# 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

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 [2–6]. 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 [7–9].

It is clear that understanding the structure of amorphous materials is key to this aim. However, due to the relative recentness of these experimental discoveries, much of the existing theory arises from studies of systems on greater length scales. Specifically, in the second half of the $20^{\text{th}}$ century, much work was 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 [10, 11]. 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 [12].

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 log-normal or maximum entropy distribution [13–15]. 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.

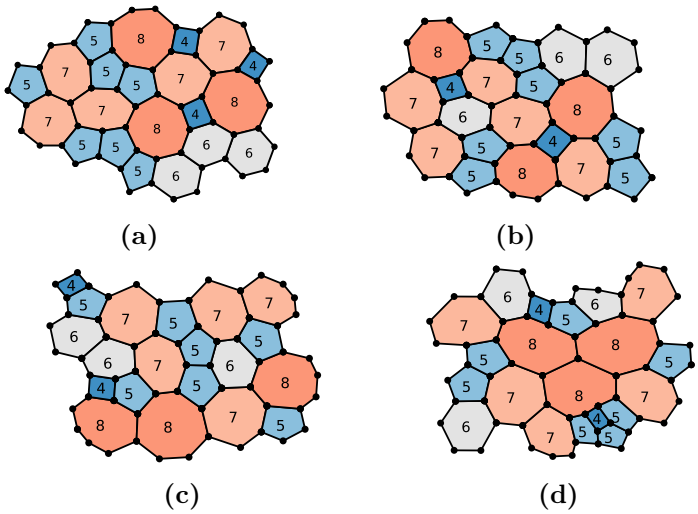



**Figure 1.1:** Panel (a) shows Zachariasen's glass and panels (b)-(d) three different periodic arrangements based on the glass (with one square removed to satisfy Euler's formula). Moving from panel (b)-(d) there is increased clustering of similar sized rings. The size of the rings are highlighted numerically 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 [16, 17]. 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 [18–25]. 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 [26, 27]. 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 [28, 29]. 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 [30–34]. There are however countless more occurrences, with drying blood, stratocumulus clouds, crocodile scales and geopolitical borders all being the subject of studies [35–38]. More intriguingly, although these systems are incredibly physically diverse, they still have similar structures [39]. 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 [40, 41]; as well as to curiosities such as explaining the arrangement of the stones in Giant's Causeway or cracking in famous artworks [42, 43].

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 [44–46]. 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 [47]. The second are the ring structures in so-called "procrystalline" lattices [48]. Paragraph to wrap up

## 1.1 Thesis Structure

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

# 2 | Network Theory & Computational Methods

This chapter discusses the theory underpinning complex networks, 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 . In addition some of the existing computational approaches for generating realisations of two-dimensional networks are presented. A broad overview of Monte Carlo methods is given, before specific methods are covered in detail. These include 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 . Finally some measures are discussed...

## 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 [49]. 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.

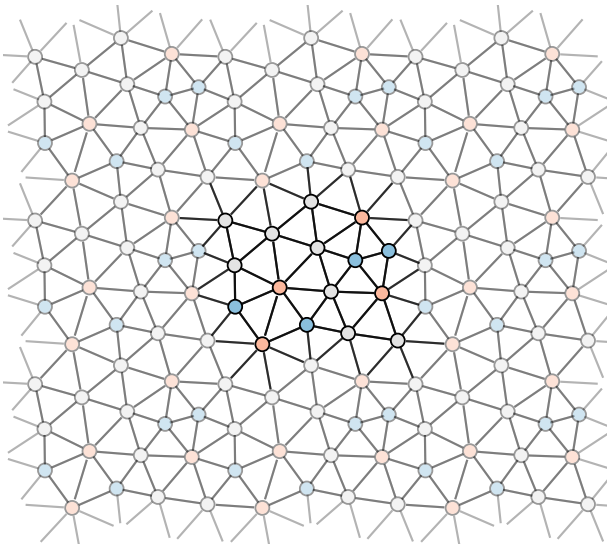


**Figure 2.1:** Example of a periodic two-dimensional network where nodes are represented by circles and links as lines. Nodes are coloured similarly according to their degree, whilst periodic images are greyed out to highlight the central repeating unit.

