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 20th century, much work was done 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.
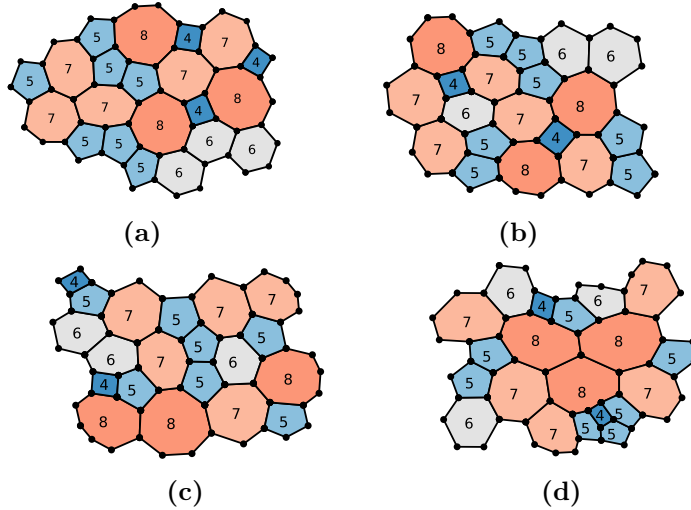
Figure 1.1: Panel (a) shows Zachariasen's glass and panels (b)-(d) three different periodic arrangements based on the glass (with one square removed to satisfy Euler's formula). Moving from panel (b)-(d) there is increased clustering of similar sized rings. The size of the rings are highlighted numerically 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 [32–36]. There are however countless more occurrences, with drying blood, stratocumulus clouds, crocodile scales and geopolitical borders all being the subject of studies [37–40]. More intriguingly, although these systems are incredibly physically diverse, they still have similar structures [41]. 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 [42, 43]; as well as to curiosities such as explaining the arrangement of the stones in Giant's Causeway or cracking in famous artworks [44, 45].

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 [46–48]. 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 [49]. The second are the ring structures in so-called "procrystalline" lattices [50]. Paragraph to wrap up

1.2 Thesis Structure

Add thesis structure to introduction Test: Mark can add notes with the command \marknote

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, Lemâitre's law and the Aboav-Weaire law are also explored. Further developments linking to modern network science etc. are given in chapter Link to network theory later .

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 [51]. 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.
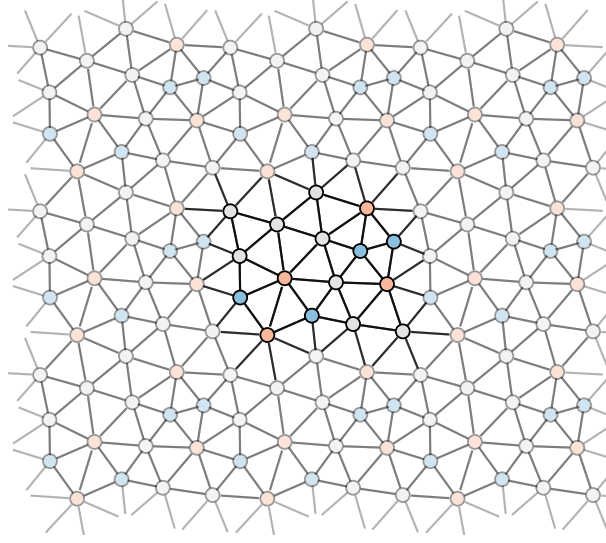
Figure 2.1: Example of a periodic two-dimensional network where nodes are represented by circles and links as lines. Nodes are coloured similarly according to their degree, whilst periodic images are greyed out to highlight the central repeating unit.

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 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 nth 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:

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

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.

(a) Atomic network.   **(b)** Ring network.   **(c)** Dual relationship.
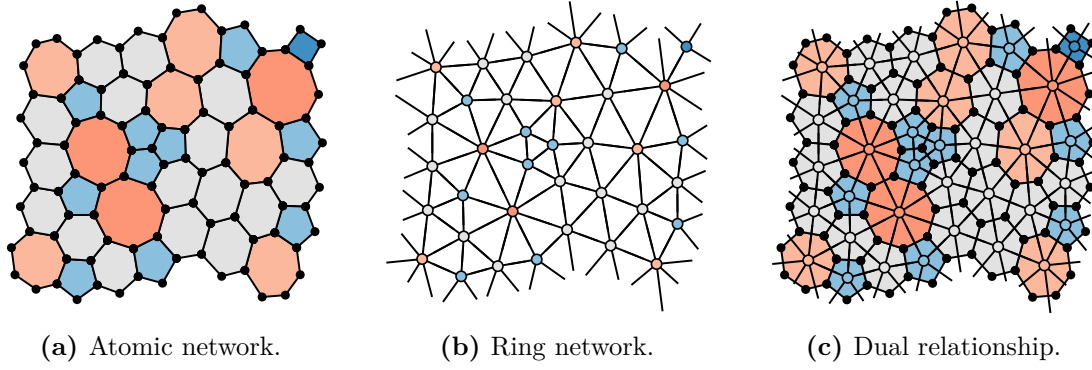
Figure 2.2: Panel (a) gives an example of a 3- coordinate periodic atomic network with disordered ring structure. Nodes and links represent atoms and bonds respectively where rings are coloured by size. Panel (b) gives the corresponding ring network where nodes and links represent rings and their adjacencies, 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.

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 [52]. A 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 [53]. 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, Lemâitre'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 $\langle c \rangle = \sum_c c x_c$. 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 [36, 54].

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) [55]. 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.
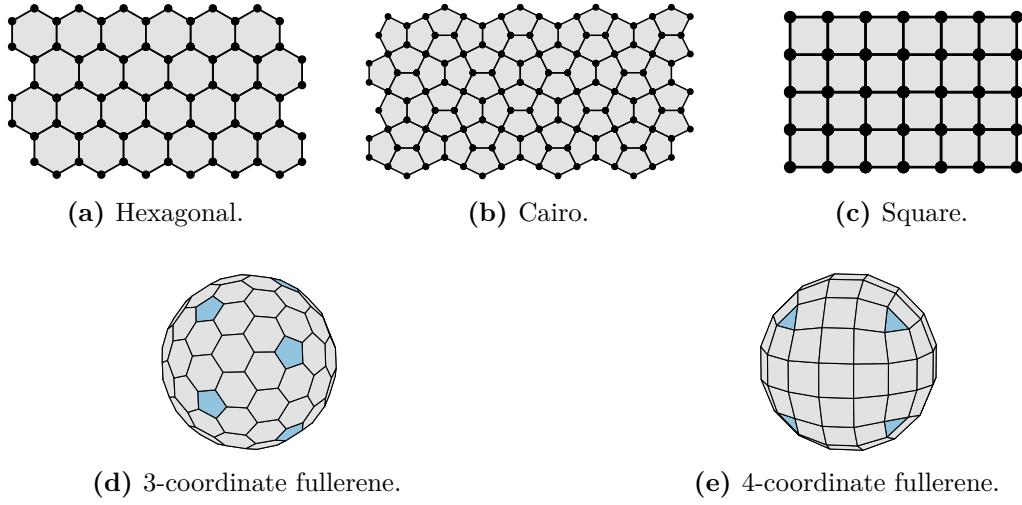
Figure 2.3: Panels (a)-(c) give regular planar tilings of 6-, 5- and 4- rings, where the ring size is related to the underlying atomic coordination. Panels (d) and (e) show the 3- and 4- coordinate tilings in spherical topology, where the mean ring size is reduced due to the change in the Euler characteristic.

