkman et al. “Defects in bilayer silica and graphene: Common trends in diverse hexagonal two-dimensional systems”. In: *Sci. Rep.* 3 (2013), p. 3482.

[21] Andrei Malashevich, Sohrab Ismail-Beigi, and Eric I. Altman. “Directing the structure of two-dimensional silica and silicates”. In: *J. Phys. Chem. C* 120 (2016), pp. 26770–26781.

[22] Mark Wilson et al. “Modeling vitreous silica bilayers”. In: *Phys. Rev. B* 87 (2013), p. 214108.

[23] Mark Wilson and Harry Jenkins. “Crystalline thin films of silica : modelling , structure and energetics”. In: *J. Phys. Condens. Matter* 30 (2018), p. 475401.

[24] Jin Zhang. “Phase-dependent mechanical properties of two-dimensional silica films: A molecular dynamics study”. In: *Comput. Mater. Sci.* 142 (2018), pp. 7–13.

[25] Franz Bamer, Firaz Ebrahem, and Bernd Markert. “Athermal mechanical analysis of Stone-Wales defects in two-dimensional silica”. In: *Comput. Mater. Sci.* 163 (2019), pp. 301–307.

[26] Projesh Kumar Roy and Andreas Heuer. “Ring Statistics in 2D Silica: Effective Temperatures in Equilibrium”. In: *Phys. Rev. Lett.* 122 (2019), p. 016104.

[27] Nina F. Richter et al. “Characterization of Phonon Vibrations of Silica Bilayer Films”. In: *J. Phys. Chem. C* 123 (2019), pp. 7110–7117.

[28] Projesh Kumar Roy, Markus Heyde, and Andreas Heuer. “Modelling the atomic arrangement of amorphous 2D silica: a network analysis”. In: *Phys. Chem. Chem. Phys.* 20 (2018), pp. 14725–14739.

[29] Avishek Kumar et al. “Ring statistics of silica bilayers”. In: *J. Phys. Condens. Matter* 26 (2014), p. 395401.

[30] D. A. Aboav. “The arrangement of cells in a net. I”. In: *Metallography* 13 (1980), pp. 43–58.

[31] B. N. Boots. “Comments on "Aboav’s Rule" for the Arrangement of Cells in a Network”. In: *Metallography* 17 (1984), pp. 411–418.

[32] J. C. Earnshaw and D. J. Robinson. “Topological correlations in colloidal aggregation”. In: *Phys. Rev. Lett.* 72.23 (1994), p. 3682.

[33] C Allain and L Limat. “Regular Patterns of Cracks Formed by Directional Drying of a Collodial Suspension”. In: *Phys. Rev. Lett.* 74.15 (1995), p. 2981.

[34] A Moncho-Jorda, F Martinez-Lopez, and R Hidalgo-Alvarez. “Simulations of aggregation in 2D . A study of kinetics , structure and topological properties”. In: *Physica A* 282 (2000), pp. 50–64.

[35] Marc Durand et al. “Statistical mechanics of two-dimensional shuffled foams: Prediction of the correlation between geometry and topology”. In: *Phys. Rev. Lett.* 107 (2011), p. 168304.

[36] Mingming Tong et al. “Geometry and Topology of Two-Dimensional Dry Foams : Computer Simulation and Experimental Characterization”. In: *Langmuir* 33 (2017), pp. 3839–3846.

[37] Lucas Goehring and Stephen W Morris. “Cracking mud, freezing dirt, and breaking rocks”. In: *Phys. Today* 67.11 (2014), p. 39.

[38] D Brutin et al. “Pattern formation in drying drops of blood”. In: *J. Fluid Mech.* 667 (2011), pp. 85–95.

[39] Franziska Glassmeier and Graham Feingold. “Network approach to patterns in stratocumulus clouds”. In: *PNAS* 114.40 (2017), pp. 10578–10583.

[40] Michel C Milinkovitch et al. “Crocodile Head Scales Are Not Developmental Units But Emerge From Physical Cracking”. In: *Science (80-.).* 339 (2019), pp. 78–81.

[41] G. Le Caër and R. Delannay. “The administrative divisions of mainland France as 2D random cellular structures”. In: *J. Phys. Fr.* 3 (1993), p. 1777.

[42] G Schliecker and S Klapp. “Why are the equilibrium properties of two-dimensional random cellular structures so similar?” In: *Europhys. Lett.* 48.2 (1999), pp. 122–128.

[43] William T. Gibson et al. “Control of the mitotic cleavage plane by local epithelial topology”. In: *Cell* 144 (2011), pp. 427–438.

[44] M Kokalj Ladan, P Ziherl, and A Šiber. “Topology of dividing planar tilings : Mitosis and order in epithelial tissues”. In: *Phys. Rev. E* 100 (2019), p. 012410.

