

Contents

1	Introduction	1
2	Statistical Physics	4
2.1	The Need for Statistical Physics	4
2.2	The Ising Model	6
3	Monte Carlo Simulations	9
3.1	Monte Carlo Sampling	9
3.1.1	Sampling	9
3.1.2	Markov Process	10
3.1.3	Detailed Balance	11
3.2	Algorithms	13
3.2.1	Metropolis Algorithm	13
3.2.2	Cluster Algorithm	16
3.3	Thermalization	18
3.4	Autocorrelations and Correlations	19
3.5	Critical Exponents	22
3.6	Statistical Errors	23
3.6.1	Errors of Independent Measurements	24
3.6.2	Blocking	25
4	Exact and Statistical Calculations	26
4.1	An Exact Computational Approach.	26
4.2	Testing the Algorithms	27
5	Simulations	29
5.1	Monte Carlo Histories	29
5.2	Sample Size	31

5.3	Observables and their critical exponents	32
5.4	Critical Slowing Down	38
6	Conclusions	43
6.1	Acknowledgement	44

Monte Carlo Simulations in the 2-dimensional Ising Model.

Anthony Bourached

Supervisor: Professor Stefan Sint

A thesis presented for
module MA4492



School of Mathematics
Trinity College Dublin
Ireland

15 February 2015

Abstract

We perform Monte Carlo simulations on the 2-dimensional Ising model in a zero magnetic field, using the Metropolis and Wolff Cluster algorithms to obtain computational results for thermodynamic quantities and the dynamic critical exponents of the algorithms. The Ising model in 2-dimensions and zero magnetic field undergoes a second order phase transition from a phase with no magnetisation to one with spontaneous magnetisation. This phase transition is modelled by our simulations with the behaviour of magnetisation, energy, magnetic susceptibility and heat capacity at phase transition all being accurately valued. We used finite size scaling of the lattice at critical temperature to derive the value of the dynamic critical exponent for both algorithms of $z = 2.11 \pm 0.073$ for Metropolis and 0.47 ± 0.051 for Wolff Cluster which compares well with accepted quantities of $z_{metropolis} = 2.17$ and $z_{cluster} = 0.52$. Critical exponents of derived observables magnetic susceptibility (γ) and heat capacity (α) were found to be $\gamma = 1.76 \pm 0.023$ and $\alpha = 0.29 \pm 1.5$. γ compared well with the accepted value of $\gamma = 7/4$. The discrepancy of α from the accepted value 0 was discussed.

Chapter 1

Introduction

The Ising model, named after Ernst Ising who first studied it in 1923, has played a central role in the development of statistical physics and quantum field theory [1] and is one of the most famous statistical systems. Of particular interest is the two dimensional Ising model which is complicated enough to possess non-trivial properties yet still simple enough to have an exact analytical solution. For several years the Ising model was considered as a greatly over-simplified representation of intermolecular forces and as being inapplicable to any real systems [2]. In fact Ernst Ising gave up on his research in physics after he believed that he had proved that the model had no physical usefulness, yet twenty years later he was to become famous as a result of his work on the model [2].

In 1923 Ising showed that the one-dimensional model does not exhibit magnetic behaviour in the absence of a magnetic field [3]. He assumed that this result could be extrapolated in higher dimensions. However, this was incorrect. The Ising model zero magnetic field, in two or more dimensions, undergoes a second-order phase transition at a critical temperature T_c from a disordered, high temperature state with zeroed magnetization; to an ordered, low temperature state that has spontaneous magnetization. This is the case only for a finite size lattice; in the thermodynamic limit there is no spontaneous magnetization.

It was only in the late 1940s that the Ising model began to rise in popularity. The explicit solution of the two-dimensional problem involved such difficult mathematics that it stumped all physicists who attempted it. In fact, it was a chemist (Onsager 1948) who found the full analytical solution to the two dimensional model in zero magnetic field, which remains one of the most important achievements in statistical mechanics. The result is interesting from a physical point of view since it is the simplest, most phenomenologically interesting, model

of a ferromagnetic material. Due to universality, the model describes also the liquid/vapor phase transition at the triple point [1].

Another approach to studying the Ising model, which is the focus of this report, is Monte Carlo simulations. These are highly efficient computational methods for solving dynamic systems in thermal equilibrium. The goal of this report is to examine in detail the errors and efficiencies of two Monte Carlo methods which greatly differ in their approach to solving the model. Their strengths and weaknesses will be compared and contrasted. The two algorithms are the local change Metropolis algorithm and the non-local Wolff Cluster algorithm.

In Chapter 2 we shall review the fundamentals of statistical physics that are essential in understanding the Ising model and, in section 2.2, how this basis in statistical physics leads us to derive bulk properties of the Ising model from the hamiltonian of the system.

Chapter 3 shall give an introduction to the Monte Carlo computation methods and a detailed discription of the algorithms used. We will also discuss in detail the method of deriving bulk properties from our functioning algorithms and how to become confident that our results are representative of the equilibrium state of the system. We shall further discuss the concept of autocorrelation and how this effects our results. Chapter 3 will be concluded with a discussion of how error analysis was conducted in this project.

In Chapter 4 we shall discuss an exact (non-statistical) computational solution to the model which, for reasons that shall be discussed is only practical for very small systems. We shall also compare these exact results to those found by our statistical approach to ensure that they are in strong agreement within our calculated margin of error.

In Chapter 5 we shall present a discussion on data produced by our simulations. We shall use the methods described in 3 to calculate expectation values of various observables and their associated errors. In particular we shall focus on the computation of autocorrelation times of the algorithms and how they handle critical slowing down. This is expected to be the greatest variant between the algorithms as the Wolff Cluster algorithm was introduced as an effective way of beating the effects of critical slowing down. We shall further discuss the implications of these comments on the efficiencies of the system and our choice of algorithm later.

Finally, in Chapter 6 the overall achievements and observations of this project will be summarised. The most important discovery will be highlighted and its implications underlined. This will conclude the project.

Chapter 2

Statistical Physics

2.1 The Need for Statistical Physics

A simple macroscopic system generally has between 10^{23} and 10^{44} number of degrees of freedom. This could describe anything from water vapour to an entire planet. Hence, it is impractical to determine exact results from the evolution of the system. It is enough, however, to know a small number of bulk properties of the system in order to have a useful description of it.

For example, we have sufficient information from the energy and density of a fluid without needing to know the position; momentum; energy and angular momentum of each individual molecule which would provide us with more information than we need, or than it would be possible to analyse. It is possible to use Monte Carlo simulations to find these bulk properties from the hamiltonian of the system.

We can consider a system to have N discrete states each with a discrete energy $E_0 < E_1 < \dots < E_\mu < \dots < E_{N-1}$. This system is in contact with a thermal reservoir of temperature T . When the system is in thermodynamic equilibrium with the reservoir, its bulk properties are described by the canonical ensemble. Weights describe the probability for the system to be in state μ at time t . The probability for state μ with energy E_μ is given by p_μ and is the connection between the microscopic and statistical descriptions of the system. The Boltzmann distribution gives:

$$p_\mu = \frac{e^{-\beta E_\mu}}{Z} \quad (2.1)$$

β is the known as the coupling temperature and is given by:

$$\beta = \frac{1}{kT} \quad (2.2)$$

Where k is Boltzman's constant. β is referred to as the inverse temperature and is sometimes used as the independent variable in calculation rather than T . In this paper, it is indeed β to which we often refer. Where high temperature refers to low β and low temperature refers to high β . As a matter of convention we also choose natural units i.e. $k = 1$.

Z is a normalising factor known as the partition function. It encrypts all the thermodynamic information about a system.

$$Z = \sum_{\mu} e^{-\beta E_{\mu}} \quad (2.3)$$

All thermodynamic properties of a system can be calculated from the partition function for that system. It is simply the sum of the weights of each state of the system. We can see that for a very high β ($T \rightarrow 0$) equation (2.1) becomes $p_{\mu} = 1/(\text{number of states})$ i.e. each state equally likely.

Thermodynamic systems have a stochastic (non-deterministic) character. Thus if we want to measure a physical quantity, or *observable* (O), we compute the thermal average, as the probability of measuring a quantity to take a value significantly different from its average is ridiculously small. This average is referred to as the expectation value of the observable $\langle O \rangle$. The expectation value is found by averaging the observable over all possible states of the system, weighted by Boltzmann probabilities:

$$\langle O \rangle = \frac{1}{Z} \sum_{\mu} O_{\mu} e^{-\beta E_{\mu}} \quad (2.4)$$

where O_{μ} is the value of O in state μ .

The observables of particular interest in this paper are magnetization and the internal energy of the system to which we, for convenience, refer to simply as energy. For a statistical system we may calculate the expectation values of these observables using equation (2.4).

$$\langle M \rangle = \frac{1}{Z} \sum_{\mu} M_{\mu} e^{-\beta E_{\mu}} \quad (2.5)$$

$$\langle E \rangle = \frac{1}{Z} \sum_{\mu} E_{\mu} e^{-\beta E_{\mu}} \quad (2.6)$$

These quantities are almost always measured as other quantities can be derived from them. Such as the susceptibility:

$$\chi = k\beta(\langle M^2 \rangle - \langle M \rangle^2) \quad (2.7)$$

and the heat capacity:

$$C = k\beta^2(\langle E^2 \rangle - \langle E \rangle^2) \quad (2.8)$$

which are expressed above as the variation of magnetization and energy respectively and are hence referred to as derived observables since when we calculate them in this way it is not necessary to take any further measurements from the system. k is a constant specific to the system.

2.2 The Ising Model

Consider a d -dimensional lattice with a scalar valued elementary spin at each lattice point of $s_i = \pm 1$. This may be considered to be an atom or a magnet due to the universality of the model. Each side is of length L giving a total $N = L^d$ number of spin sites. Sites are connected to one another via links to their nearest neighbours with whom they directly interact. This spin-spin interaction is what determines the dynamics of the system.

At the boundary of the lattice periodic boundary conditions are usually employed such that the rightmost site's nearest neighbour to the right is the leftmost site. This is clearly illustrated in figure 2.1.