2.2.2 Lemâitre'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]. Lemâitre *et al.* demonstrated that the distribution in 3-coordinate networks systems can be well described by a maximum entropy distribution [56]. Lemâitre's maximum entropy method is summarised here, trivially extended to arbitrary coordination.

The entropy of a probability distribution is defined as

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

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 Lemâitre'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 [57–59]. 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 [60]. Link to procrystal chapter . The resulting Lemâitre curves for a selection of these modifications are given in figure 2.4c. explain these here or later? 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 for many observable networks [61–63]. Despite this, the universality of the Lemâitre law suggests that there must be a physical basis to (2.16), and

(a) Maximum entropy distributions.

(b) Lemâitre's law.

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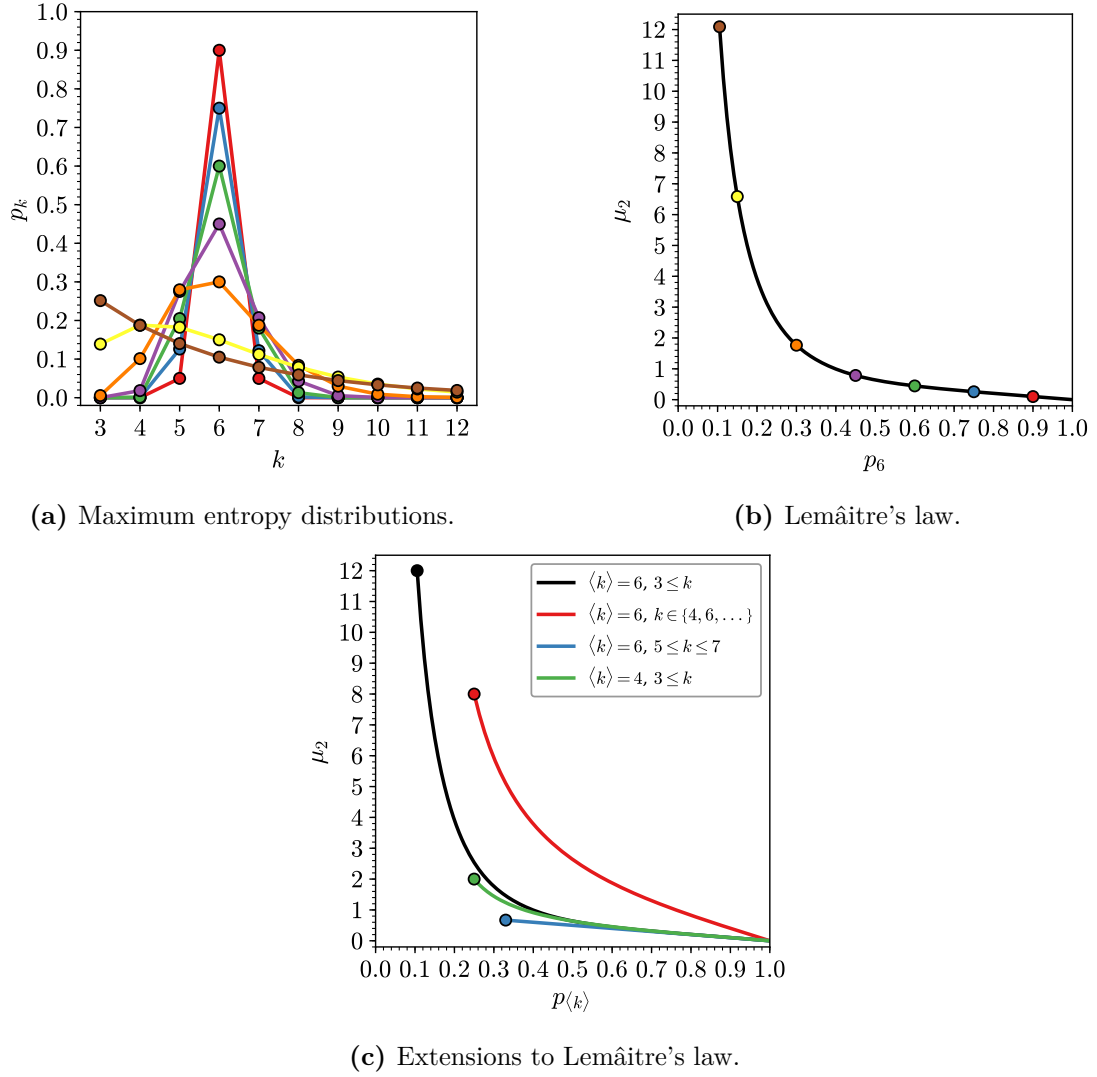
Figure 2.4: Illustration of Lemâitre's maximum entropy method. Panel (a) gives examples of explicit maximum entropy distributions with different values of p_6. Panel (b) shows how these distributions can be summarised in a plot of p_6 *vs.* μ_2 (Lemâitre's law). Panel (c) provides extensions to the law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

in the section   Link to later networks   it will be demonstrated that it can be regenerated by considering ring adjacencies.