### 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 $n^{\text{th}}$ moments of this distribution are then given by:

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

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

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

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

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:

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

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

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

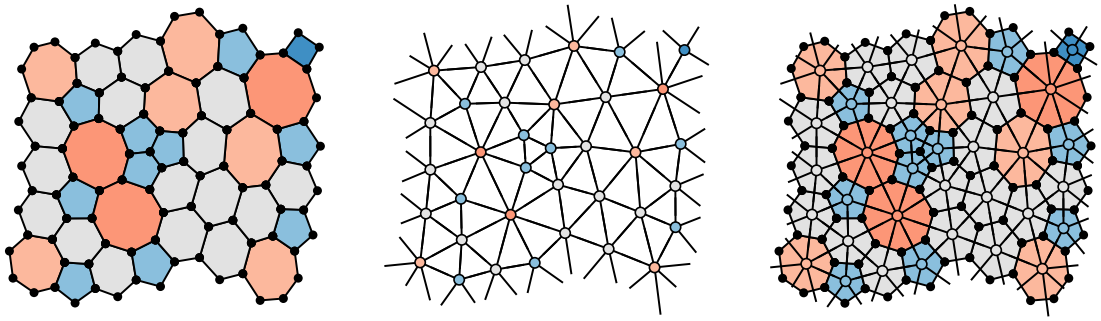
where the final result arises from reciprocal nature of the links in an undirected network. As an example, these three probability distributions are provided for the network in figure 2.1:

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

### 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 [50]. 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 [27]. 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 [51]. 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 $\chi$ 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:

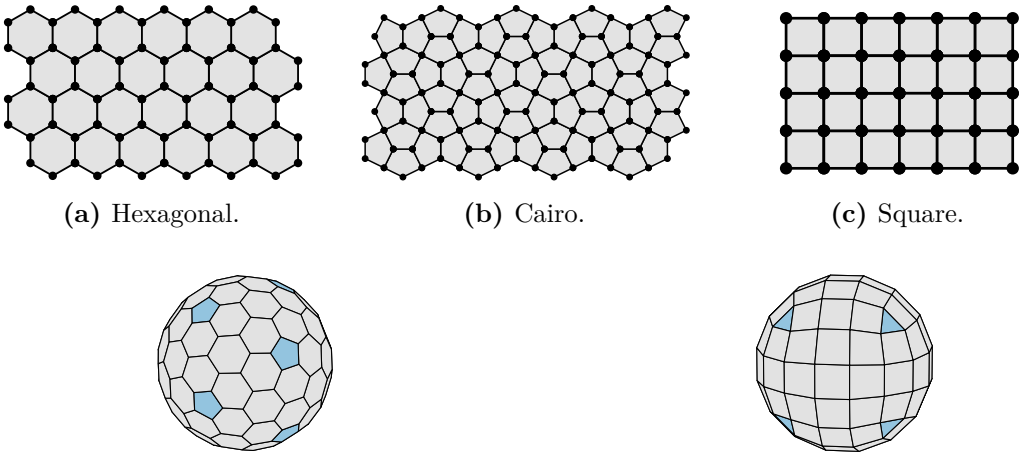
$$\langle k \rangle = \begin{cases} 6, & x_3 = 1 \\ 4, & x_4 = 1 \\ 5, & x_3 = 2/3, x_4 = 1/3. \end{cases} \tag{2.11}$$

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 $\chi$. 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 [34, 52].

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) [53]. 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 [9, 13]. Lemâitre *et al.* demonstrated that the distribution in 3-coordinate networks systems can be well

described by a maximum entropy distribution [54]. Lemâitre's maximum entropy method is summarised here, trivially extended to arbitrary coordination.