[45] D. Weaire and N. Rivier. “Soap, cells and statistics-random patterns in two dimensions”. In: *Contemp. Phys.* 50.1 (2009), pp. 199–239.

[46] J C Flores. “Mean-field crack networks on desiccated films and their applications : Girl with a Pearl Earring”. In: *Soft Matter* 13 (2017), pp. 1352–1356.

[47] Steven H Strogatz. “Exploring complex networks”. In: *Nature* 410 (2001), p. 268.

[48] S Boccaletti et al. “Complex networks : Structure and dynamics”. In: *Phys. Rep.* 424 (2006), pp. 175–308.

[49] Albert-László Barabási. “The network takeover”. In: *Nat. Phys.* 8 (2012), pp. 14–16.

[50] Alice L. Thorneywork et al. “Two-Dimensional Melting of Colloidal Hard Spheres”. In: *Phys. Rev. Lett.* 118 (2017), p. 158001.

[51] Andrew B Cairns et al. “Design of crystal-like aperiodic solids with selective disorder–phonon coupling”. In: *Nat. Commun.* 7 (2016), p. 10445.

[52] Albert-László Barabási and Márton Pósfai. *Network science.* Cambridge: Cambridge University Press, 2016.

[53] Xianglong Yuan and A N Cormack. “Efficient algorithm for primitive ring statistics in topological networks”. In: *Comput. Mater. Sci.* 24 (2002), pp. 343–360.

[54] D. A. Aboav. “The Arrangement of Cells in a Net. III”. In: *Metallography* 17 (1984), pp. 383–396.

[55] E Ressouche et al. “Magnetic Frustration in an Iron-Based Cairo Pentagonal Lattice”. In: *Phys. Rev. Lett.* 103 (2009), p. 267204.

[56] P W Fowler et al. “Energetics of Fullerenes with Four-Membered Rings”. In: *J Phys Chem* 100 (1996), pp. 6984–6991.

[57] A. Gervois, J. P. Troadec, and J. Lemaitre. “Universal properties of Voronoi tessellations of hard discs”. In: *J. Phys. A* 25 (1992), pp. 6169–6177.

[58] G. Le Caër and R. Delannay. “Correlations in Topological Models of 2d Random Cellular Structures”. In: *J. Phys. A* 26 (1993), pp. 3931–3954.

[59] P Cerisier, S Rahal, and N Rivier. “Topological correlations in Benard-Marangoni convective structures”. In: *Phys. Rev. E* 54.5 (1996), pp. 5086–5094.

[60] Matthew P. Miklius and Sascha Hilgenfeldt. “Analytical results for size-topology correlations in 2D disk and cellular packings”. In: *Phys. Rev. Lett.* 108 (2012), p. 015502.

[61] N Rivier, D Weaire, and R Romer. “Tetrahedrally Bonded Random Networks Without Odd Rings”. In: *J. Non. Cryst. Solids* 105 (1988), pp. 287–291.

[62] F. T. Lewis. “The correlation between cell division and the shapes and sizes of prismatic cell in the epidermis of cucumis”. In: *Anat. Rec.* 38.3 (1928), pp. 341–376.

[63] M. A. Fortes. “Applicability of the Lewis and Aboav-Weaire laws to 2D and 3D cellular structures based on Poisson partitions”. In: *J. Phys. A* 28 (1995), pp. 1055–1068.

[64] Sangwoo Kim, Muyun Cai, and Sascha Hilgenfeldt. “Lewis’ law revisited: the role of anisotropy in size-topology correlations”. In: *New J. Phys.* 16 (2014), p. 015024.

[65] S. N. Chiu. “Aboav-Weaire’s and Lewis’ laws - A review”. In: *Mater. Charact.* 34 (1995), pp. 149–165.

[66] Renaud Delannay and Gérard Le Caër. “Topological characteristics of 2D cellular structures generated by fragmentation”. In: *Phys. Rev. Lett.* 73.11 (1994), pp. 1553–1556.

[67] S Le Roux and F Rezai-Aria. “Topological and metric properties of microscopic crack patterns : application to thermal fatigue of high temperature”. In: *J. Phys. D* 46 (2013), p. 295301.

[68] David A Noever. “Statistics of emulsion lattices”. In: *Colloids and Surfaces* 62 (1992), pp. 243–247.

[69] J. C. M. Mombach, R. M. C. de Almeida, and J. R. Iglesias. “Two-cell correlations in biological tissues”. In: *Phys. Rev. E* 47.5 (1993), pp. 3712–3717.

[70] P Pedro et al. “Polygonal terrains on Mars : A contribution to their geometric and topological characterization”. In: *Planet. Space Sci.* 56 (2008), pp. 1919–1924.