2.2.3   Aboav-Weaire Law

The ring statistics given by Lemâitre'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 3.3.3). 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 [64]:

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

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 [65]:

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

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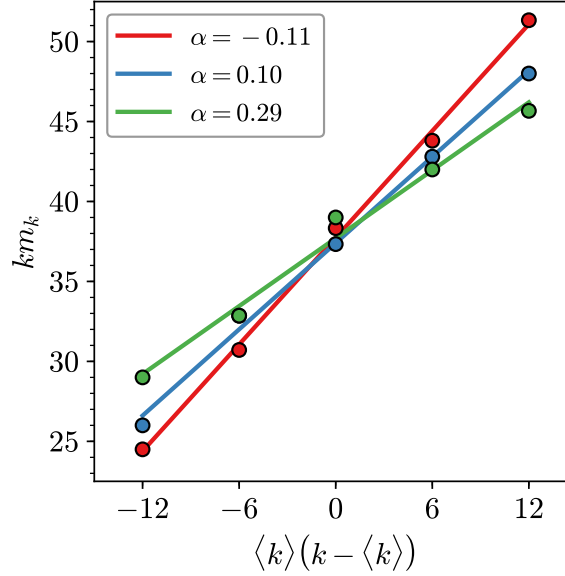
Figure 2.5: Calculation of an Aboav-Weaire fit for three configurations (shown in figure 1.1(b)-(d)). The value of the α parameter quantifies the tendency of small rings to be adjacent to large rings, with a larger value indicating stronger small-large ring correlations.

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, 66–69]. 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.

2.3 Persistent Homology?

2.4 Percolation Theory?

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. Extensions to these methods and additional approaches are outlined in the relevant chapters Link to bond switching/Voronoi/mx2/procrystals later .

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-20th 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 [70–75]. In this thesis these Monte Carlo methods will be used in a variety of contexts chapter xxx fill this in . 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,

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

where $\mathcal{K}\left(\mathbf{p}\right)$ is the kinetic energy as a function of all particle momenta and $\mathcal{U}\left(\mathbf{r}\right)$ is the potential energy as a function of all particle positions [76]. 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:

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

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

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

where

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

is the configurational integral [77]. 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:

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

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

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

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.

(a) Quadrature.

(b) Uniform Sampling.

(c) Importance Sampling.

(d) Markov Chain Monte Carlo
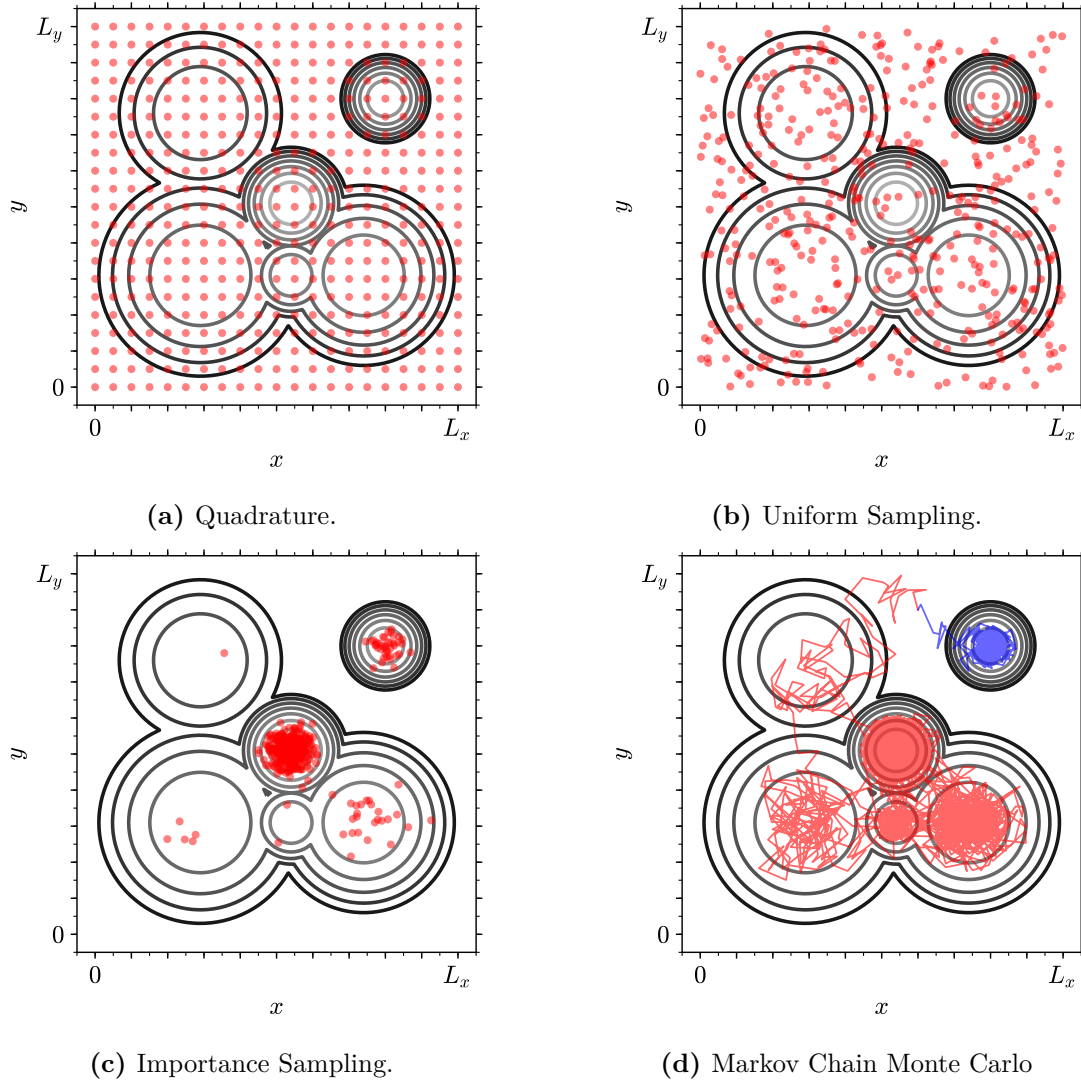
Figure 3.1: Demonstration of different sampling methods with an example two-dimensional potential energy surface (contour lines). Panels(a)-(c) display the same number of (red) sampling points. Panel (a) shows conventional quadrature where the surface is divided into a regular grid of sampling points 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.

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 [78]. 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 [76–78].

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 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 [79]. 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 [80–82], 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:

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

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

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

which can be expressed more succinctly as

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

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 $\mathcal{U}(X_s)$

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 $\mathcal{U}(X_t)$

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 [83, 84].

### 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 [85–87]. This thesis will use simulated annealing, which can be considered an extension to Metropolis Monte Carlo [88]. 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.

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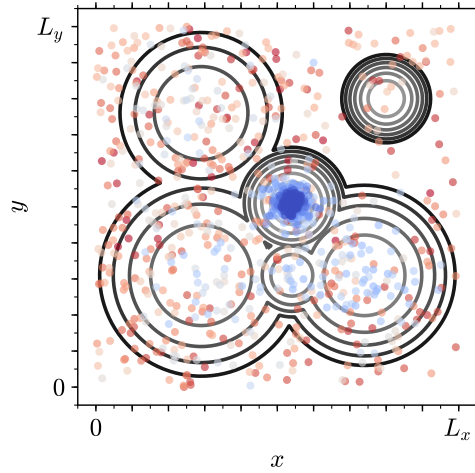
Figure 3.2: Demonstration of the simulated annealing algorithm on a two-dimensional potential energy surface, with states coloured by temperature (red→blue indicating hot→cold. As the temperature is reduced the state converges on the global minimum.