The entropy of a probability distribution is defined as $S = -\sum_k p_k \log p_k$. In addition, the degree distribution has the following constraints:

$$\sum_k p_k = 1, \tag{2.13}$$

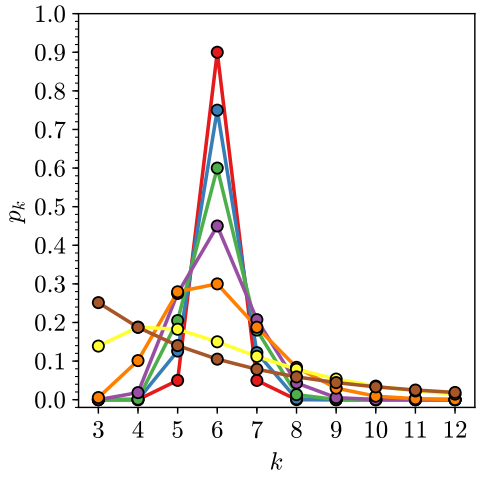
$$\sum_k k p_k = \langle k \rangle, \tag{2.14}$$

$$\sum_k \frac{p_k}{k} = \text{constant}, \tag{2.15}$$

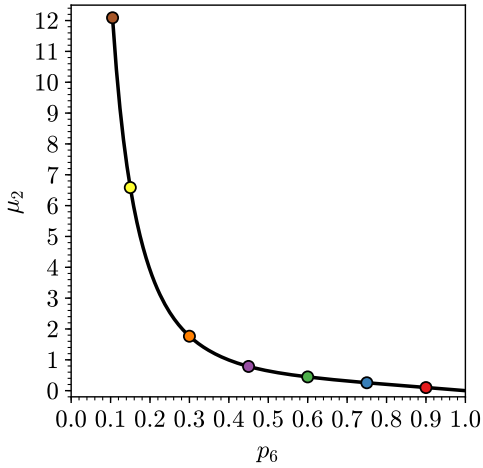
where the first two constraints correspond to the normalisation condition and the fixed mean ring size, and the final constraint will be discussed below. The entropy can then be maximised using Lagrange's method of undetermined multipliers to yield the result:

$$p_k = \frac{e^{-\lambda_1 k - \lambda_2/k}}{\sum_k e^{-\lambda_1 k - \lambda_2/k}}, \tag{2.16}$$

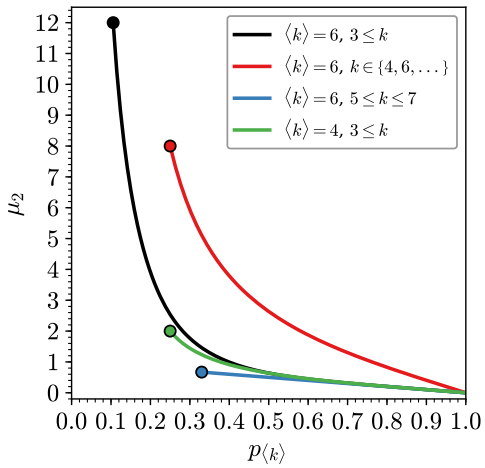
which can be solved numerically by substitution into equations (2.14),(2.15). By allowing the chosen constant to vary, a family of maximum entropy curves can be generated, as in figure 2.4a. The resulting distributions can be summarised by relating the variance, $\mu_2 = \langle k^2 \rangle - \langle k \rangle^2$, 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 [55–57]. 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 [58]. 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 (2.15), equivalent to setting $\lambda_2 = 0$ in equation (2.16).




**(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.* $\mu_2$ (Lemâitre's law). Panel (c) provides extensions to the law by modifying the underlying constraints of the mean ring size and allowable $k$-range.

The only somewhat puzzling aspect of this successful theory is the choice of constraint (2.15). 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 [59–61]. Despite this, the universality of the Lemâitre law suggests that there must be a physical basis to (2.15), and

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 [16]. 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

$$\sum_k k m_k p_k = \sum_k k^2 p_k = \mu_2 + \langle k \rangle^2 , \tag{2.17}$$

known as Weaire's sum rule [17]. 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 link to Poisson-Voronoi ). Aboav then proposed that if a generic form of $m_k = A + B/k$ was used in conjunction with Weaire's sum rule then

$$m_k = A + \frac{\mu_2 + \langle k \rangle^2 - A\langle k \rangle}{k} . \tag{2.18}$$

This is now more commonly expressed in the linear form [62]:

$$k m_k = \mu_2 + \langle k \rangle^2 + \langle k \rangle (1 - \alpha)(k - \langle k \rangle) . \tag{2.19}$$

Equation 2.19 is known as the Aboav-Weaire law and relates the mean ring size about a given central ring to a single fitting parameter, $\alpha$. The value of $\alpha$ 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 [63]:

$$\alpha = -\frac{\mu_2}{\langle k \rangle^2} \,. \tag{2.20}$$

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.