[71] David P Landau and Kurt Binder. *A Guide to Monte Carlo Simulations in Statistical Physics.* 4th ed. Cambridge University Press, 2014.

[72] David J Wales and Harold A Scheraga. “Global Optimization of Clusters, Crystals, and Biomolecules”. In: *Science (80-.).* 285 (1999), pp. 1368–1372.

[73] Andrea C Levi and Miroslav Kotrla. “Theory and simulation of crystal growth”. In: *J. Phys. Condens. Matter* 9 (1997), p. 299.

[74] C Ratsch and J A Venables. “Nucleation Theory and the Early Stages of Thin Film Growth”. In: *J. Vac. Sci. Technol. A* 21 (2003), S96.

[75] Wlater Kob. “Computer simulations of supercooled liquids and glasses”. In: *J. Phys. Condens. Matter* 11 (1999), R85.

[76] Pablo Jensen. “Growth of nanostructures by cluster deposition: Experiments and simple models”. In: *Rev. Mod. Phys.* 71.5 (1999), pp. 1695–1735.

[77] Daan Frenkel and Berend Smit. *Understanding Molecular Simulation: from Algorithms to Applications.* 2nd ed. Academic Press, 2002.

[78] M P Allen and D J Tildesley. *Computer simulation of liquids.* 2nd ed. Oxford Science Publications, 2017.

[79] Steve Brooks et al. *Handbook of Markov Chain Monte Carlo.* CRC Press, 2011.

[80] N Metropolis et al. “Equation of State Calculations by Fast Computing Machines”. In: *J. Chem. Phys.* 21.6 (1953), pp. 1087–1092.

[81] Vasilios I Manousiouthakis and Michael W Deem. “Strict detailed balance is unnecessary in Monte Carlo simulation”. In: *J. Chem. Phys.* 110 (1999), p. 2753.

[82] Hidemaro Suwa and Synge Todo. “Markov Chain Monte Carlo Method without Detailed Balance”. In: *Phys. Rev. Lett.* 105 (2010), p. 120603.

[83] Manon Michel, Sebastian C Kapfer, and Werner Krauth. “Generalized event-chain Monte Carlo:” in: *J. Chem. Phys.* 140 (2014), p. 054116.

[84] G M Torrie and J P Valleau. “Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling”. In: *J. Comput. Phys.* 23 (1977), pp. 187–199.

[85] David J Earl and Michael W Deem. “Parallel tempering: Theory, applications, and new perspectives”. In: *Phys. Chem. Chem. Phys.* 7 (2005), pp. 3910–3916.

[86] Bernd Hartke. “Global Geometry Optimization of Clusters Using Genetic Algorithms”. In: *J. Phys. Chem.* 97 (1993), pp. 9973–9976.

[87] J A Niesse and Howard R Mayne. “Global geometry optimization of atomic clusters using a modified genetic algorithm in space-fixed coordinates”. In: *J. Chem. Phys.* 105 (1996), p. 4700.

[88] David J Wales and Jonathan P K Doye. “Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms”. In: *J. Phys. Chem. A* 101 (1997), pp. 5111–5116.

[89] S . Kirkpatrick, C . D . Gelatt Jr., and M . P . Vecchi. “Optimization by Simulated Annealing”. In: *Science (80-.).* 220.4598 (1983), pp. 671–680.

[90] Darrall Henderson, Sheldon H Jacobson, and Alan W Johnson. “The Theory and Practice of Simulated Annealing”. In: *Handb. Metaheuristics*. Ed. by Fred Glover and Gary A Kochenberger. Boston, MA: Springer US, 2003, pp. 287–319.

[91] F Wooten, K Winer, and D Weaire. “Computer Generation of Structural Models of Amorphous Si and Ge”. In: *Phys. Rev. Lett.* 54.13 (1985), pp. 1392–1395.

[92] M M J Treacy and K B Borisenko. “The Local Structure of Amorphous Silicon”. In: *Science (80-.).* 335 (2012), pp. 950–953.

[93] Yuhai Tu et al. “Properties of a Continuous-Random-Network Model for Amorphous Systems”. In: *Phys. Rev. Lett.* 81.22 (1998), pp. 4899–4902.

[94] B R Djordjevic, M F Thorpe, and F Wooten. “Computer model of tetrahedral amorphous diamond”. In: *Phys. Rev. B* 52.8 (1995), pp. 5685–5690.

[95] Normand Mousseau and G T Barkema. “Binary continuous random networks”. In: *J. Phys. Condens. Matter* 16 (2004), S5183–S5190. arXiv: `0408705 [cond-mat]`.

[96] E M Huisman, C Storm, and G T Barkema. “Monte Carlo study of multiply crosslinked semiflexible polymer networks”. In: *Phys. Rev. E* 78 (2008), p. 051801.