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 [90]. 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
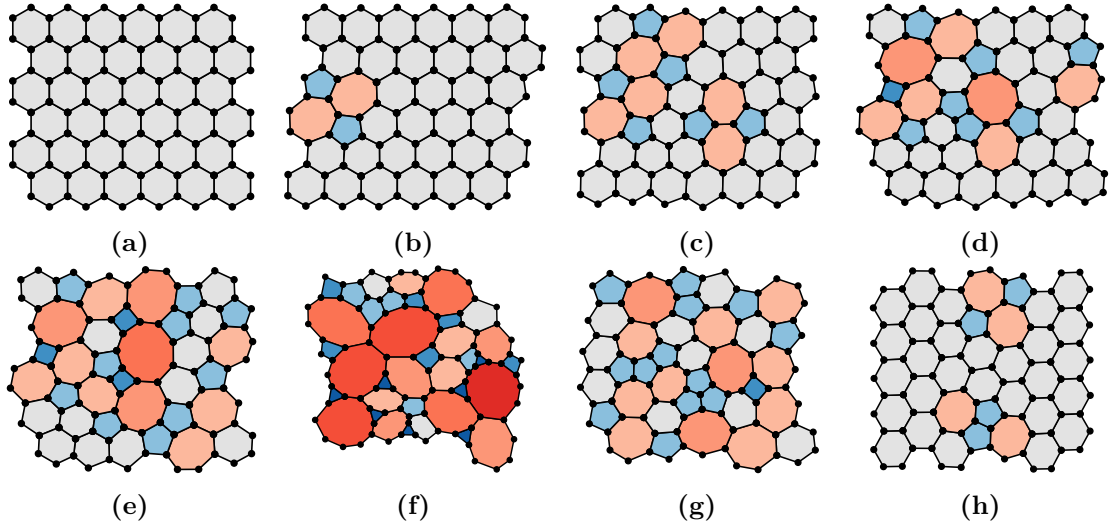
Figure 3.3: Configurations taken from stages of the two-dimensional bond switching algorithm. A crystalline lattice (a) is first thermalised to generate a random high energy network (f) by sequential overlapping Stone-Wales defects (b)-(e). Sampling then occurs as the system is slowly annealed (g)-(h), allowing access to defect states that are not initially obtainable from the crystal structure.

alternative systems such as three-dimensional amorphous carbon, binary glasses and biological polymers [91–96]. 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 ref to later chapters [29, 97]. 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 [98]. 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 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 [99, 100]. For a two-dimensional system the Keating potential has the form:

$$\mathcal{U}\left(\mathbf{r}\right) = \frac{3}{16}\frac{k_S}{r_0^2} \sum_{\substack{i,j \in \\ \text{stretches}}} \left(r_{ij}^2 - r_0^2\right)^2 + \frac{3}{8}\frac{k_A}{r_0^2} \sum_{\substack{ijk \in \\ \text{angles}}} \left(r_{ij}r_{ik}\cos\theta_{i,j,k} + d_0^2\cos\theta_0\right)^2 , \quad (3.21)$$

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 [98]. 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 [98, 101].

However, a more generic potential model is sometimes required which captures the same essential physics. This is provided through the simplified Keating potential [102],

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

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 [103]:

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

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, $\mathcal{U}(\mathbf{r})$, 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 [104]. It has the following scheme:

1. Calculate the potential energy of the system $\mathcal{U}_i = \mathcal{U}(\mathbf{r}_i)$

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 Add appendix for force calculations...eugh . 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 [105].

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 [106–109]. In this thesis it will be used to generate ring systems in the form of Voronoi tessellations (see section 3.3.3), in analogy to experimental colloidal systems [49].

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 centroids by r_{ij}, the pair potential is:

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

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 [110]. For higher packing fractions a more sophisticated algorithm is needed Find refs .

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
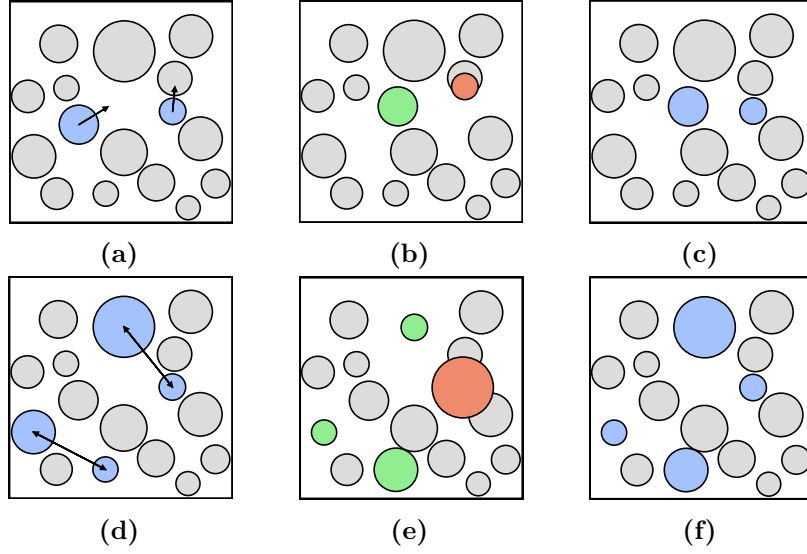
Figure 3.4: Demonstration of two displacement (a)-(c) and two swap (d)-(f) moves in hard particle Monte Carlo. In displacement moves, particles are randomly selected and assigned a trial random displacement vector (a). In swap moves, two particles are randomly selected and their radii trial swapped (b). 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).
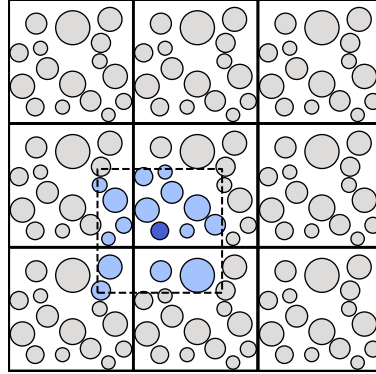
Figure 3.5: Simulation of bulk system is achieved using periodic boundary conditions, where a central cell is surrounded by repeated images of itself. A particle of interest (dark blue) then interacts with the nearest images of every other particle (light blue).