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 [26, 64–67]. 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 $\alpha$ 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 Monte Carlo Methods

Monte Carlo methods are a class of computational algorithms designed to solve complex problems stochastically. These normally fall into the broad categories of calculating integrals, sampling probability distributions and finding global minima of very high dimensional functions - tasks which are often incredibly hard to compute deterministically. Since their initial development in the mid-20$^{\text{th}}$ century, such methods have become an invaluable tool for solving problems in the physical sciences. Monte Carlo methods are used in this context for calculating thermodynamic averages of properties in equilibrium systems; finding the minima in potential energy surfaces of small molecules, glasses, crystals and biomolecules; as well as non-equilibrium simulations such as growth of crystals and thin-films [68–73]. In

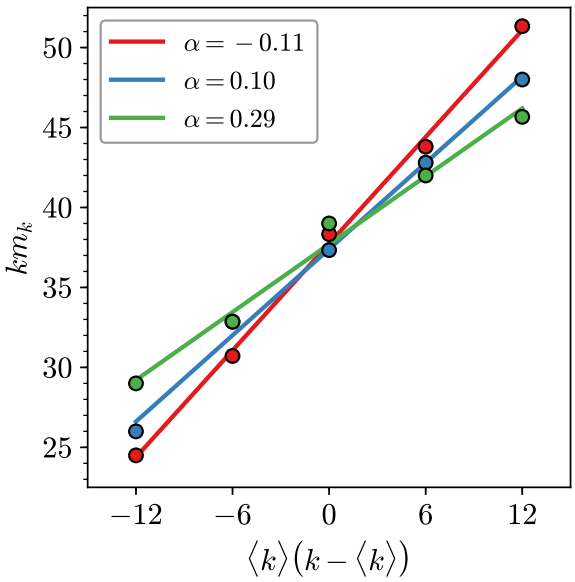



**Figure 2.5:** Calculation of an Aboav-Weaire fit for three configurations (shown in figure 1.1(a)-(c)). The value of the $\alpha$ parameter quantifies the tendency of small rings to be adjacent to large rings, with a larger value indicating stronger small-large ring correlations.

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 along with specific details of two established methods: bond switching and hard particle Monte Carlo.

### 2.3.1 Statistical Mechanics

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

$$\mathcal{H}(\mathbf{p}, \mathbf{r}) = \mathcal{K}(\mathbf{p}) + \mathcal{U}(\mathbf{r}) \,, \tag{2.21}$$

where $\mathcal{K}(\mathbf{p})$ is the kinetic energy as a function of all particle momenta and $\mathcal{U}(\mathbf{r})$ is the potential energy as a function of all particle positions [74]. The positions and momenta comprise the phase space of the system. At fixed volume, $\mathcal{V}$, and temperature, $T$, all the the essential thermodynamic information is then provided through the classical canonical partition function:

$$Q = \frac{1}{h^{D\mathcal{N}}\mathcal{N}!} \int \mathrm{d}\mathbf{p}\, \mathrm{d}\mathbf{r}\, \exp\left[-\mathcal{H}(\mathbf{p}, \mathbf{r}) / k_{\mathrm{B}}T\right], \tag{2.22}$$