The Hamiltonian for the Ising model is given by:

$$H = -J \sum_{\langle i,j \rangle} s_i s_j - B \sum_i s_i \quad (2.9)$$

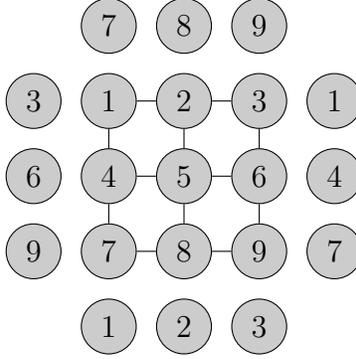


Figure 2.1: 2-dimensional $L = 3$ lattice, $N = L \times L = 9$. Links between nearest neighbours are seen as lines connecting sites and periodic boundaries are shown for this small lattice. Note, for illustration purposes the links across the boundary are not shown here but do exist.

where B is the external magnetic field. We consider the 2-dimensional ising model in a zeroth magnetic field ($B=0$) and so the Hamiltonian is given by:

$$H = -J \sum_{\langle i,j \rangle} s_i s_j \quad (2.10)$$

where the notation $\langle i,j \rangle$ is a sum over nearest neighbours for each given point, or in other words: $\sum_{\langle i,j \rangle} s_i s_j = \sum_{i,j} (s_{i,j} \cdot s_{i,j+1} + s_{i,j} \cdot s_{i+1,j})$. We only take the two links (up and right in this case) since once we have summed over all spins we will have accounted for each link; two vertices are connected by one link and four links intersect at one vertex:

$$2N_l = 4N \Rightarrow N_l = 2N \quad (2.11)$$

J is the exchange energy (for interaction between spins). If $J > 0$ then the energy of the system is lowered when spins align which is known as ferromagnetic behaviour. $J < 0$ causes the system to have lowest energy when spins are anti-parallel, referred to as anti-ferromagnetism. The system is considered to be in thermodynamic equilibrium when its energy reaches a steady state minimum and for this reason the system tends to ferromagnetism for $J > 0$ and paramagnetism for $J < 0$ for a system with a non-zero external magnetic field.

We can consider the total Magnetisation of the system in state μ to be given by:

$$M = \left| \sum_i s_i \right| \quad (2.12)$$

The probability of the appearance of a state depends only on the Hamiltonian H of the system. Since the Hamiltonian is symmetric under reflection of all the spins two configurations of opposite sign are equally likely. These configurations will have opposite magnetizations, therefore the average magnetization $M = \langle \sum_i s_i \rangle$ will be zero. This is the reason for the absolute value sign in equation (2.12).

Spin interaction makes it possible that the model will end up in a state with most of the spins aligned despite the absence of an external magnetic field. This is called spontaneous magnetisation. The question of whether or not the Ising model exhibits spontaneous magnetisation in a particular dimension is a large motivation in the study of the ising model. Furthermore, it is the focus of this curiosity to determine if there exists a critical temperature $\beta_c = 1/T_c$ at which the model undergoes a phase transition from exhibiting no spontaneous magnetisation to a spontaneously magnetised region.

From the Hamiltonian it is clear to see that energy of the system in a given state may be calculated by:

$$E = - \sum_{\langle i,j \rangle} s_i s_j \quad (2.13)$$

We can derive quantities of susceptibility and heat capacity, as discussed in the previous section, using equations (2.7) and (2.8) where k for the Ising model is the number of sites N .

It can be derived analytically [1], [4] that in 2-dimensions (with $B = 0$), the Ising Model has a phase transition at critical temperature $\beta = \beta_c$ where:

$$\beta_c = \frac{1}{2} \ln(1 + \sqrt{2}) = 0.4406867935... \quad (2.14)$$

Chapter 3

Monte Carlo Simulations

3.1 Monte Carlo Sampling

For the most interesting statistical mechanical system, the partition function can't be computed analytically. Monte Carlo simulations are a class of algorithms that rely on random sampling to obtain numerical results that approximate, to a high degree of accuracy, the observable quantities in a dynamical system.

3.1.1 Sampling

We would like to calculate expectation values $\langle O \rangle$ using equation (2.4). However for large systems, such as the 2 dimensional ising model where we generally want to consider systems with $L = 100$ and so $2^{100 \times 100}$ states it is impossible (with even the most advanced super computers) to calculate the hamiltonian for each individual state. (This is possible for a very small lattice as will be discussed in Chapter 4.) Thus we need to estimate the expectation value using sampling.

We construct a sample of N states $\mu_1, \mu_2, \dots, \mu_N$ which are distributed according to a *chosen* probability distribution, P_μ , which should depend on the temperature of the system. Then we can define an estimator (O_N) of $\langle O \rangle$ to be:

$$O_N = \frac{\sum_{i=1}^N O_{\mu_i} e^{-\beta E_{\mu_i}} P_{\mu_i}^{-1}}{\sum_{i=1}^N e^{-\beta E_{\mu_i}} P_{\mu_i}^{-1}} \quad (3.1)$$

This is easily understood since, for a sufficiently large sample, P_μ will simply be the frequency that the system is in state μ and so we expect that:

$$\langle O \rangle = \lim_{N \rightarrow \infty} O_N \quad (3.2)$$

We wish to determine a value P_μ which causes the convergence in equation (3.2) as fast as possible. Simple sampling sets P_μ as a constant. But since our sample is generally tiny compared to the number of states in the system the probability of getting many states that make a large contribution to the sum in equation (2.4) is low since this is a vast minority of states.

Importance sampling chooses a state μ and adds it to the sample with probability $P_\mu = \frac{e^{-\beta E_\mu}}{Z}$. In this way the states chosen should have the most significant contribution to the calculation of $\langle O \rangle$. This is the sampling process used in most Monte Carlo Simulations. Putting this probability into equation (3.1) we get the expression

$$O_N = \frac{1}{N} \sum_{i=1}^N O_{\mu_i} \quad (3.3)$$

3.1.2 Markov Process

It is not possible to choose a state μ and add it to the sample with probability $P_\mu = \frac{e^{-\beta E_\mu}}{Z}$ in a direct way as it is clearly not possible to determine the partition function in a large system. We need to find a method that produces states according to the Boltzmann probabilities without needing to compute the partition function. To do this we introduce a *Markov chain*. This is a sequence of states that shall make up our sample.

A Markov chain begins in a state μ and uses a transition probability $P_{\mu \rightarrow \nu}$ which is dependent on P_μ and P_ν to change the system from state μ to state ν . $P_{\mu \rightarrow \nu}$ is chosen such that the convergence to equilibrium¹ is quick. Hence, proper choice of $P_{\mu \rightarrow \nu}$ is one of the most important considerations in the Markov process and is one of the defining features of an algorithm that employs it.

The transition from state μ to ν must obey the following criteria:

¹We will consider the system to have reached equilibrium when the observables of the system converge the region of their expectation values and henceforth fluctuate finitely

1. It is independent of time.
2. It is memoryless: $P_{\mu \rightarrow \nu}$ does not depend on the path taken to reach state μ
3. The relation

$$\sum_{\nu} P_{\mu \rightarrow \nu} = 1$$

holds. Where $\nu = \mu$ is allowed.

4. For $\lim_{N \rightarrow \infty}$ the sample μ_1, \dots, μ_N follows the $P_{\mu} = \frac{e^{-\beta E_{\mu}}}{Z}$ distribution.

Since each state (ν) in the sample is reached from another state (μ) it is a requirement that the process is ergodic². That is, any state ν can be reached by any other state μ in a finite number of steps.

It is possible that we start our system in a state μ_1 that is far from equilibrium. In this case it is clear that several of our initial measurements shall not be close to the expectation value as it will take many steps along the chain. In this case we must consider a criteria for discarding false results and deciding when we have reached states which are typical of the simulated temperature. This is the problem of thermalization and shall be considered in section 3.3.

3.1.3 Detailed Balance

Finally we need to determine $P(\mu \rightarrow \nu)$ for our process such that the system will reach equilibrium as quickly as possible and only fluctuate from it by small amounts. If the system is in equilibrium, the rate of transition into and out of any state μ should be equal:

$$\sum_{\nu} p_{\mu} P(\mu \rightarrow \nu) = \sum_{\mu} p_{\nu} P(\nu \rightarrow \mu)$$

. This may be guaranteed by requiring that the rate of transition from state μ to state ν is the same as the rate of transition from ν to μ .

$$p_{\mu} P(\mu \rightarrow \nu) = p_{\nu} P(\nu \rightarrow \mu) \tag{3.4}$$

This condition is known as detailed balance.

²The system has the same behaviour averaged over time as it does averaged over phase space

Taking the sum over ν on both sides of equation (3.4) above we have:

$$p_\mu = \sum_\nu p_\nu P(\nu \rightarrow \mu) \quad (3.5)$$

Which we can consider this in matrix notation as:

$$\vec{p} = \vec{p} \cdot P \quad (3.6)$$

where $p = [p_{\mu_1}, \dots, p_{\mu_{2N}}]$ is a probability vector, with number of sites N , and P is the Markov Matrix given by:

$$P = \begin{pmatrix} p(\mu_1 \rightarrow \mu_1) & \dots & p(\mu_1 \rightarrow \mu_2^N) \\ \vdots & \dots & \vdots \\ p(\mu_N \rightarrow \mu_1) & \dots & p(\mu_N \rightarrow \mu_2^N) \end{pmatrix} \quad (3.7)$$

where we must have an eigenvalue of 1 to satisfy equation (3.6).

The Perron Frobenius theorem [6] ensures that every stochastic matrix has a vector \vec{p} that satisfies (3.6) and that the largest absolute value of any eigenvalue is always 1. For a matrix with strictly positive entries, which is guaranteed by the fact that the entries are probabilities, this vector is unique and each of the other $2^N - 1$ eigenvalues of P are < 1 .