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 [111, 112]. 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 [113]. A two-dimensional Voronoi diagram is formed through the placement of dividing lines between the centroids 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 centroids are required for its construction, with no requirement 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

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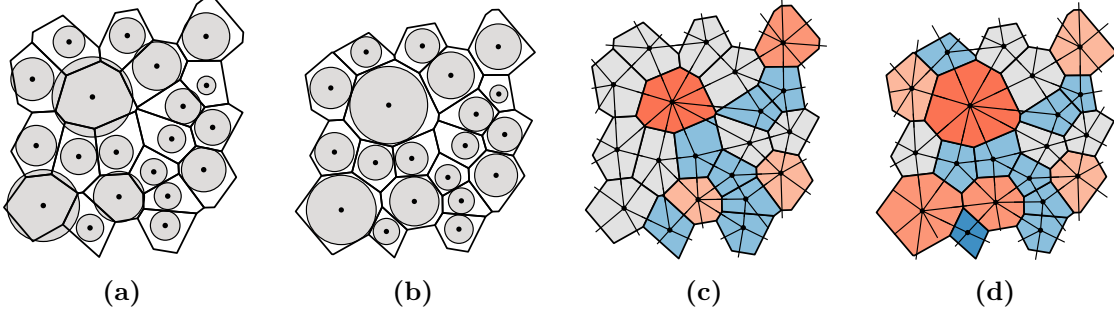
Figure 3.6: Voronoi construction of a polydisperse hard disk system. Panels (a) and (b) compare the unweighted and weighted (radical) Voronoi tessellations respectively. The radical Voronoi assigns more volume to the larger particles to ensure a more equitable distribution of space, which can affect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

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 [114].

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 introduced by Finney [115]. In this modification, the dividing line is placed a distance d_i from particle i, given by:

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

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 [116]. 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 $d_i = R_i$ *i.e.* the radical dividing line sits exactly between the two disks, producing the most equitable distribution of volume (see figure 3.6b). Furthermore, when the radii are equal, $d_i = r_{ij}/2$ 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.

(a) $\phi = 0.8$ **(b)** $\phi = 0.6$ **(c)** $\phi = 0.4$ **(d)** $\phi = 0.2$
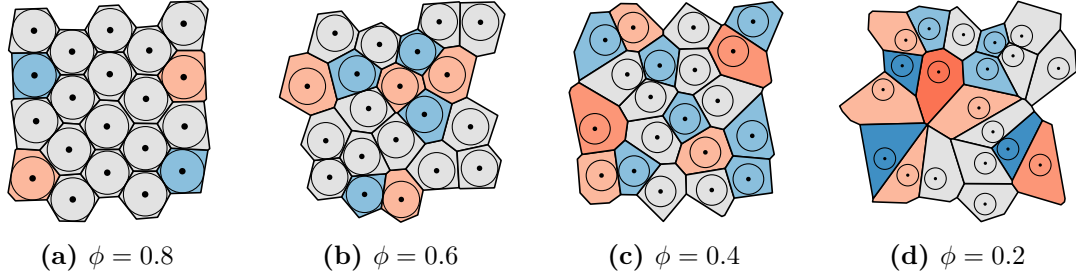
Figure 3.7: The ring structure in Voronoi diagrams is controlled through the packing fraction, ϕ, of the underlying hard particle system. Ring diversity increases as packing fraction is lowered from $0.8 \rightarrow 0.2$ in (a)-(d).

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 $\phi \rightarrow 0$ is well studied as the Poisson Voronoi diagram [117, 118]. 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.

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 Lemâitre'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 21st 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. Put more bilayer stuff in intro?

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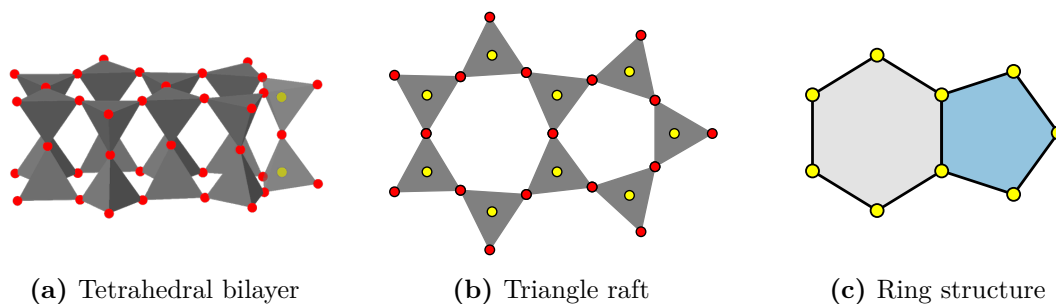
Figure 4.1: Silica bilayers of vertex sharing tetrahedra, (a), can be represented as a two-dimensional triangle raft, (b), (silicon and oxygen atoms are coloured yellow and red respectively). The ring structure then emerges from the three-coordinate network comprising the silicon atoms, (c).

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 [119, 120].

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.

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"

enough for use in computational studies. 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, 121], 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, as will be shown in sec x   [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 [98]. 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 [122, 123]. 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 [124], 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 [125].

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 $r_{OO}^0 = \sqrt{3}\, r_{SiO}^0$, maintaining a set of ideal 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 $r_0 = 2r_{SiO}$. 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:

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

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

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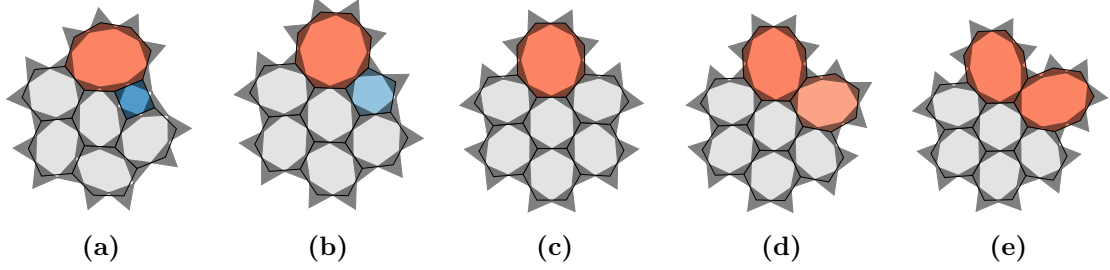
Figure 4.3: Geometry optimised structures for trial rings in the range $k = 4 - 8$. The ring structure is shown along with the SiO_3 triangle