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

$$Q = \frac{1}{h^{D\mathcal{N}}\mathcal{N}!}\int \mathrm{d}\mathbf{p}\, \exp\left[-\mathcal{K}\left(\mathbf{p}\right)/k_{\mathrm{B}}T\right]\int \mathrm{d}\mathbf{r}\, \exp\left[-\mathcal{U}\left(\mathbf{r}\right)/k_{\mathrm{B}}T\right], \tag{2.23}$$

where

$$Z = \int \mathrm{d}\mathbf{r}\, \exp\left[-\mathcal{U}\left(\mathbf{r}\right)/k_{\mathrm{B}}T\right] \tag{2.24}$$

is the configurational integral [75]. 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:

$$P\left(\mathbf{r}\right) = \frac{\exp\left[-\mathcal{U}\left(\mathbf{r}\right)/k_{\mathrm{B}}T\right]}{Z}. \tag{2.25}$$

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

$$\langle A\rangle = \int \mathrm{d}\mathbf{r}\, \mathcal{A}\left(\mathbf{r}\right) P\left(\mathbf{r}\right). \tag{2.26}$$

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.

### 2.3.2 Importance Sampling

An integral of form (2.26) 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 2.6. To calculate the expectation value of the potential energy one must evaluate the integral

$$\langle \mathcal{U}\rangle = \int_0^{L_y}\int_0^{L_x} \mathrm{d}x\mathrm{d}y\, \mathcal{U}\left(x,y\right)\mathcal{P}\left(x,y\right). \tag{2.27}$$

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 2.6a, 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 2.6b. Weighting these according to the Boltzmann distribution and averaging gives an estimation to the integral:

$$\langle \mathcal{U} \rangle = \frac{L_x L_y}{S} \sum_{i=1}^{S} \mathcal{U}(x_i, y_i)\, \mathcal{P}(x_i, y_i) \ , \tag{2.28}$$

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 2.6b. The expectation value of the observable can then be calculated using a simple average:

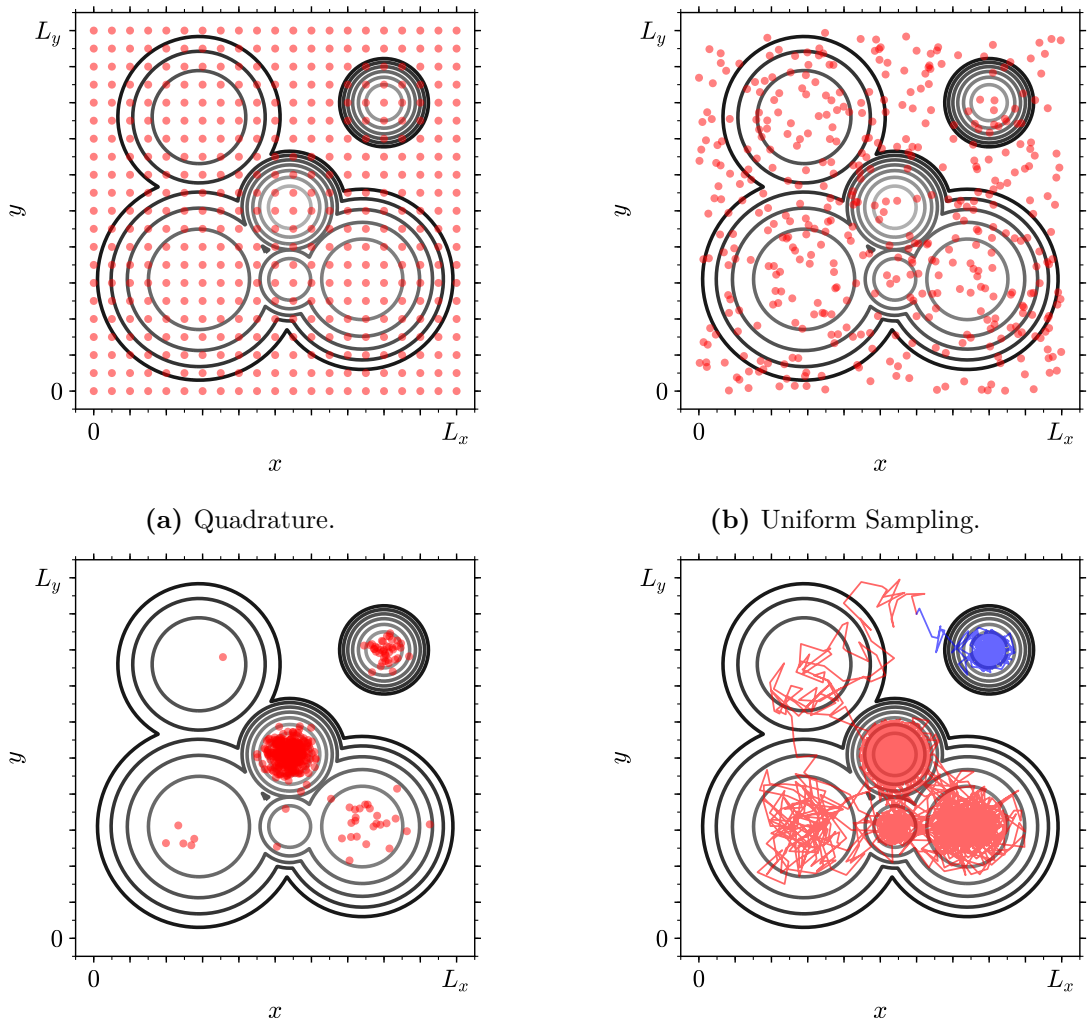
$$\langle \mathcal{U} \rangle = \frac{1}{S} \sum_{i=1}^{S} \mathcal{U}(x_i, y_i) \ . \tag{2.29}$$