Thus we must have that

$$\vec{p} = \vec{v}(\lim_{k \rightarrow \infty} P^k) \quad (3.8)$$

Which guarantees our desired probability distribution³ given that all eigenvalues of P will fall off to zero while our unique eigenvalue of 1 that satisfied equation (3.6) remains. It is also clear from this that we approach the desired distribution asymptotically with each step of our algorithm as each eigenvalue falls off with the power k .

³This is the Boltzmann distribution 2.1 discussed previously.

We have showed that detailed balance is a sufficient condition to ensure that the asymptotic fixed point distribution is the desired distribution. Thus from equation (2.1), the detailed balance condition for Boltzmann distribution is:

$$\frac{P(\mu \rightarrow \nu)}{P(\nu \rightarrow \mu)} = \frac{p_\nu}{p_\mu} = \frac{\frac{1}{Z}e^{-\beta E_\nu}}{\frac{1}{Z}e^{-\beta E_\mu}} = e^{-\beta(E_\nu - E_\mu)} \quad (3.9)$$

Where we can express $P(\mu \rightarrow \nu)$ as a product of two terms, $P(\mu \rightarrow \nu) = g(\mu \rightarrow \nu)A(\mu \rightarrow \nu)$. Now $g(\mu \rightarrow \nu)$ is the probability of selecting the transition from state μ to ν and $A(\mu \rightarrow \nu)$ is the probability of accepting that transition. Thus we have our final expression of detailed balance:

$$\frac{g(\mu \rightarrow \nu)A(\mu \rightarrow \nu)}{g(\nu \rightarrow \mu)A(\nu \rightarrow \mu)} = e^{-\beta(E_\nu - E_\mu)} \quad (3.10)$$

The reason that we choose to express it this way is so that we may decide any selection process we desire for long we fix the acceptance probability such that equation (3.10) holds. We change the acceptance ratio such that the maximum number of selected states are accepted.

3.2 Algorithms

We use two algorithms that greatly differ in their approach to a statistical solution to the Ising Model. The first is the Metropolis algorithm. Which is well documented over the last fifty years and is most frequently used in Monte Carlo simulations [1]. It is based on a local change procedure. The second is the Wolff Cluster algorithm [7]. This is a non-local clustering algorithm which was especially designed for spin systems and has efficiencies which greatly exceeds Metropolis in the critical region.

3.2.1 Metropolis Algorithm

As discussed in section 3.1.3 if equation (3.10) is satisfied for each change of state then the sample will converge to equation (2.1). We also discussed that we want the system to change often enough so that we reach equilibrium as quickly as possible. Thus we wish to maximise $P(\mu \rightarrow \nu)$, our aim shall be to make it of order one. This will be the case if the differences in energy $\Delta E = E_\nu - E_\mu$ isn't too large. The way we accomplish this in Metropolis algorithm is by only considering single spin changes. $s_{i,j} = \pm 1 \rightarrow s_{i,j} = \mp 1$

If we consider equation (2.13) we see that change of a single spin results in the change of sign for four terms $s_{i,j}s_{i,j+1}$, $s_{i,j}s_{i+1,j}$, $s_{i-1,j}s_{i,j}$, $s_{i,j-1}s_{i,j}$ in the sum. Thus the change in the energy for each bond connected to $s_{i,j}$ is ± 2 ie. from -1 to $+1$ or vice versa. In two dimensions we have four nearest neighbours and thus the maximum change in energy for the single spin flip is ± 8 . Or, more generally if the model has z nearest neighbours then

$$|\Delta E| \leq 2z \Leftrightarrow E_\mu \leq E_\nu - 2z \quad (3.11)$$

The selection probability $g(\mu \rightarrow \nu)$ may be chosen arbitrarily as discussed in the previous section and so we decide to choose it systematically; we'll start with the first site in the lattice and then go to the second, and so on. We can clearly reach any possible state of the system by this process thus it satisfies ergodicity. We now have

$$\frac{A(\mu \rightarrow \nu)}{A(\nu \rightarrow \mu)} = e^{-\beta(E_\nu - E_\mu)} \quad (3.12)$$

Where we set $A(\nu \rightarrow \mu)$ to 1, as discussed and use the algorithm proposed by Metropolis:

$$A(\mu \rightarrow \nu) = \begin{cases} e^{-\beta(E_\nu - E_\mu)} & E_\nu - E_\mu > 0 \\ 1 & E_\nu - E_\mu \leq 0 \end{cases} \quad (3.13)$$

So if the change of state lowers the energy then we always accept the new state. To implement this acceptance ratio in our algorithm we first choose a spin site; say, the first in our lattice. Then we calculate the change in energy that would result from flipping the spin. If this satisfies the Metropolis condition (3.13) then the new state is accepted. This process is repeated for each site as we 'sweep' through the lattice. A large number of sweeps through the lattice are made so that equilibrium is reached and so that we have many states in our sample to measure quantities like energy and magnetisation. Each 'sweep' of the lattice is taken to be a single 'timestep' of the Metropolis algorithm. This is not always the case but is the convention used in this paper.

Currently, our Metropolis condition (3.13) requires us to calculate the energy of the system for each state it is in, or which it selects as a possible transition. This is very costly since a calculation of the energy involves every site in the lattice. We are considering an entire lattice sweep as a single timestep and thus are calculating energy almost L^2 more

times than we'd like to. However, since spin flips and the associated change in energy is local we can consider this more simply.

Assume that the system is in state μ and considering the transition to state ν , which differs by value of one spin $s_k^\nu = -s_k^\mu$, $s_j^\nu = s_j^\mu \forall j \neq k$. Then

$$\begin{aligned}
E_\nu - E_\mu &= \left(-\sum_{\langle ij \rangle} s_i^\nu s_j^\nu\right) - \left(-\sum_{\langle ij \rangle} s_i^\mu s_j^\mu\right) \\
&= -\sum_{\langle ik \rangle} s_i^\mu (s_k^\nu - s_k^\mu) \\
&= 2\sum_{\langle ik \rangle} s_i^\mu s_k^\mu \\
\Delta E &= 2s_k^\mu \left(\sum_{\langle ik \rangle} s_i^\mu\right) \tag{3.14}
\end{aligned}$$

Now we only need to calculate the sum over the nearest neighbours to determine the result of the Metropolis condition and hence we have cut computation time considerably.

At the end of each timestep we measure observables magnetisation and energy using equations (2.12) and (2.13). Once the simulation has ran over a sufficient number of timesteps, the criteria for which will be determined by thermalisation and correlation time discussed later, we then estimate expectation values for these observables using equation (3.3). The error for which scales as $1/\sqrt{N}$ for sample size N , as will be discussed later.

Now energy per bond is

$$\langle e \rangle = \frac{1}{N_l} \langle E \rangle = \frac{1}{2N} \langle E \rangle$$

(from equation (2.11)) where $-1 \leq e \leq 1$. We also have that the magnetization per site is

$$\langle m \rangle = \frac{1}{N} \langle M \rangle$$

where $0 \leq m \leq 1$ and $\beta = 0 \implies$ a perfect disorder ($m = 0$) and the ground state is at $\beta = \infty$, ($m = 1$).

The magnetic susceptibility and heat capacity can finally be derived using equations (2.7) and (2.8) where the constant k is now the number of sites N - not to be confused with the number of elements in the sample. Finally, β may be changed and analysis of of observables at the new temperature may commence.

3.2.2 Cluster Algorithm

The cluster algorithm [7], is a non-local spin flip algorithm that changes the system from state μ to state ν by flipping a group, or cluster, of like spins. A single site is chosen from the lattice at random with probability $1/N$. This site is called the seed (s_{seed}). We then look at the nearest neighbours; if they have the same spin as the seed, they are added to the cluster with the probability P_{add} . This process is repeated for each new member of the cluster until there are no new members, at which point all the elements of the cluster are flipped. Sites may be considered more than once for addition to the cluster if they appear as neighbours to more than one site in the cluster.

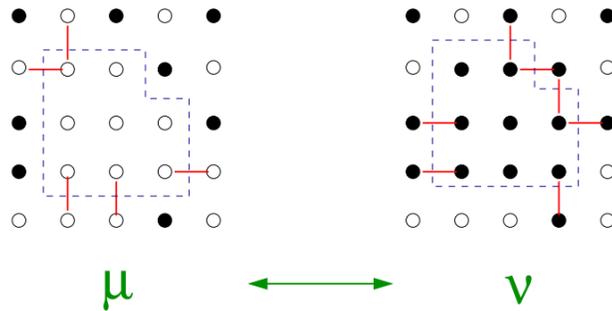


Figure 3.1: Two spin configurations that differ by the flip of a Wolff cluster. The bonds that are destroyed/created during the transition belong to the boundary of the cluster. Figure from Computational Physics [1].

Now, we are left with the task of determining the value of P_{add} that satisfies detailed balance. Similarly to our plan in the Metropolis algorithm, we want to determine if the resulting state ν decreases the energy. So we want to add a member to the cluster if it is likely that it will decrease the energy of the system to have that spin flipped. From figure 3.1 we can see that the bonds between mutual elements of the cluster do not change sign, as they are flipped together. The links whose signs change (ie. bonds that are broken or created by the flip) are highlighted in this figure. So it is easy to see that the change in energy is determined by the situation at the boundary of the cluster. Therefore, when we break a bond we incur an energy difference of $+2J$, and similarly creating a bond incurs a difference of $-2J$.

We can also see, from figure 3.1, that any bonds that are broken/created in the change from state $\mu \rightarrow \nu$ is created/broken in the change from $\nu \rightarrow \mu$.