[97] C P Broedersz and F C Mackintosh. “Modeling semiflexible polymer networks”. In: *Rev. Mod. Phys.* 86 (2014), pp. 995–1036.

[98] Sandeep K Jain and Gerard T Barkema. “Rupture of amorphous graphene via void formation”. In: *PCCP* 20 (2018), pp. 16966–16972.

[99] Avishek Kumar, Mark Wilson, and M F Thorpe. “Amorphous graphene: a realization of Zachariasen’s glass”. In: *J. Phys. Condens. Matter* 24 (2012), p. 485003.

[100] P. N. Keating. “Effect of invariance requirements on the elastic strain energy of crystals with application to the diamond structure”. In: *Phys. Rev.* 145.2 (1966), pp. 637–645.

[101] G. Barkema and Normand Mousseau. “High-quality continuous random networks”. In: *Phys. Rev. B* 62.8 (2000), pp. 4985–4990.

[102] D A Drabold. “Topics in the theory of amorphous materials”. In: *Eur Phys J B* 68 (2009), pp. 1–21.

[103] S. von Alfthan, A. Kuronen, and K. Kaski. “Realistic models of amorphous silica: A comparative study of different potentials”. In: *Phys. Rev. B* 68 (2003), p. 073203.

[104] Monica Bulacu et al. “Improved Angle Potentials for Coarse-Grained Molecular Dynamics Simulations”. In: *J. Chem. Theory Comput.* 9 (2013), pp. 3282–3292.

[105] Jorge Nocedal and Stephen J Wright. *Numerical Optimization*. 2nd ed. Springer, 2006.

[106] Normand Mousseau and G. T. Barkema. “Fast bond-transposition algorithms for generating covalent amorphous structures”. In: *Curr. Opin. Solid State Mater. Sci.* 5 (2001), pp. 497–502.

[107] Masaharu Isobe. “Hard sphere simulation in statistical physics - methodologies and applications”. In: *Mol. Simul.* 42.16 (2016), pp. 1317–1329.

[108] Etienne P Bernard, Werner Krauth, and David B Wilson. “Event-chain Monte Carlo algorithms for hard-sphere systems”. In: *Phys. Rev. E* 80 (2009), p. 056704.

[109] Joshua A Anderson et al. “Massively parallel Monte Carlo for many-particle simulations on GPUs”. In: *J. Comput. Phys.* 254 (2013), pp. 27–38.

[110] Masaharu Isobe and Werner Krauth. “Hard-sphere melting and crystallization with event-chain Monte Carlo”. In: *J. Chem. Phys.* 143 (2015), p. 084509.

[111] B Widom. “Random Sequential Addition of Hard Spheres to a Volume”. In: *J Chem Phys* 44 (1966), p. 3888.

[112] T S Grigera and G Parisi. “Fast Monte Carlo algorithm for supercooled soft spheres”. In: *Phys. Rev. E* 63 (2001), 045102(R).

[113] Andrea Ninarello, Ludovic Berthier, and Daniele Coslovich. “Models and Algorithms for the Next Generation of Glass Transition Studies”. In: *Phys. Rev. X* 7 (2017), p. 021039.

[114] A Okabe, B Boots, and K Sugihara. *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams.* Wiley, 1992.

[115] Anne Poupon. “Voronoi and Voronoi-related tessellations in studies of protein structure and interaction”. In: *Curr. Opin. Struct. Biol.* 14 (2004), pp. 233–241.

[116] B. J. Gellatly and J. L. Finney. “Characterisation of Models of Multicomponent Amorphous Metals: the Radical Alternative to the Voronoi Polyhedron”. In: *J. Non. Cryst. Solids* 50 (1982), pp. 313–329.

[117] F Aurenhammer. “Power diagrams: properties, algorithms and applications”. In: *SIAM J Comput* 16.1 (1987), pp. 78–96.

[118] FM Richards. “The Interpretation of Protein Structures : Total Volume, Group Volume Distributions and Packing Density”. In: *J Mol Biol* 82 (1974), pp. 1–14.

[119] B N Boots. “The Spatial Arrangement of Random Voronoi Polygons”. In: *Comput. Geosci.* 9.3 (1983), pp. 351–365.

[120] Masaharu Tanemura. “Statistical Distributions of Poisson Voronoi Cells in Two and Three Dimensions”. In: *Forma* 18 (2003), pp. 221–247.

[121] Chris H Rycroft. “VORO++ : A three-dimensional Voronoi cell library in C ++”. In: *Chaos* 19 (2009), p. 041111.

[122] D. Löffler et al. “Growth and structure of crystalline silica sheet on Ru(0001)”. In: *Phys. Rev. Lett.* 105 (2010), p. 146104.