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

Whilst this scheme is ideal theoretically, it is impracticable for physical systems. This is because for any problem of real interest one lives in a "black box" where the functional form of the potential energy surface in its hundreds if not thousands of dimensions is unknown. In this case often the only way of learning about the form is by on-the-fly exploration of the surface [76]. This can be achieved by talking a random walk through configurational space using Markov chain Monte Carlo.

### 2.3.3 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)




**(a)** Quadrature.

**(b)** Uniform Sampling.

**(c)** Importance Sampling.

**(d)** Markov Chain Monte Carlo

**Figure 2.6:** 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.

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,

$\pi_{ij}$, gives the probability of moving from the state $\mathbf{r}_i$ to another state $\mathbf{r}_j$. This leads to the two relationships:

$$0 \leq \pi_{ij} \leq 1\,, \tag{2.30}$$

$$\sum_j \pi_{ij} = 1\,, \tag{2.31}$$

the first being a statement of the probabilistic nature of the elements whilst the second ensures all transfer remains within the state space [74–76].

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

$$\mathbf{P} = \mathbf{P}\boldsymbol{\pi}\,, \tag{2.32}$$

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 (2.26), the expectation value of an observable, $A$, can be calculated from the ensemble average:

$$\langle A \rangle = \sum_{i=1}^{M} A\left(\mathbf{r}_i\right) \mathcal{P}\left(\mathbf{r}_i\right)\,, \tag{2.33}$$

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:

$$\langle A \rangle = \frac{1}{S} \sum_{i=1}^{S} A\left(X_i\right)\,, \tag{2.34}$$

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 (2.25). A practical solution to this problem is provided by the Metropolis algorithm.

### 2.3.4 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 [77]. Firstly, combining equations (2.31) and (2.32) gives a condition on the transition matrix known as global balance:

$$\sum_j \mathcal{P}\left(\mathbf{r}_i\right) \pi_{ij} = \sum_j \mathcal{P}\left(\mathbf{r}_j\right) \pi_{ji} \,. \tag{2.35}$$

Whilst it is possible to construct transition matrices which satisfy only global balance [78–80], it is practically simpler to satisfy global balance by applying the stronger condition of detailed balance:

$$\mathcal{P}\left(\mathbf{r}_i\right) \pi_{ij} = \mathcal{P}\left(\mathbf{r}_j\right) \pi_{ji} \,. \tag{2.36}$$

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

$$\pi_{ij} = \begin{cases} \tau_{ij} P_{ij} & i \neq j \\ 1 - \sum\limits_{j \neq i} \tau_{ij} P_{ij} & i = j \end{cases} , \tag{2.37}$$