Suppose that the number of bonds broken in going from $\mu \rightarrow \nu$ is r , and the number of bonds broken in the transition $\nu \rightarrow \mu$ is s . Then

$$E_\nu - E_\mu = 2J(r - s)$$

. Now equation (3.10) becomes

$$\frac{g(\mu \rightarrow \nu)A(\mu \rightarrow \nu)}{g(\nu \rightarrow \mu)A(\nu \rightarrow \mu)} = e^{-2J\beta(r-s)} \quad (3.15)$$

So the probability of 'not' adding a site at the boundary is $1 - P_{add}$. Thus $g(\mu \rightarrow \nu) = (1 - P_{add})^r$ and $g(\nu \rightarrow \mu) = (1 - P_{add})^s$. Putting these two equations into (3.15) we have

$$(1 - P_{add})^{r-s} \frac{A(\mu \rightarrow \nu)}{A(\nu \rightarrow \mu)} = e^{-2J\beta(r-s)} \quad (3.16)$$

which implies

$$\frac{A(\mu \rightarrow \nu)}{A(\nu \rightarrow \mu)} = (e^{-2J\beta}(1 - P_{add}))^{(s-r)} \quad (3.17)$$

We now choose P_{add} such that equation (3.17) is equal to 1. This means that we can take the acceptance ratios $A(\mu \rightarrow \nu)$ and $A(\nu \rightarrow \mu)$ also to be one, so that once the cluster has been constructed, the new configuration (state ν) will always be accepted. Hence we never build a cluster that we don't flip and our algorithm shall be as efficient as possible. Thus we use

$$P_{add} = 1 - e^{-2\beta J} \quad (3.18)$$

There is a non-zero probability, and actually typical of high temperatures, that the cluster shall only consist of the seed site. Since the seed site is chosen at random, and any site may be chosen, any state of the system can be obtained in a finite number of steps. Therefore the Wolff Cluster algorithm satisfies ergodicity.

Observables are measured in the same way as using the Metropolis algorithm, however now a timestep is taken to be a single cluster flip.

3.3 Thermalization

When a system is in thermal equilibrium with a reservoir of a given temperature, a typical state has energy that differs very little from its average value and belongs to a restricted region of phase space. Thus when we choose an initial state which is far from this region then the system must perform a thermalising procedure through a series of states until it finds the region of typical states. This process in Monte Carlo simulation is known as thermalization.

This gives rise to two problems; choosing an appropriate initial configuration, ie. hot or cold, that will reduce the length of this thermalisation, and finding a criteria to determine when the system is thermalised so that we can start measuring observables. The difference that a choice of initial configuration can make is demonstrated in figure (3.2). In addition to hot and cold we may also use an old state start. This is usually done when we are calculating observables across a large heat spectrum. A small number of sweeps are carried out on a particular temperature, then the temperature is changed slightly and the final configuration at the last step is used as the first configuration for the next. This reduces thermalization time as for the most part the difference in expectation values of observables when we change the temperature by small amounts are small.

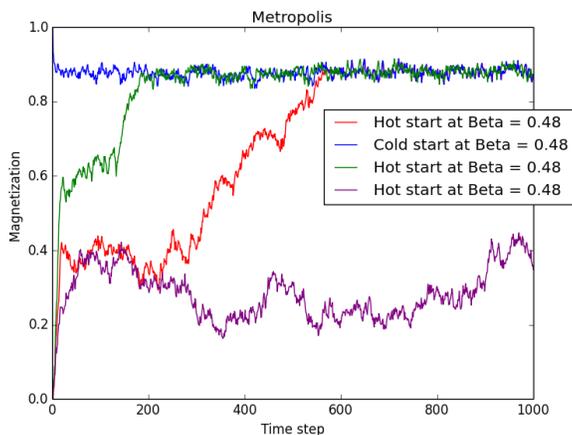
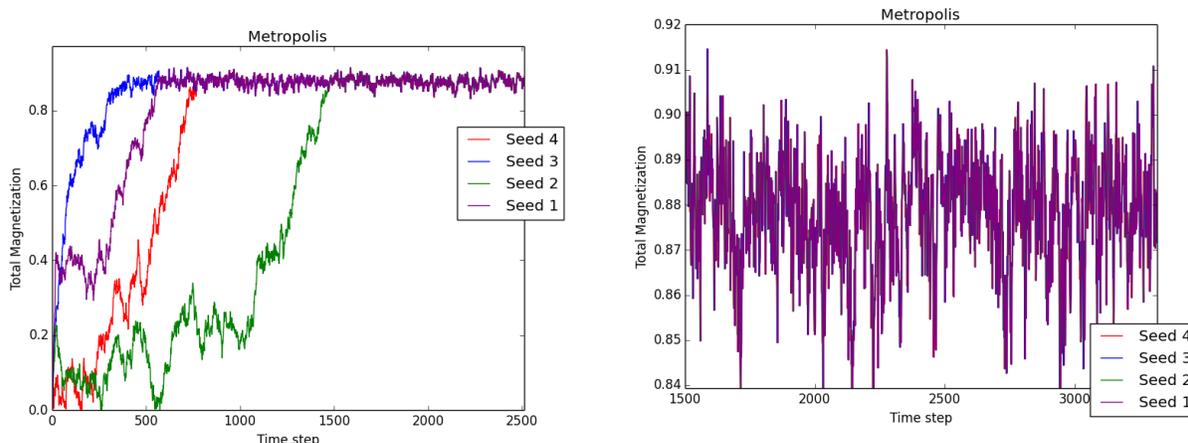


Figure 3.2: $\beta = 0.48$, Thermalization for three hot starts and one cold start in a cold region. Note how the thermalization time for the cold start is much less. The only change in each of the hot starts is the random number seed. One of the simulations has not even had enough time to thermalize.

From figure 3.2 we can also see that we have different thermalization times for different random number seeds, despite having the same initial start. This gives rise to a potential

criteria for testing for thermalization. In figure (3.3) we have the same initial start at the same temperature for four independent sets of random numbers. This time enough timesteps have been given for each to converge. This value of convergence is the systems magnetization for the given temperature. The expectation value of magnetisation is then given by equation (3.3).



(a) Simulations all converge after 1500 monte carlo steps. (b) Zoomed in on thermalized region. Observables magnetisation can be measured at this point.

Figure 3.3: Thermalization for four hot starts cold region at $\beta = 0.48$, $L=100$. When the timestep histories of the monitored quantities converge, we are confident that the system has been thermalized.

3.4 Autocorrelations and Correlations

So far we have been assuming that the measurements we have taken have come from independent states. However, since the algorithms change a finite number of spins at each iteration, the states used at each timestep may be statistically correlated. For example, in the Metropolis algorithm we have defined one timestep to be a sweep of the entire lattice. Thus we may expect that we have reached new state that is in no way correlated to the previous state. This is true for certain temperatures, however when we are close to a phase transition, in the *critical region*, the behaviour is not so trivial. Consider figure 3.4a, which shows a spin configuration for the $L=100$ Ising model, where black dots are spin down sites and white are spin up sites. It appears to have a random (disordered) nature. A local change algorithm such as Metropolis has no problem changing individual spins in this case as the

nearest neighbours of each individual spin are, in general, a even mix of 'up' and 'down' spins. Thus, according to equation (3.14) the probability of flipping will likely be of order 1. However, as we can see from figure 3.4b, the neighbours of any given spin are likely to have four neighbours with like spins and thus we have a small ΔE which tends to make the probability of flipping very large. This causes the Metropolis algorithm to mostly change the spins at the boundary of the clusters (in figure 3.4b). It should now be clear that an entire sweep of the lattice is not enough to obtain an independent spin configuration.

In the way we have just described it can be understood how spins that are quite far apart in the lattice may still have mutual influence over one another, this is seen to be significant at close distances, as in figure 3.4a, or at large distances, as in figure 3.4b. This introduces a new variable that is known as the correlation length ξ .



(a) $\beta = 0.2$ Correlation lengths are small and it is easy to change individual spins.



(b) $\beta = 0.43$ Like spins have clustered together as the correlation length diverges.

Figure 3.4: Spin configurations for non-critical and critical regions. $L=100$.

To quantify the magnitude of correlations between configurations separated by t timesteps we construct the *autocorrelation function*. If we consider an observable O , and let $O(t)$ be it's value after t Monte Carlo timesteps. Then the autocorrelation function $p_O(t)$ of O is

$$p_O(t) = \frac{\langle (O(t') - \langle O \rangle)(O(t' + t) - \langle O \rangle) \rangle_{t'}}{\langle (O - \langle O \rangle)^2 \rangle} \quad (3.19)$$

Where $\langle \dots \rangle'_t$ is the average value over the configurations in the sample ie. for each different t' with $t' < t_{max} - t$. We clearly have $p_O(0) = 1$ which is required as a configuration is

clearly fully correlated itself. When $p_O(t)$ fluctuates about zero, we may consider that we have reached a statistically independent configuration.

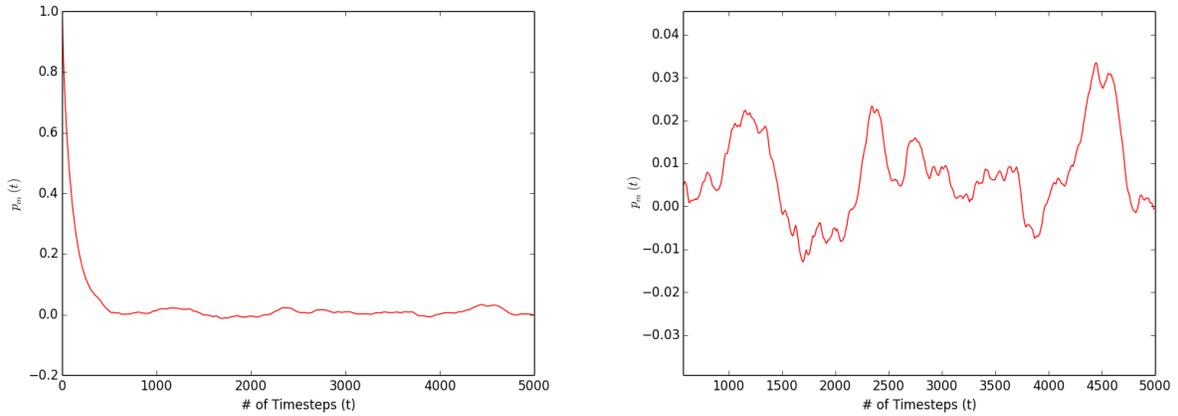
The autocorrelation function should be the same, within statistical error, for a any run using the same temperature and observable⁴. $p_O(t)$ drops exponentially:

$$p_O(t) \approx e^{-t/\tau_O} \quad (3.20)$$

Where τ_O is known as the *autocorrelation time*, and is also a constant for an observable at a particular temperaure. equation (3.20) is not exact because we have only accounted for the drop off between 1 and the next highest eigenvalue of the Markov Matrix. We generally consider that we have an independent measurement when $t = 2\tau_O$ ie. when

$$p_O(t) = e^{-2} \approx 0.14$$

Therefore, if t_{max} is the number of timesteps taken then we have $t_{max}/2\tau_O$ statistically independent measurements.