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 [120]. 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.
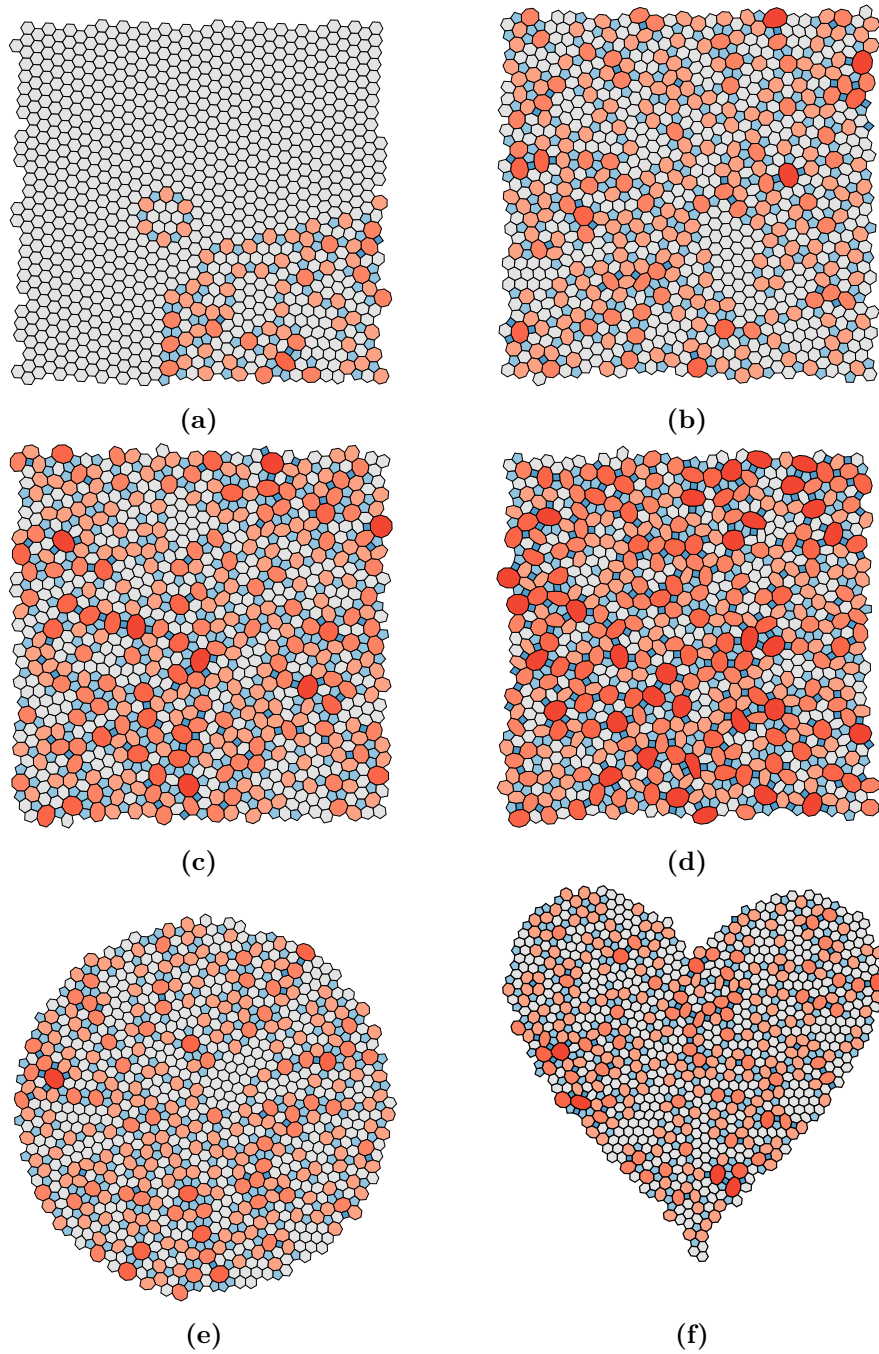
Figure 4.4: Example $1,000$ ring configurations generated with different temperatures and shapes. Panels (a) through (d) show square lattices grown at $T = 10^{-4.0}$, $10^{-3.0}$, $10^{-2.5}$, $10^{-2.0}$ respectively. The samples show the increasing diversity in ring structure as temperature is increased. 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 Lemâitre'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
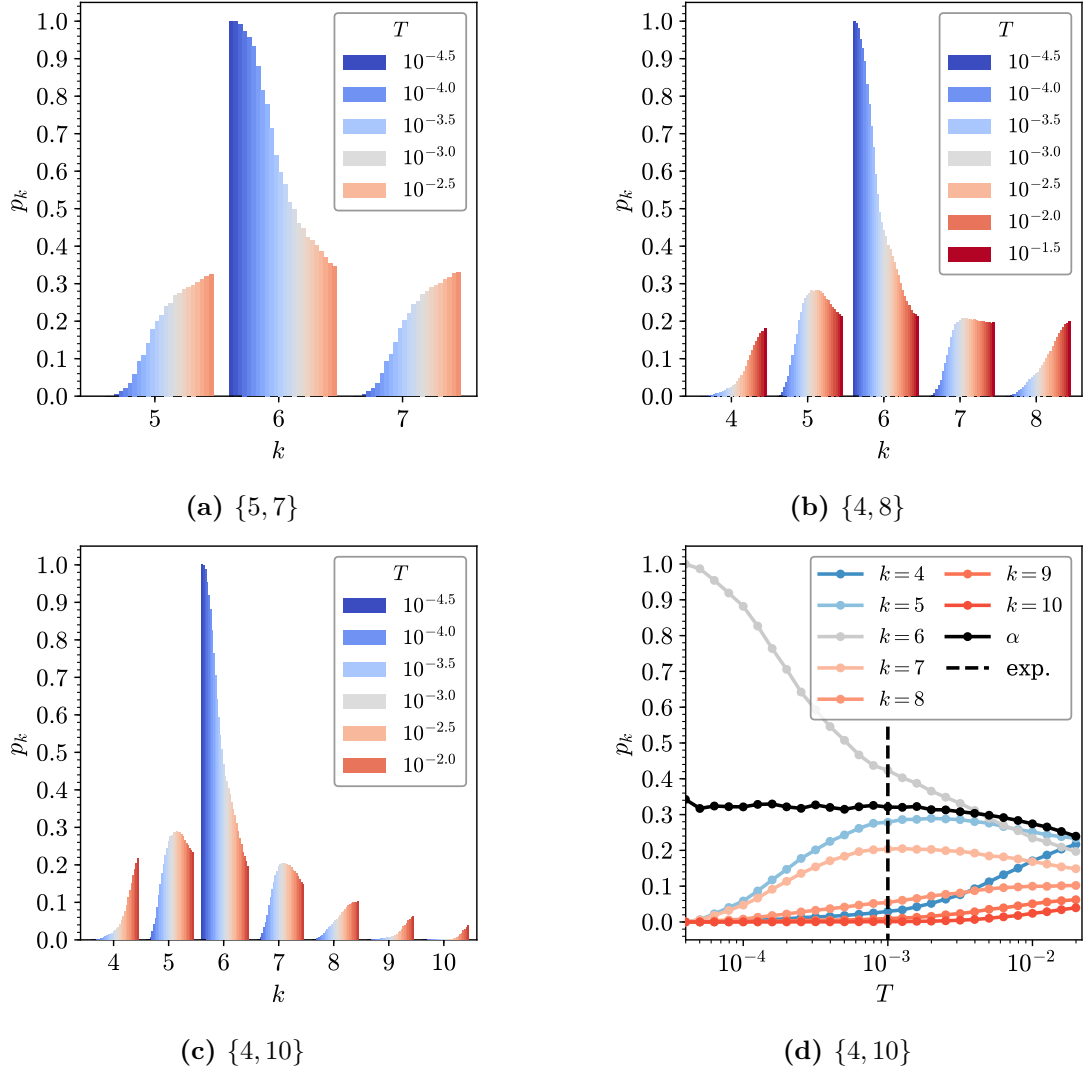
Figure 4.5: Variation in ring statistics with temperature over a given allowable k-range. Panels (a)-(c) show bar graph representations of the ring statistics, coloured by temperature, for the $\{5, 7\}$, $\{4, 8\}$ and $\{4, 10\}$ systems, respectively. Panel (d) gives an alternative line graph representation 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) [123]	Graphene [4]		MC[a] [29]	MC[a] [29]	MD[b] [28]	TR[c]	Lemâitre [56]
$\mathcal{N}$	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