[123] Leonid Lichtenstein, Markus Heyde, and Hans Joachim Freund. “Crystalline-vitreous interface in two dimensional silica”. In: *Phys. Rev. Lett.* 109 (2012), p. 106101.

[124] Mahdi Sadjadi et al. “Refining glass structure in two dimensions”. In: *Phys. Rev. B* 96 (2017), 201405(R).

[125] James F. Shackelford. “Triangle rafts - extended Zachariasen schematics for structure modeling”. In: *J. Non. Cryst. Solids* 49 (1982), pp. 19–28.

[126] Christin Büchner et al. “Building block analysis of 2D amorphous networks reveals medium range correlation”. In: *J. Non. Cryst. Solids* 435 (2016), pp. 40–47.

[127] Louis Theran et al. "Anchored boundary conditions for locally isostatic networks". In: *Phys. Rev. E* 92 (2015), p. 053306.

[128] P. Tangney and S. Scandolo. "An ab initio parametrized interatomic force field for silica". In: *J. Chem. Phys.* 117 (2002), pp. 8898–8904.

[129] I. Zsoldos and A. Szasz. "Appearance of collectivity in two-dimensional cellular structures". In: *Comput. Mater. Sci.* 15 (1999), pp. 441–448.

[130] Mahdi Sadjadi and M. F. Thorpe. "Ring correlations in random networks". In: *Phys. Rev. E* 94 (2016), p. 062304.

[131] S R Broadbent and J M Hammersley. "Percolation processes I. Crystals and Mazes". In: *Proc. Camb. Phil. Soc.Phil. Soc.* 53 (1956), pp. 629–641.

[132] Duncan S Callaway et al. "Network Robustness and Fragility : Percolation on Random Graphs". In: *Phys. Rev. Lett.* 85.25 (2000), p. 5468.

[133] Dietrich Stauffer and Amnon Aharony. *Introduction to percolation theory.* 2nd ed. 2014.

[134] M F Sykes and J W Essam. "Exact Critical Percolation Probabilities for Site and Bond Problems in Two Dimensions". In: *J. Math. Phys.* 5.8 (1964), p. 1117.

[135] Scott Kirkpatrick. "Percolation and Conduction". In: *Rev. Mod. Phys.* 45.4 (1973), p. 574.

[136] H L Frisch, J M Hammersley, and D J A Welsh. "Monte Carlo Estimates of Percolation Probabilities for Various Lattices". In: *Phys. Rev.* 126.3 (1962), pp. 949–951.

[137] B Y P Dean and N F Bird. "Monte Carlo estimates of critical percolation probabilities". In: *Proc. Camb. Phil. Soc.* 63 (1967), p. 477.

[138] M E J Newman. "Mixing patterns in networks". In: *Phys. Rev. E* 67 (2003), p. 026126.

[139] Di Zhou et al. "Assortativity decreases the robustness of interdependent networks". In: *Phys. Rev. E* 86 (2012), p. 066103.

[140] C Schmeltzer et al. "Percolation of spatially constrained Erdos-Renyi networks with degree correlations". In: *Phys. Rev. E* 89 (2014), p. 012116.

[141] V Kapko, D A Drabold, and M F Thorpe. "Electronic structure of a realistic model of amorphous graphene". In: *Phys. Status Solidi* 247.5 (2010), pp. 1197–1200.

[142] Taishan Zhu and Elif Ertekin. "Phonons, Localization, and Thermal Conductivity of Diamond Nanothreads and Amorphous Graphene". In: *Nano Lett.* 16 (2016), pp. 4763–4772.

[143] Franz Bamer, Firaz Ebrahem, and Bernd Markert. "Elementary plastic events in a Zachariasen glass under shear and pressure". In: *Materialia* 9 (2020), p. 100556.

[144] Firaz Ebrahem, Franz Bamer, and Bernd Markert. "Vitreous 2D silica under tension : From brittle to ductile behaviour". In: *Mater. Sci. Eng. A* 780 (2020), p. 139189.

[145] Oliver Whitaker. "Modelling of complex ring networks in two- and three-dimensions". PhD thesis. University of Oxford, 2019.

[146] Xiaolei Ma, Janna Lowensohn, and Justin C Burton. "Universal scaling of polygonal desiccation crack patterns". In: *Phys. Rev. E* 99 (2019), p. 012802.

[147] Hisao Honda. "Description of cellular patterns by Dirichlet domains: The two-dimensional case". In: *J. Theor. Biol.* 72 (1978), pp. 523–543.

[148] Ross Carter et al. "Pavement cells and the topology puzzle". In: *Development* 144 (2017), pp. 4386–4397.