(a) Exponential decay according to equation (3.20). (b) Zoomed in region. We can see fluctuations, averaging at zero.

Figure 3.5: Autocorrelation Function L=50. For magnetization with the Metropolis algorithm in the non-critical region.

⁴ $p_O(t)$ can be quite different for different O or β and lattice sizes L

An accurate measurement of τ_O is not easy as it requires $t_{max} \gg \tau_O$. For measurements that give us a very large autocorrelation time, we try instead to take measurements every τ_O steps. To avoid being swamped by data.

Beating autocorrelations is crucial in Monte Carlo simulations since they are the main obstacle for studying large systems, which in turn is essential for taking the thermodynamic limit without the systematic errors introduced by finite size effects.

We may also measure the correlation between two points on the lattice. That is, how two points s_i and s_j vary together. This can be simply calculated by the product of the variance of these two sites, known as the *correlation function*:

$$g_{ij} = \langle (\sigma_i - \langle \sigma_i \rangle)(\sigma_j - \langle \sigma_j \rangle) \rangle \quad (3.21)$$

Similarly to the autocorrelation function, this has an exponential decay given by

$$g_{ij} = e^{-|i-j|/\xi} \quad (3.22)$$

which we can see falls off as we examine points further apart, scaled by the value of the *correlation length*: $\xi = \xi(\beta, L)$.

3.5 Critical Exponents

We must introduce a new variable known as the 'reduced temperature', t

$$t \equiv \frac{\beta_c - \beta}{\beta_c} \quad (3.23)$$

This approaches zero at the critical point ($\beta \rightarrow \beta_c$), and so is used to describe how observables behave as they approach the critical temperature.

$$\chi \sim t^{-\gamma} \quad (3.24)$$

$$C \sim t^{-\alpha} \quad (3.25)$$

$$\xi \sim t^{-\nu} \quad (3.26)$$

$$\tau \sim t^{-\nu z} \quad (3.27)$$

Note: τ has two critical exponents as convention so that it may be easier to relate τ and ξ .

These observables clearly diverge for exponents > 0 , however this may only be seen in the thermodynamic limit; as for a finite system, these observables are clearly analytic⁵ as they can be derived using the partition function. However, on a finite lattice these observables have a maximum⁶ in a *pseudocritical temperature* $\beta_c(L)$

$$\lim_{L \rightarrow \infty} \beta_c(L) = \beta_c \quad (3.28)$$

While $\xi \ll L$ the model should behave approximately as an infinite system. However, at critical temperature we have $\xi \sim L$, so finite size effect becomes significant. So, on a finite lattice we take the pseudocritical region, $\beta = \beta_c(L)$, and so we have that $\xi(t, L) \sim L$. This implies that in the region of $\beta_c(L)$:

$$t \sim L^{-1/\nu} \quad (3.29)$$

and thus equations (3.24 - 3.27) become

$$\chi \sim L^{\gamma/\nu} \quad (3.30)$$

$$C \sim L^{\alpha/\nu} \quad (3.31)$$

$$\xi \sim L \quad (3.32)$$

$$\tau \sim L^z \quad (3.33)$$

This is referred to as finite size scaling, and provides us with a convenient way of calculating critical exponents, as shall be discussed later.

3.6 Statistical Errors

The determination of statistical errors is of central importance in order to assess the quality of a measurement and predict the amount of resources needed for reaching a specific accuracy goal.

⁵The sum in equation (2.4) is as set of analytic functions, however impractical it may be to solve.

⁶They appear to diverge at the critical point but do have an analytic peak. This can be seen in figures 5.7 and 5.8.

3.6.1 Errors of Independent Measurements

Using the assumption that the source of statistical errors are the thermal fluctuations around the average value of an observable, we conclude that its expectation value can be estimated by the mean of the sample and its error by the error of the mean. Thus we calculate the mean of the fluctuations in the thermalised region: We take n samples in the thermalised region and thus our sample mean is given by:

$$\bar{O}_n = \frac{1}{n} \sum_{i=1}^n O_i \quad (3.34)$$

This is an unbiased⁷ estimator of the mean. Now the error of the unbiased estimator \bar{O}_n is an estimator of the statistical error which is given by:

$$\begin{aligned} \sigma_O^2 &= \frac{1}{n-1} \left\{ \frac{1}{n} \sum_{i=0}^{n-1} (O_i - \bar{O})^2 \right\} \\ \sigma_O^2 &= \frac{1}{n-1} (\langle O^2 \rangle - \bar{O}^2) \end{aligned} \quad (3.35)$$

Following our discussion in section 3.4 it is possible to see that each successive value O_i are not guaranteed to be statistically independent. If τ_O is the autocorrelation time, with unit of number of measurements \hat{n} , then according to section 3.4, we have

$$n_O = \frac{n}{2\tau_O} \quad (3.36)$$

independent measurements. For this reason the error given in equation (3.35) is referred to as *naïve error*.

The real error should be given by

$$\begin{aligned} \sigma_O^2 &= \frac{1 + 2\tau_O}{n-1} (\langle O^2 \rangle - \langle O \rangle^2) \\ &\approx \frac{1}{n/2\tau_O} (\langle O^2 \rangle - \langle O \rangle^2) \\ &\approx \frac{1}{n_O - 1} (\langle O^2 \rangle - \langle O \rangle^2) \end{aligned} \quad (3.37)$$

⁷It gives the correct mean for finite n and not only in the limit $n \rightarrow \infty$

where we have gotten the second and third lines by assuming $\tau_O \gg 1$ and using equation (3.36).

3.6.2 Blocking

A natural way to ensure that our measurements are statistically independent is to divide the n measurements into n_b number of bins and increase the size of these bins until the error levels off ie. when the number of elements in a bin b is $> 2\tau_O$ and $n_b = n_O$. This will ensure that each bin contains a statistically independent set of results. This process of error handling is referred to as the blocking method.

We define $O_{b,i}$ as the average value of observable O in bin i . Following equation (3.37) we can calculate error using:

$$\sigma_{O_b}^2 = \frac{1}{n_b - 1} (\langle O_b^2 \rangle - \langle O_b \rangle^2) \quad (3.38)$$

Where we can be confident that we have a correct representation of error if $b \geq 2\tau_O$. If b is too small then the error will be underestimated (of course $b = 1$ is the naïve error) then we shall underestimate the error. In this case the b should be increased until the error levels off, stops increasing, and remains at a constant value, then we have discovered the true statistical error for the sample n .

This also gives us a method of calculating errors for our derived observables. If we consider again the equations for calculating susceptibility and heat capacity from magnetisation (2.7) and energy (2.8), we see that a sample of the measured observables are required to get a measurement of these derived observables. Thus, using the blocking method we may calculate these derived observables for each bin of data, as if that bin were an entire sample. From this we have a new sample of n_b values of the derived observables from which we may calculate the error once again by using equation (3.35). Note that, we may consider each of these measurements as having come from sufficiently uncorrelated configurations.

Chapter 4

Exact and Statistical Calculations

As discussed, our metropolis and cluster algorithms give statistical approximations of the expectation value of an observable at a given temperature. These estimates have errors as explained in section 3.6, which can be beaten down by higher Monte Carlo steps. The error however remains non-zero. As discussed in section 2.1 we can calculate our expectation value by using equation (2.4). Since this requires us to examine each possible state computation time goes up as quickly as number of states, which for lattice size of $L = 100$ we have $2^{100 \times 100}$ states. Thus, far beyond numerical calculation. For a smaller lattice, this is possible.

4.1 An Exact Computational Approach.

In this section we calculate exact solutions for observables in the 3×3 and 4×4 models. The exact solution of the observables were calculated using equation (2.4) where it was necessary calculate the energy (and magnetization) in every possible state of the system. The corresponding number of states are: $2^{3 \times 3} = 512$ and $2^{4 \times 4} = 65\,536$. Thus, it was impractical to lattice sizes greater than this as a $L = 5$ already gives $2^{5 \times 5} = 33\,554\,432$ states!

A comparison for the exact solutions for spontaneous magnetization and energy with different lattice sizes is shown in figures 4.1a and 4.1b. To illustrate what these might look like closer to the thermodynamic limit ($N \rightarrow \infty$) we include a statistical solution at $L = 100$ from data generated by the Metropolis algorithm.

We can see that for smaller lattices there appears to be a 'smoothing out' of the critical region so that the change in magnetization and energy is less extreme. This implies that

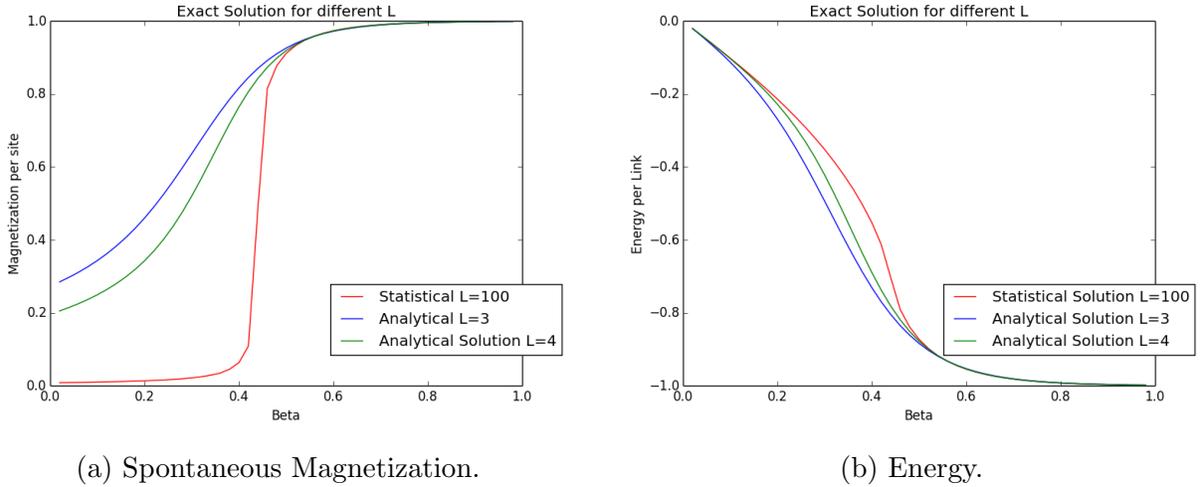


Figure 4.1: Exact solutions for different lattice sizes with statistical $L = 100$ for comparison to a large system.