(a)

(b)

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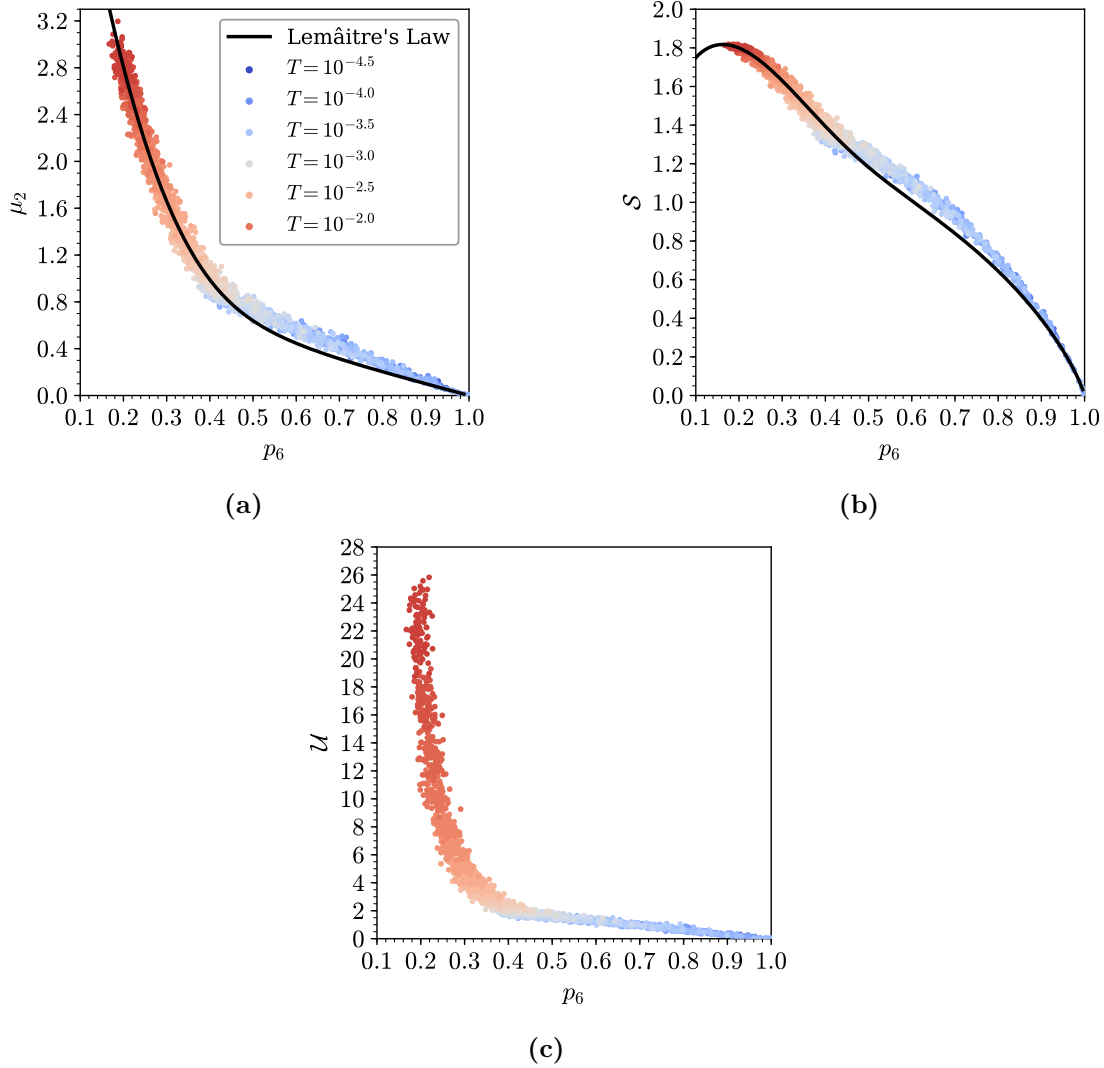
Figure 4.6: Evolution of ring statistics (a), entropy (b) and potential energy (c) of triangle rafts with temperature. The experimental value of $p_6 \approx 0.4$ occurs just before the exponential increase in potential energy, reflecting the balance of energetic and entropic factors.