[149] Sangwoo Kim et al. "Hexagonal Patterning of the Insect Compound Eye : Facet Area Variatio , Defects, and Disorder". In: *Biophys. J.* 111 (2016), pp. 2735–2746.

[150] C. J. Lambert and D. L. Weaire. "Theory of the arrangement of cells in a network". In: *Metallography* 14.4 (1981), pp. 307–318.

[151] Susmit Kumar, Stewart K. Kurtz, and Denis Weaire. "Average number of sides for the neighbours in a Poisson-Voronoi tesselation". In: *Philos. Mag. B* 69.3 (1994), pp. 431–435.

[152] M. Blanc and A. Mocellin. "Grain coordination in plane sections of polycrystals". In: *Acta Metall.* 27 (1979), pp. 1231–1237.

[153] N. Rivier. "Statistical crystallography structure of random cellular networks". In: *Philos. Mag. B* 52.3 (1985), pp. 795–819.

[154] Michael A. Peshkin, Katherine J. Strandburg, and Nicolas Rivier. "Entropic predictions for cellular networks". In: *Phys. Rev. Lett.* 67.13 (1991), pp. 1803–1806.

[155] S N Chiu. "Mean-Value Formulae for the Neighbourhood of the Typical Cell of a Random Tessellation". In: *Adv. Appl. Probab.* 26 (1994), pp. 565–576.

[156] J K Mason, R Ehrenborg, and E A Lazar. "A geometric formulation of the law of Aboav – Weaire in two and three dimensions". In: *J. Phys. A* 45 (2012), p. 065001.

[157] H. J. Hilhorst. "Planar Voronoi cells: The violation of Aboav's law explained". In: *J. Phys. A* 39 (2006), pp. 7227–7243.

[158] M. E.J. Newman. "Assortative Mixing in Networks". In: *Phys. Rev. Lett.* 89.20 (2002), pp. 1–4.

[159] Rogier Noldus and Piet Van Mieghem. "Assortativity in complex networks". In: *J. Complex Networks* 3 (2015), pp. 507–542.

[160] Alexandros Chremos and Philip J. Camp. "Neighbor network in a polydisperse hard-disk fluid: Degree distribution and assortativity". In: *Phys. Rev. E* 76 (2007), p. 056108.

[161] Nelly Litvak and Remco van der Hofstad. "Uncovering disassortativity in large scale-free networks". In: *Phys. Rev. E* 87 (2013), p. 022801.

[162] J M Greneche and J M D Coey. "The topologically-disordered square lattice". In: *J. Phys. Fr.* 51 (1990), pp. 231–242.

[163] Franz R Eder et al. "A journey from order to disorder - atom by atom transformation from graphene to a 2D carbon glass". In: *Sci. Rep.* 4 (2014), p. 4060.

[164] Ordnance Survey. *Boundary-Line Data © Crown copyright and database right 2018*. 2018.

[165] Federal Office of Topography. *swissBOUNDARIES3D*. 2019.

[166] Eurostat. *NUTS Geodata © EuroGeographics for the administrative boundaries.* 2016.

[167] Etienne P Bernard and Werner Krauth. “Two-Step Melting in Two Dimensions : First-Order Liquid-Hexatic Transition”. In: *Phys. Rev. Lett.* 107 (2011), p. 155704.

[168] Andrew H Marcus and Stuart A Rice. “Phase transitions in a confined quasi-two-dimensional colloid suspension”. In: *Phys. Rev. E* 55.1 (1997), p. 637.

[169] Han-Rui Tian et al. “An Unconventional Hydrofullerene C66H4 with Symmetric Heptagons Retrieved in Low-Pressure Combustion”. In: *J. Am. Chem. Soc.* 141 (2019), pp. 6651–6657.

[170] Runnan Guan et al. “Stable C92(26) and C92(38) as Well as Unstable C92(50) and C92(23) Isolated-Pentagon-Rule Isomers As Revealed by Chlorination of C92 Fullerene”. In: *Inorg. Chem.* 58 (2019), pp. 5393–5396.

[171] Victor A Brotsman, Daria V Ignateva, and Sergey I Troyanov. “Chlorination-promoted Transformation of Isolated Pentagon Rule C78 into Fused-pentagons- and Heptagons-containing Fullerenes”. In: *Chem Asian J* 12 (2017), pp. 2379–2382.

[172] Victor A Brotsman et al. “Rebuilding C60 : Chlorination-Promoted Transformations of the Buckminsterfullerene into Pentagon-Fused C60 Derivatives”. In: *Inorg. Chem.* 57 (2018), pp. 8325–8331.

[173] Daishi Fujita et al. “Self-assembly of tetravalent Goldberg polyhedra from 144 small components”. In: *Nature* 540 (2016), pp. 563–566.