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

$$P_{ij} = \begin{cases} 1 & \mathcal{P}\left(\mathbf{r}_j\right) \geq \mathcal{P}\left(\mathbf{r}_i\right) \\ \frac{\mathcal{P}(\mathbf{r}_j)}{\mathcal{P}(\mathbf{r}_i)} & \mathcal{P}\left(\mathbf{r}_j\right) < \mathcal{P}\left(\mathbf{r}_i\right) \end{cases} = \begin{cases} 1 & \mathcal{U}\left(\mathbf{r}_j\right) \leq \mathcal{U}\left(\mathbf{r}_i\right) \\ \frac{\exp[-\mathcal{U}(\mathbf{r}_j)/k_\mathrm{B}T]}{\exp[-\mathcal{U}(\mathbf{r}_i)/k_\mathrm{B}T]} & \mathcal{U}\left(\mathbf{r}_j\right) > \mathcal{U}\left(\mathbf{r}_i\right) \end{cases} , \tag{2.38}$$

which can be expressed more succinctly as

$$P_{ij} = \min\Big[1, \exp\left[-\Delta\mathcal{U}/k_{\mathrm{B}}T\right]\Big], \tag{2.39}$$

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 $\tau_{st}$
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 (2.39)
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 2.6d. 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 2.6c. 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 [81, 82].

### 2.3.5 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 2.6d until a solution is found, using the more general Metropolis criterion:

$$P_{ij} = \min\Big[1, \exp\left[-\Delta\Omega/k_{\mathrm{B}}T\right]\Big]. \tag{2.40}$$

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 2.6d). 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 [83–85]. This thesis will use simulated annealing, which can be considered an extension to Metropolis Monte Carlo [86]. 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 (2.40) 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 [87]. In practice this is not realisable and so a cooling rate must be empirically selected. Still it is possible for trapping to occur in local minima, especially if the transition between low energy states is very slow. As before, one can then cycle the simulated annealing, repeatedly heating and cooling the system until the global minimum is found. The simulated annealing algorithm is demonstrated with the two-dimensional potential energy surface in figure 2.7. 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.

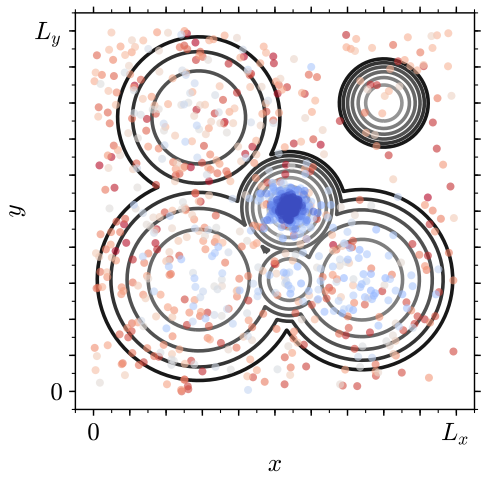



**Figure 2.7:** Demonstration of the simulated annealing algorithm on a two-dimensional potential energy surface, with states coloured by temperature (red→blue indicating hot→cold. As the temperature is reduced the state converges on the global minimum.

# 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. “Crystalline-vitreous interface in two dimensional silica”. In: *Phys. Rev. Lett.* 109 (2012), p. 106101.

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

[7] 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.

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

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

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

[11] 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.

[12] 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.

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

[14] 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.

[15] 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.

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

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

[18] 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.

[19] 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.

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

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

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

[23] 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.

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

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

[26] 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.

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

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

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

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

[31] 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.

[32] 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.

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

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

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

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

[37] 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.

[38] 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.

[39] 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.

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

[41] 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.

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

[43] 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.

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

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

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

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

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

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

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

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

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

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

[54] 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.

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

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

[57] 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.

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

[59] 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.

[60] 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.

[61] 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.

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

[63] 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.

[64] 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.

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

[66] 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.

[67] 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.

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

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

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

[71] 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.

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

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

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

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

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

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

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

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

[80] Manon Michel, Sebastian C Kapfer, and Werner Krauth. "Generalized event-chain Monte Carlo: Constructing rejection-free global-balance algorithms from infinitesimal steps". In: *J. Chem. Phys.* 140 (2018), p. 054116.

[81] 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.

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

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

[84] 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.

[85] 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.

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

[87] 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.