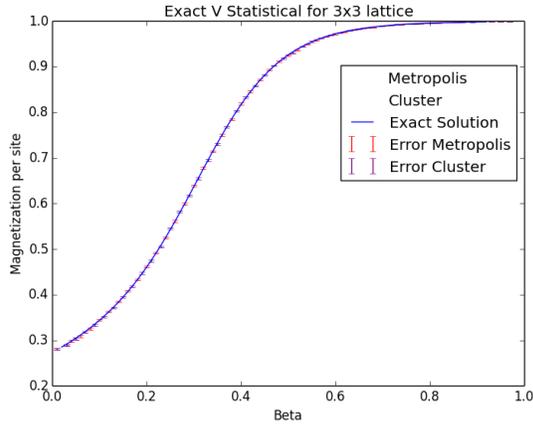
there would be a smaller divergence of the magnetic susceptibility and heat capacity in the critical region as they are the variance of magnetization and energy respectively. This is indeed shown to be the case and discussed further in section 5.3.

It is also noticable from figure 4.1a that in the cold region ($\beta > \beta_c$) lattice size has no significant effect and small magnetization acts very similarly to how it would in the thermodynamic limit.

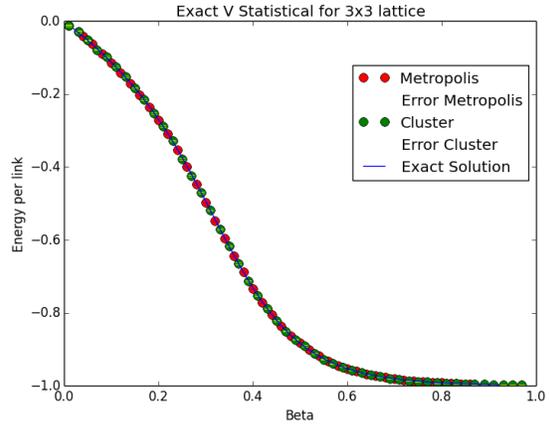
4.2 Testing the Algorithms

The analytical solution to the $L=3$ Ising model is shown by in figure 4.2. Expectation values for the magnetization (figure 4.2a) and energy (figure 4.2b) and their associated errors were calculated using methods discussed in section 3.6 for both the Cluster and Metropolis Algorithm.

An agreement to the exact solution for these observables within the range of the error bars can be seen giving convincing evidence that our algorithms and error data analysis give an accurate statistical description of the model at these lattice sizes.



(a) Spontaneous Magnetization $L = 3$.



(b) Energy $L = 3$.

Figure 4.2: In both figures the results calculated using each algorithm agree with the exact solution within statistical error as can be seen from the error bars. The points for the two algorithms are taken at different values of β so that they are both clearly visible on the graph.

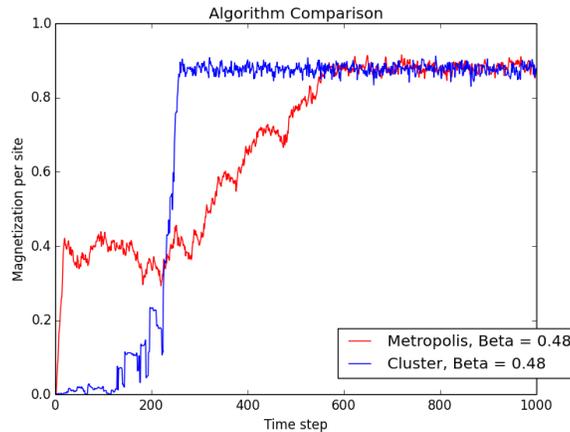


Figure 4.3: $\beta = 0.48$, $L = 100$ Algorithm comparison in cold region.

From figure 4.3 we see an agreement between algorithms once thermalised despite the very different paths taken to reach equilibrium. It is obvious, without the plot legend, which algorithm has created each result. Clearly we have a set of successive large 'jumps' from the cluster algorithm where the first large cluster of like spins has been formed.

Chapter 5

Simulations

5.1 Monte Carlo Histories

Computing Monte Carlo histories ie. the value of an observable at a particular temperature as a function of number of Monte Carlo steps, gives us a very good idea about thermalisation times. It is shown in figure 5.1 that for both algorithms longer thermalisation is required for larger lattice sizes.

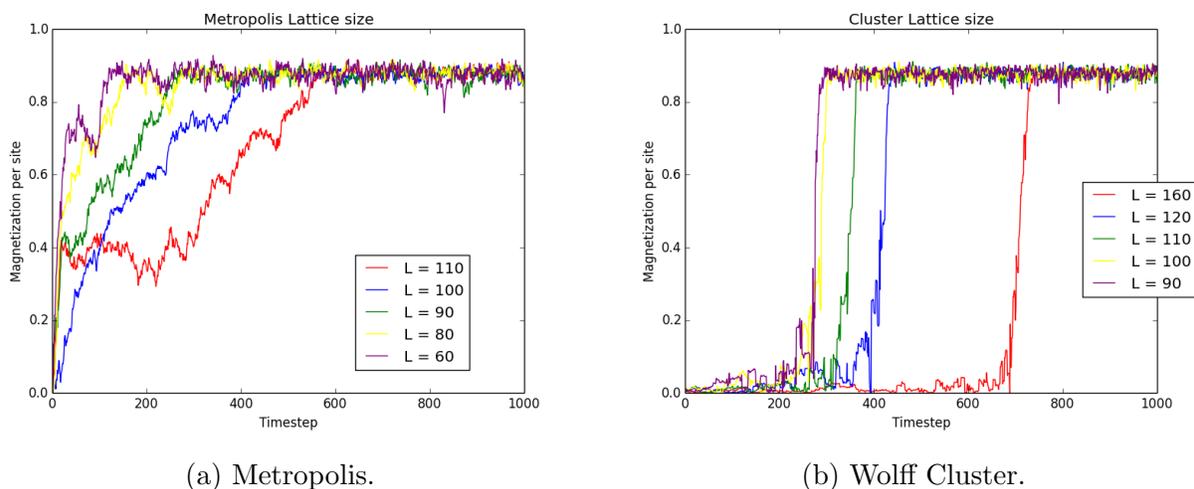


Figure 5.1: Magnetization per site for the Ising model for both algorithms with $\beta = 0.48$. Thermalization from a hot start takes longer for a large system.

In both cases the system undergoes a thermalisation process through a set of states until the observables of the system converge to a value about which they will fluctuate only slightly.

In figure 5.2a we notice that there is no major difference in thermalisation time when we compare different lattice sizes. This can be understood by recalling that we defined our Metropolis timestep as an entire sweep through the lattice. A disordered state has lower energy than an ordered state so every potential change, when all spins are aligned, will be accepted by equation (3.12). Thus the difference of lattice size is irrelevant. Hot region thermalisation in the Cluster algorithm, however, appears to depend on lattice size. This may also be understood if we assume that cluster sizes are small in this region. Which can be seen to be the case in figure 5.3. Since a timestep in this case would be flipping a single spin per step and hence for larger systems this will be very costly. We can now see the importance of clearly defining a Monte Carlo step.

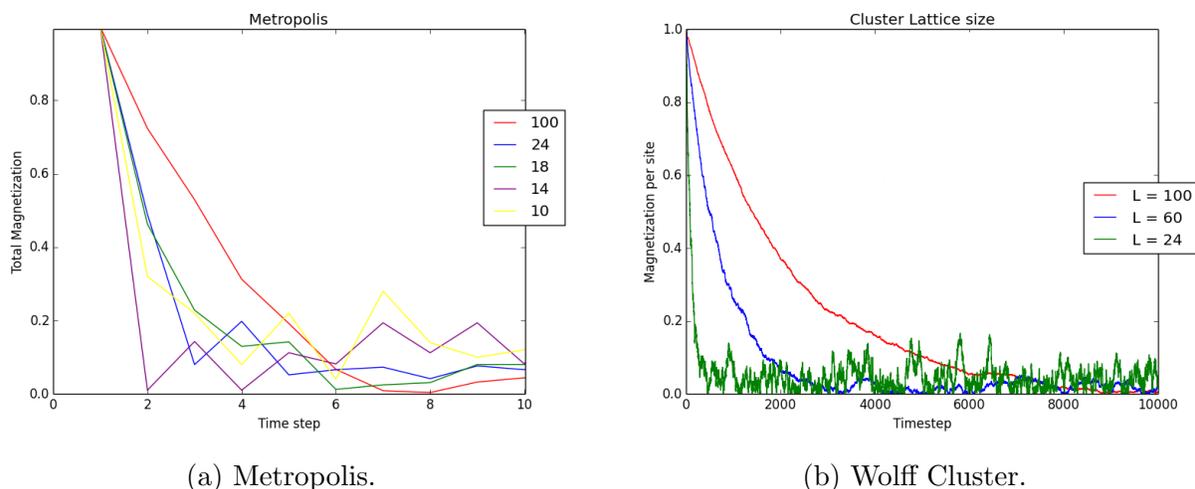


Figure 5.2: Magnetisation per site for the Ising model for both algorithms with $\beta = 0.1$. Thermalization from a cold start takes longer for a large system for the cluster algorithm given how we have defined our Monte Carlo steps and the fact that cluster don't grow large in the hot region.