[174] Zhi Wang et al. “Assembly of silver Trigons into a buckyball-like Ag180 nanocage”. In: *PNAS* 114.46 (2017), pp. 12132–12137.

[175] A E Roth, C D Jones, and D J Durian. “Coarsening of a two-dimensional foam on a dome”. In: *Phys. Rev. E* 86 (2012), p. 021402.

[176] P N Pusey and W Van Megen. “Phase behaviour of concentrated suspensions of nearly hard colloidal spheres”. In: *Nature* 320 (1986), p. 340.

[177] Alice L Thorneywork et al. “Radial distribution functions in a two-dimensional binary colloidal hard sphere system”. In: *J. Chem. Phys.* 140 (2014), p. 161106.

[178] Alice L Thorneywork et al. “Structure factors in a two-dimensional binary colloidal hard sphere system”. In: *Mol. Phys.* 116 (2018), pp. 3245–3257.

[179] R Roth, R Evans, and A A Louis. “Theory of asymmetric nonadditive binary hard-sphere mixtures”. In: *Phys. Rev. E* 64 (2001), p. 051202.

[180] R Y Yang, R P Zou, and A B Yu. “Voronoi tessellation of the packing of fine uniform spheres”. In: *Phys. Rev. E* 65 (2002), 041302 Voronoi.

[181] VS Kumar and V Kumaran. “Voronoi neighbor statistics of hard-disks and hard-spheres”. In: *J Chem Phys* 123 (2005), p. 074502.

[182] A Jaster. “Computer simulations of the two-dimensional melting transition using hard disks”. In: *Phys. Rev. E* 59.3 (1999), pp. 2594–2602.

[183] Sander Pronk and Daan Frenkel. “Melting of polydisperse hard disks”. In: *Phys. Rev. E* 69 (2004), p. 066123.

[184] Sebastian C Kapfer and Werner Krauth. "Two-Dimensional Melting : From Liquid-Hexatic Coexistence to Continuous Transitions". In: *Phys. Rev. Lett.* 114 (2015), p. 035702.

[185] Alice L Thorneywork. "Structure and dynamics of two-dimensional colloidal hard spheres". PhD thesis. University of Oxford, 2015.

[186] Q Weikai, A P Gantapara, and Marjolein Dijkstra. "Two-stage melting induced by dislocations and grain boundaries in monolayers of hard spheres". In: *Soft Matter* 10.30 (2014), p. 5449.

[187] Y Peng et al. "Short-time self-diffusion of nearly hard spheres at an oil–water interface". In: *J Fluid Mech* 618 (2009), pp. 243–261.

[188] N Vogel et al. "Direct visualization of the interfacial position of colloidal particles and their assemblies†". In: *Nanoscale* 6 (2014), pp. 6879–6885.

[189] Elisa Tamborini, C Patrick Royall, and Pietro Cicuta. "Correlation between crystalline order and vitrification in colloidal monolayers". In: *J. Phys. Condens. Matter* 27 (2015), p. 194124.

[190] H Imai, I Masao, and K Murota. "Voronoi Diagram in the Laguerre Geometry and its Applications". In: *SIAM J* 14.1 (1985), pp. 93–105.

[191] N Rivier. "Geometry and Fluctuations of Surfaces". In: *J. Phys. Colloq.* 51 (1990), pp. 309–317.

[192] P Bhattacharyya and B K Chakrabarti. "The mean distance to the nth neighbour in a uniform distribution of random points: an application of probability theory". In: *Eur. J. Phys.* 29 (2008), p. 639.

[193] L Oger et al. "Comparison of two representations of a random cut of identical sphere packing". In: *Eur. Phys. J. B* 14 (2000), pp. 403–406.

[194] Annie Gervois, Luc Oger, and Jean-Paul Troadec. "Random cuts in binary mixtures of spheres". In: *Phys. Rev. E* 70 (2004), p. 031112.

[195] U Hahn and U Lorz. "Stereological analysis of the spatial Poisson-Voronoi tessellation". In: *J. Microsc.* 175.3 (1994), pp. 176–185.

[196] Simone Falco et al. "Generation of 3D polycrystalline microstructures with a conditioned Laguerre-Voronoi tessellation technique". In: *Comput. Mater. Sci.* 136 (2017), pp. 20–28.

[197] Dorian Depriester and Régis Kubler. "Radical Voronoï tessellation from random pack of polydisperse spheres Prediction of the cells' size distribution". In: *Comput. Des.* 107 (2019), pp. 37–49.

[198] Matthew O Blunt et al. "Random Tiling and Topological Defects in a Two-Dimensional Molecular Network". In: *Science (80-.).* 322 (2008), pp. 1077–1081.