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

However, Lemâitre'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 Lemâitre'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 Lemâitre'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 Lemâitre 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 $\{4, 10\}$ 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:

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

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

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_{SiSi}^0 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 [98], 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

(a)

(b)

(c)
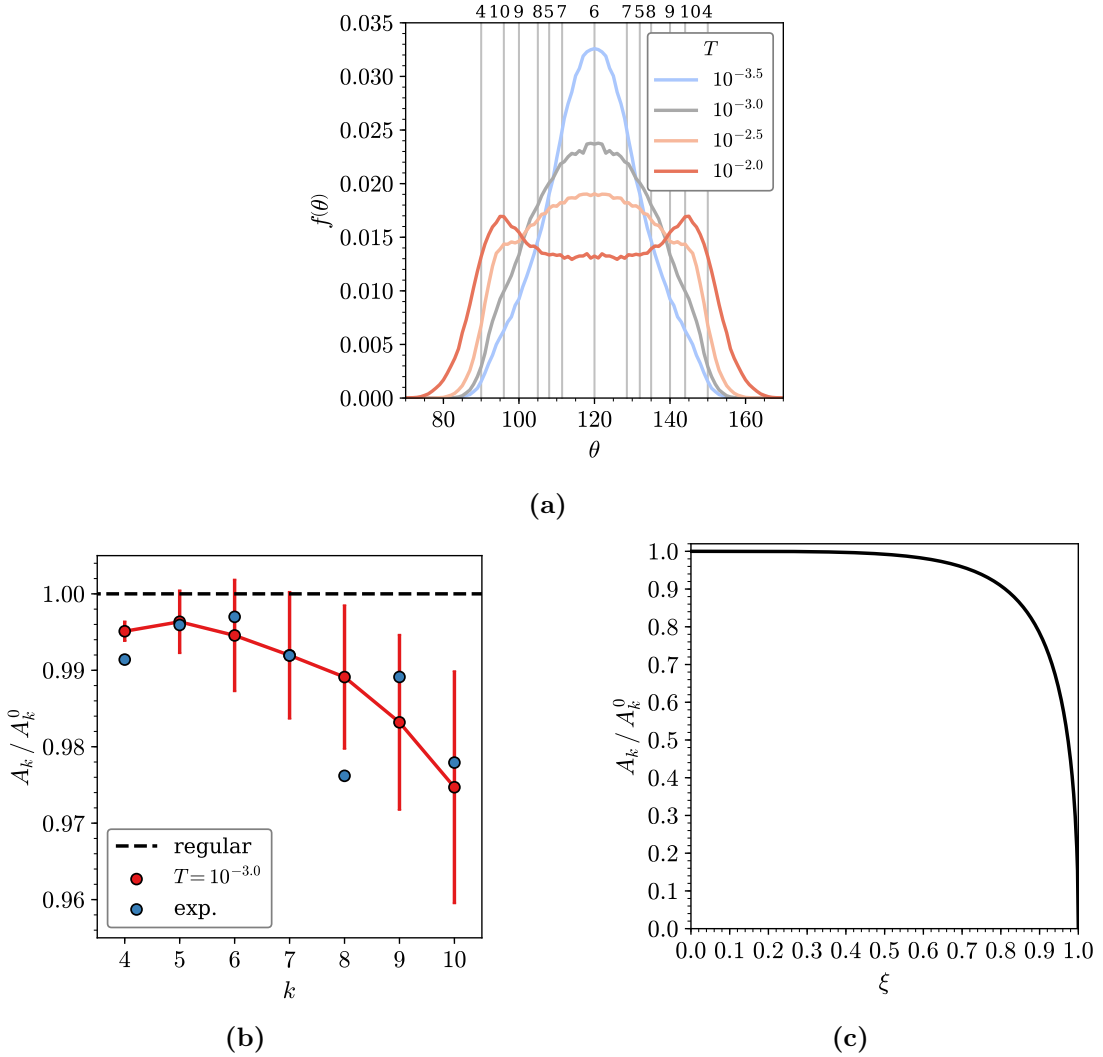
Figure 4.7: Panel (a) gives the ring angle distribution function for triangle rafts formed at different temperatures. Panel (b) compares the regularity of rings in computational and experimental amorphous configurations, with points indicating the mean value and bars corresponding to the standard deviation. 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 [125].

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 Lemâitre's law and has highlighted why the fraction of six-membered rings observed in real systems is often ≈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 the 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.

5.1 Disorder in Two-Dimensional Networks

As mentioned in previous sections link , 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

is in a tight range of $\alpha \approx 0.15 \rightarrow 0.3$ [126]. This is also effect is also seen in computational studies, including for example the previous chapter.

Whilst it is necessary to capture these measures accurately, and almost by definition good computational models capture the underlying essential physics of a system, 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
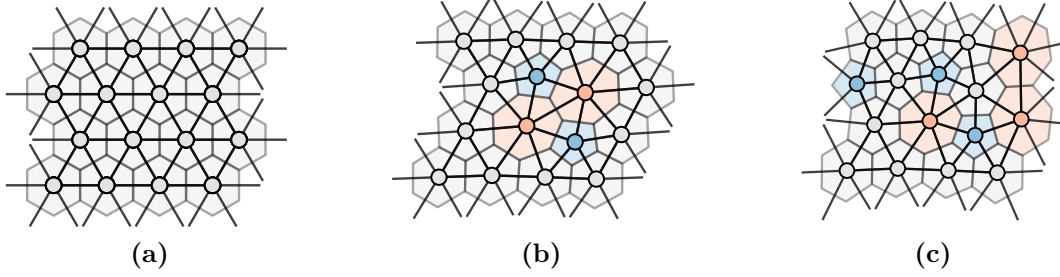
Figure 5.1: Bond switching Monte Carlo moves can be performed solely through the dual lattice. Two successive moves are shown from (a)-(b) and (b)-(c). In the dual lattice (bold circles and lines) two edge-sharing triangles are selected and the shared edge transposed. The atomic network is also shown (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:

$$m_j = \sum_k \frac{k e_{jk}}{q_k} \,.$$ (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 requirement for geometry optimisation after each step, the triangle lattice can be used to cheaply maintain a semi-phyical structure. By applying a harmonic potential between all pairs of linked nodes link   etc etc etc. physical, regenerate atomic positions, speedy

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] 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.

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

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

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

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

[39] 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.

[40] 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.

[41] 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.

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

[43] 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.

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

[45] 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.

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

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

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

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

[50] Alistair R Overy et al. "Design of crystal-like aperiodic solids with selective disorder–phonon coupling". In: *Nat. Commun.* 7 (2016), p. 10445.

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

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

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

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

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

[56] 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.

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

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

[59] 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.

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

[61] 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.

[62] 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.

[63] 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.

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

[65] 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.

[66] 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.

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

[68] 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.

[69] 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.

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

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

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

[73] 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.

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

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

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

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

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

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

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

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

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

[83] 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.

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

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

[86] 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.

[87] 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.

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

[89] 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.

[90]   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.

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

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

[93]   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.

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

[95]   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.

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

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

[98]   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.

[99]   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.

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

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

[102] 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.

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

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

[105] 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.

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

[107] 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.

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

[109] Masaharu Isobe and Werner Krauth. "Hard-sphere melting and crystallization with event-chain Monte Carlo". In: *J. Chem. Phys.* 143 (2015), p. 084509. arXiv: `arXiv:1505.07896v2`.

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

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

[112] 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.

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

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

[115] 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.

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

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

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

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

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

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

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

[123] 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.

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

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

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