We include at the end of this section a graph of the average fraction of spins flipped per Monte Carlo step for the Wolff Cluster and Metropolis algorithms as a function of temperature as it gives a good graphical indication of the discussion above. This was simulated by a 10^2 lattice for 10,000 Monte Carlo steps in each case.

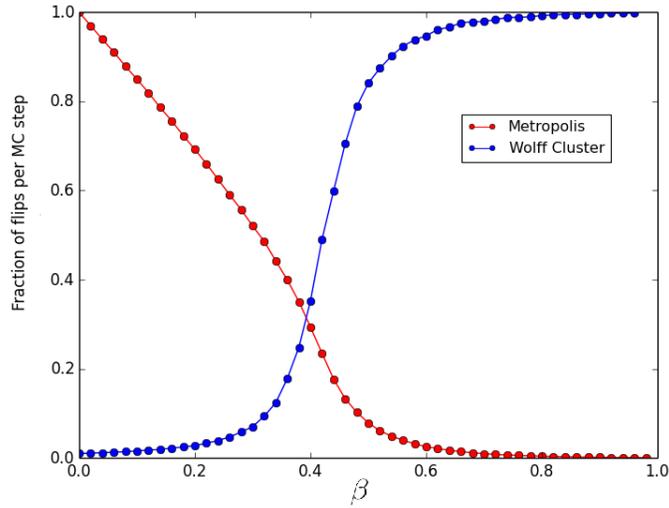


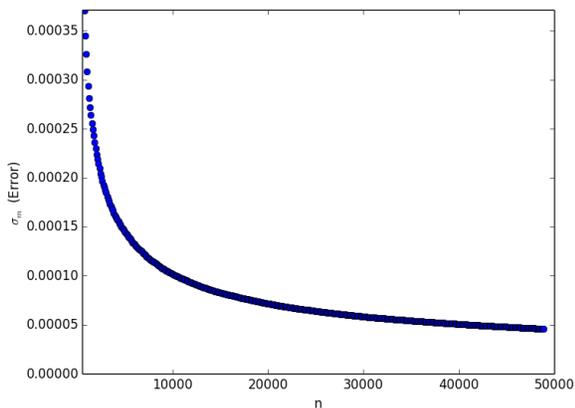
Figure 5.3: The plot shows the average cluster size for the Wolff Cluster algorithm and the number of sites flipped per sweep for the Metropolis algorithm as a fraction of the lattice size N for the range of temperatures shown.

5.2 Sample Size

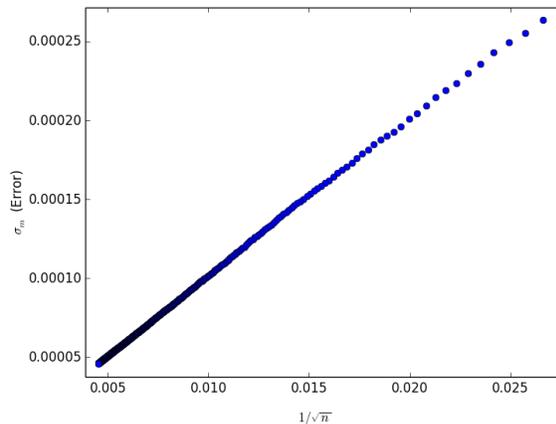
Considering equation (3.35) we may think that error will fall off as we increase the size of the sample, in a way close to $1/\sqrt{n-1}$. We consider that this is a rough approximation as the $(\langle O^2 \rangle - \langle O \rangle^2)$ should only be changing slightly, for each sample size. It can be shown that

$$\sigma_O \propto 1/\sqrt{n} \tag{5.1}$$

This is illustrated in figure 5.4.



(a) Error Versus sample size n .



(b) Error is proportional to $1/\sqrt{n}$.

Figure 5.4: Relationship between error and sample size. $L = 100$. $\beta = 0.1$ Metropolis algorithm.

This relationship gives us a way of choosing our accuracy by creating a sufficiently large sample n according to (5.1).

5.3 Observables and their critical exponents

We simulated the Ising model using the Metropolis algorithm on lattices of size $N = 10^2$, $N = 20^2$, $N = 50^2$ and $N = 100^2$, at a range of temperatures $0 < \beta < 1$. Each simulation was began in a hot (random) state with 50,000 Monte Carlo steps. For the hot region $\beta < \beta_c (\approx 0.44)$ this start was very close to equilibrium behaviour of the system, it was found that thermalisation for this region was generally less than 1,000 sweeps and so we could safely expect the system to be thermalised after 3,000 steps. The number of sweeps required elsewhere (in the pseudocritical region and above) was found to be 30,000 to have safely thermalised. This difference in thermalisation times is easily seen looking back at figure 3.2.

In the thermodynamic limit, the 2-dimensional Ising Model undergoes a second order phase transition¹. However, for a finite lattice size the free energy, and it's derivatives, must always be analytic, since the free energy is $F = F(Z)$, where the partition function for a finite lattice is a finite sum of exponentials. Thus the 'phase transition' that we might appear to see will always be continuous for a finite system. This is seen to be the case in figure 5.5

¹Has a discontinuity in the first derivative of the order parameter, generally taken to be the free energy.

which shows spontaneous magnetisation per spin for different lattice sizes. Notice how the 'phase transition' is more gradual for smaller lattices.

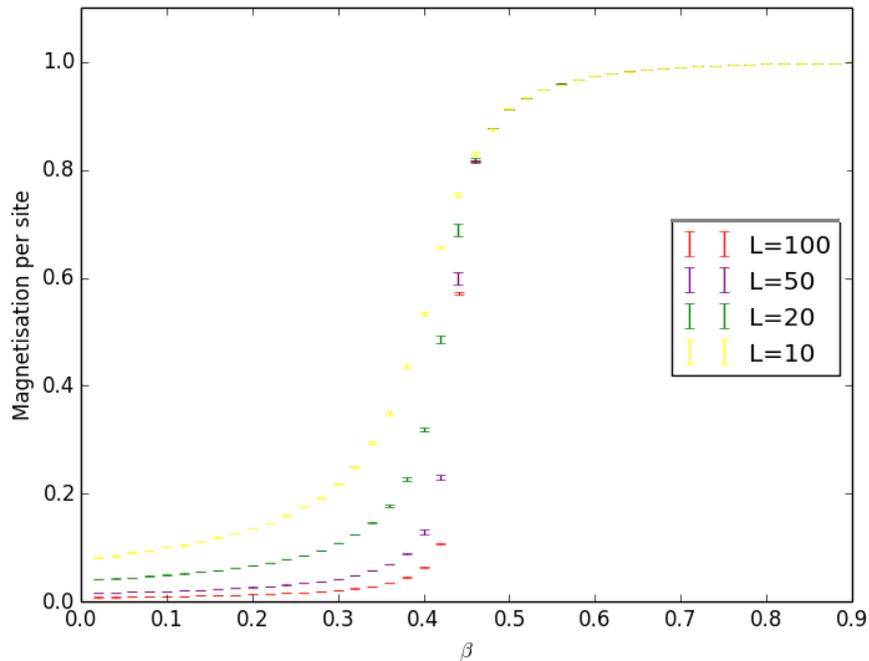


Figure 5.5: Ising Model spontaneous magnetisation per spin for lattice sizes $N = 10^2$, $N = 20^2$, $N = 50^2$ and $N = 100^2$.

As we previously discussed we take the absolute value of magnetisation, which implies that we will always slightly overestimate it's value in the hot region. This is a source of systematic error that must be noted however may not be relevant for calculating values of interest as we are generally more interested in the critical region.

The blocking method, described in section 3.6.2, was used to accurately determine the statistical error seen in figure 5.5. Derived observables, magnetic susceptibility and heat capacity were also calculated using this method. For most of the range ($\beta < 0.4$, $\beta > 0.5$) it was possible to use small bin sizes to overcome autocorrelation time. We can see from figure 5.6 that for $\beta = 0.4$ a bin size of 50 was sufficient to consider each bin to be a statistically independent measurement. This was therefore taken as the bin size for the regions above. In the critical region however, a bin size of 10,000 was required to give accurate measurements of error and the derived observables. For this reason a much larger sample size for these

temperatures was used. The highest of these was 1,000,000 for the sample sitting closest to the critical point ($\beta = 0.44$), this also helped to beat down error in this region. This is reflected in figure 5.5. However, being required to perform 1,000,000 timesteps at a single temperature is extremely costly and is the major problem caused by critical slowing down seen very significantly in the Metropolis algorithm.

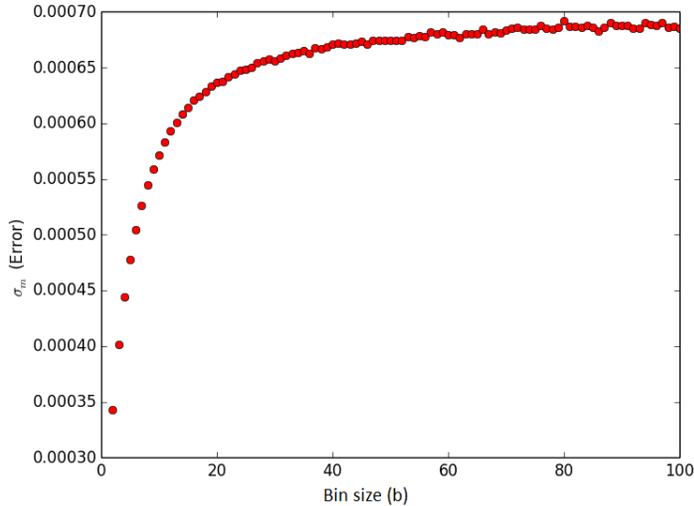


Figure 5.6: $\beta = 0.4$. $L = 100$. Metropolis. Error for increasing bin size.