[199] P W Anderson. "RESONATING VALENCE BONDS" A NEW KIND OF INSULATOR?" In: *Mat. Res. Bull.* 8 (1973), pp. 153–160.

[200] Philip J Camp, Amparo Fuertes, and J P Attfield. "Subextensive Entropies and Open Order in Perovskite Oxynitrides". In: *J Am Chem Soc* 134 (2012), pp. 6762–6766.

[201] R Comes, M Lambert, and A Guinier. "THE CHAIN STHUCTURE OF BaT1O3 AND KNbO3". In: *Solid State Commun.* 6 (1968), pp. 715–719.

[202] G Algara-Siller et al. "Square ice in graphene nanocapillaries". In: *Nature* 519 (2015), p. 443.

[203] YinBo Zhu, FengChao Wang, and HengAn Wu. "Structural and dynamic characteristics in monolayer square ice". In: *J. Chem. Phys.* 147 (2017), p. 044706.

[204] Simon J Hibble et al. "Structures of Pd (CN)2 and Pt (CN)2: Intrinsically Nanocrystalline Materials?" In: *Inorg. Chem.* 50 (2011), pp. 104–113.

[205] Qing-Na Zheng et al. "Adaptive Reorganization of 2D Molecular Nanoporous Network Induced by Coadsorbed Guest Molecule". In: *Langmuir* 30 (2014), pp. 3034–3040.

[206] Lars Postulka et al. "Spin Frustration in an Organic Radical Ion Salt Based on a Kagome-Coupled Chain Structure". In: *J. Am. Chem. Soc.* 138 (2016), pp. 10738–10741.

[207] Ting Chen et al. "2D Hexagonal Tilings Based on Triangular and Hexagonal Structural Units in the Self-Assembly of Thiacalix[4]arene Tetrasulfonate on an Au(111) Surface". In: *Chem Asian J* 6 (2011), pp. 1811–1816.

[208] Austin D Griffith and Robert S Hoy. "Densest versus jammed packings of two-dimensional bent-core trimers". In: *Phys. Rev. E* 98 (2018), p. 042910.

[209] Jose I Urgel et al. "Five-Vertex Lanthanide Coordination on Surfaces: A Route to Sophisticated Nanoarchitectures and Tessellations". In: *J Phys Chem C* 118 (2014), pp. 12908–12915.

[210] N P Kryuchkov et al. "Complex crystalline structures in a two-dimensional core-softened system". In: *Soft Matter* 14 (2018), p. 2152.

[211] Bai-Qiao Song et al. "Periodic tiling of triangular and square nanotubes in a cationic metal – organic framework for selective anion exchange †". In: *Chem Commun* 51 (2015), pp. 9515–9518.

[212] William D Piñeros, Michael Baldea, and Thomas M Truskett. "Designing convex repulsive pair potentials that favor assembly of kagome and snub square lattices". In: *J. Chem. Phys.* 145 (2016), p. 054901.

[213] Mia Baise et al. "Negative Hydration Expansion in ZrW2O8 : Microscopic Mechanism, Spaghetti Dynamics, and Negative Thermal Expansion". In: *Phys. Rev. Lett.* 120 (2018), p. 265501.

[214] V A Gorbunov, S S Akimenko, and A V Myshlyavtsev. "Cross-impact of surface and interaction anisotropy in the self-assembly of organic adsorption monolayers : a Monte Carlo". In: *Phys Chem Chem Phys* 19 (2017), pp. 17111–17120.

[215] Damien Nieckarz, Wojciech Rzysko, and Paweł Szabelski. "On-surface self-assembly of tetratopic molecular building blocks". In: *Phys Chem Chem Phys* 20 (2018), p. 23363.

[216] C Buzano et al. "Two-dimensional lattice-fluid model with waterlike anomalies". In: *Phys. Rev. E* 69 (2004), p. 061502.

[217] Philip J Camp. “Structure and phase behavior of a two-dimensional system with core-softened and long-range repulsive interactions”. In: *Phys. Rev. E* 68 (2003), p. 061506.

[218] Michael Griebel and Jan Hamaekers. “Molecular dynamics simulations of boron-nitride nanotubes embedded in amorphous Si-B-N”. In: *Comput. Mater. Sci.* 39 (2007), pp. 502–517.

[219] Arkadiy Simonov et al. “Hidden diversity of vacancy networks in Prussian blue analogues”. In: *Nature* 578 (2020), p. 256.

[220] Marian Florescu, Salvatore Torquato, and Paul J Steinhardt. “Designer disordered materials with large, complete photonic band gaps”. In: *PNAS* 106.49 (2009), pp. 20658–20663.

[221] Steven R Sellers et al. “Local self-uniformity in photonic networks”. In: *Nat. Commun.* 8 (2017), p. 14439.