Smaller bin sizes were required for the smaller lattices in both the critical and non-critical regions. In general the required bin size at each temperature was determined by systematically increasing bin size while subtracting the error found using the previous bin size from the error using the current bin size. When this converged sufficiently close to zero the bin size was accepted and that value of error was recorded. Deciding when the difference in the error was sufficiently close to zero is a source of systematic error but recall that we generally consider states that have a correlation of less than 0.14 to be de-correlated so we are not providing any additional source of statistical error by this approach.

Overcoming the effect of varying autocorrelation times was essential for deriving magnetic susceptibility and heat capacity. As these are the variance of magnetism and energy respectively, using a bin size that was too small ie. ($b < 2\tau_O$) not only underestimates the error in the derived observables, but far more significantly underestimates the values of the derived observables themselves. This problem was largely significant at the critical temperature. Thus the correct bin size was required to satisfy $b \geq 2\tau_m, 2\tau_e, 2\tau_\chi, 2\tau_C$.

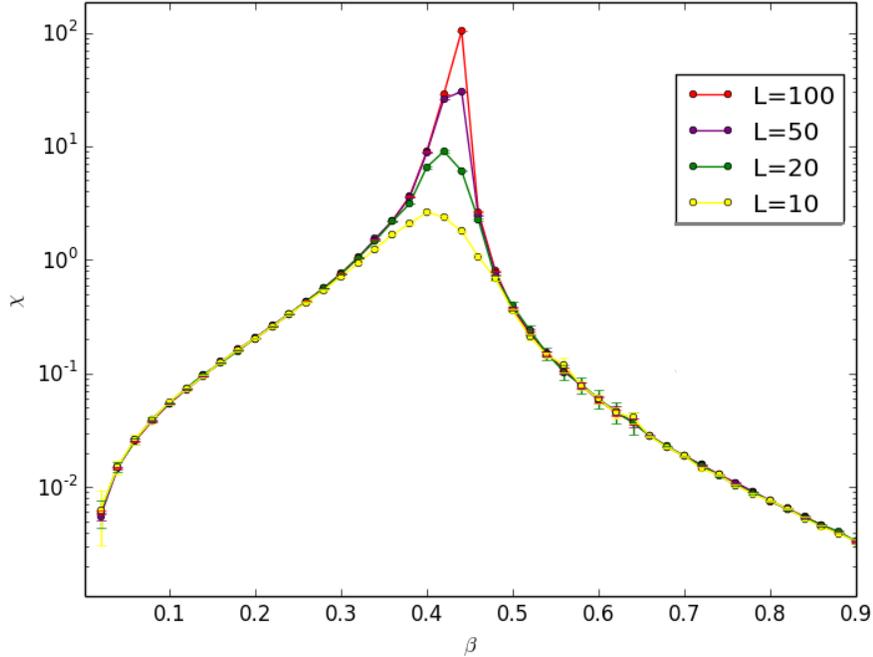


Figure 5.7: Magnetic Susceptibility of the Ising Model for lattice sizes $N = 10^2$, $N = 20^2$, $N = 50^2$ and $N = 100^2$. Where they are not visible, the error bars are smaller than the points shown.

The magnetic susceptibility, χ , is shown to be almost independent of lattice size when β takes on values away from the critical region. This is also seen to be the case for heat capacity shown in figure 5.8. Their values in this region are increasing according to equations (3.30) and (3.31). We take $\nu = 1$ as it has been found to be in literature. Thus we may estimate γ and α in the following way: from equation (3.30) we deduce

$$\frac{\chi_{L'}}{\chi_L} \cong a^\gamma \qquad \frac{C_{L'}}{C_L} \cong a^\alpha \qquad (5.2)$$

where $L' = L \cdot a$. Before preceeding to implement equation (5.2), we re-calculated the values of χ and C at the critical values using 200,000 steps of the Wolf Cluster algorithm. Where these values agreed with those of the Metropolis within the calculated error, we used the their average value to calculate the critical exponents. The agreed upon value of χ and C at the critical temperatures are shown in table 5.1.

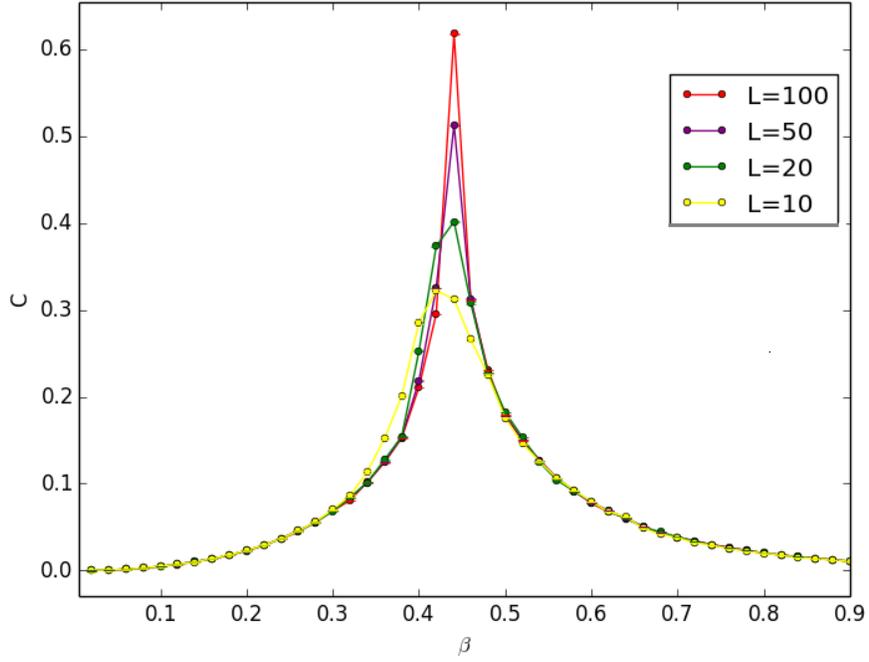


Figure 5.8: The famous divergence of heat capacity for lattice sizes $N = 10^2$, $N = 20^2$, $N = 50^2$ and $N = 100^2$. Where they are not visible, the error bars are smaller than the points shown.

Table 5.1: χ and C in the pseudo-critical region.

L	χ	C
200	349.40 ± 0.081	0.769 ± 0.0057
100	103.89 ± 0.041	0.625 ± 0.0048
50	30.306 ± 0.040	0.527 ± 0.0080
20	6.03 ± 0.013	0.406 ± 0.005
10	1.78 ± 0.014	0.318 ± 0.0090

From these values, ten estimates of each γ and α were found by using equations (5.2) to relate the results from each of the simulated lattice sizes. The average of these values were accepted and the error was calculated from the square root of the variance of those ten values:

$$\gamma = 1.76 \pm 0.023 \tag{5.3}$$

$$\alpha = 0.29 \pm 1.5 \tag{5.4}$$

It must be noted that the errors in the above equations account for the statistical fluctuations of the system, they are not systematic. Thus it is possible for the true value of these constants to lie outside the error range. This is indeed the case for α as the generally accepted value $\alpha = 0$. This is because

We notice that lattice size has a much larger effect on magnetic susceptibility than it does on heat capacity at critical temperature. This is reasonable when we recall that these are the variance of magnetism and energy respectively. As we can see from the figure below (5.9), and from figure 5.7, we have a much greater variance of magnetism due to different lattice sizes than for energy where the difference is barely visible.

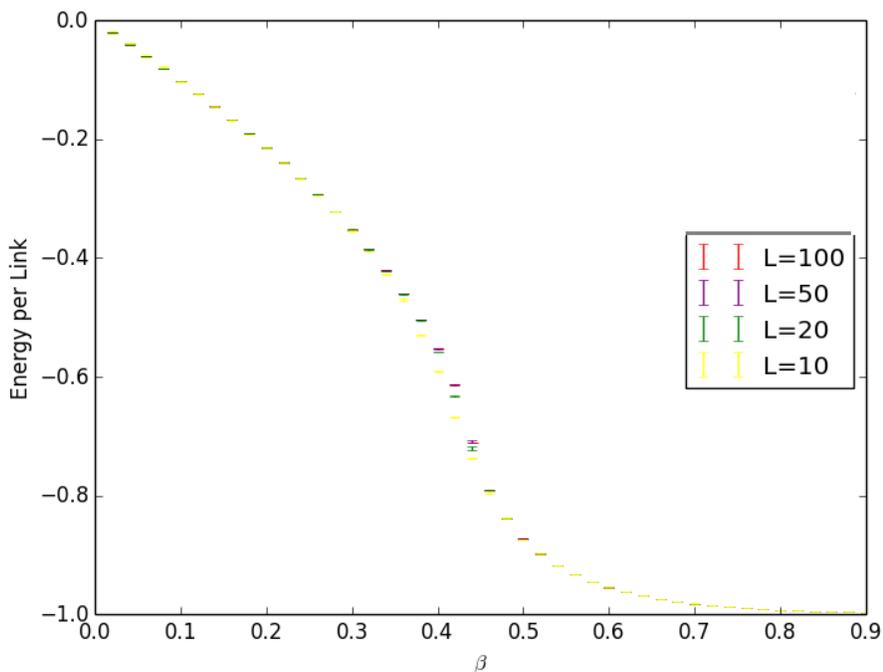


Figure 5.9: Ising Model energy per link for lattice sizes $N = 10^2$, $N = 20^2$, $N = 50^2$ and $N = 100^2$. We may recall that the $L=3$ and $L=4$ lattice in figure 4.1b were only slightly different to this.

5.4 Critical Slowing Down

The z in equations (3.27) and (3.33) is referred to as the dynamic critical exponent and is a property² of the algorithm which is used to solve the Ising Model. This is in contrast to the other critical exponents which are constants of the model.

A algorithm with $z > 0$ suffers from *critical slowing down*. From equation (3.33) we can see that this is the main obstacle for studying large systems. Thus it is our aim to calculate z for the Metropolis and Cluster algorithm to determine how much of an improvement the Cluster algorithm is at handling critical slowing down.

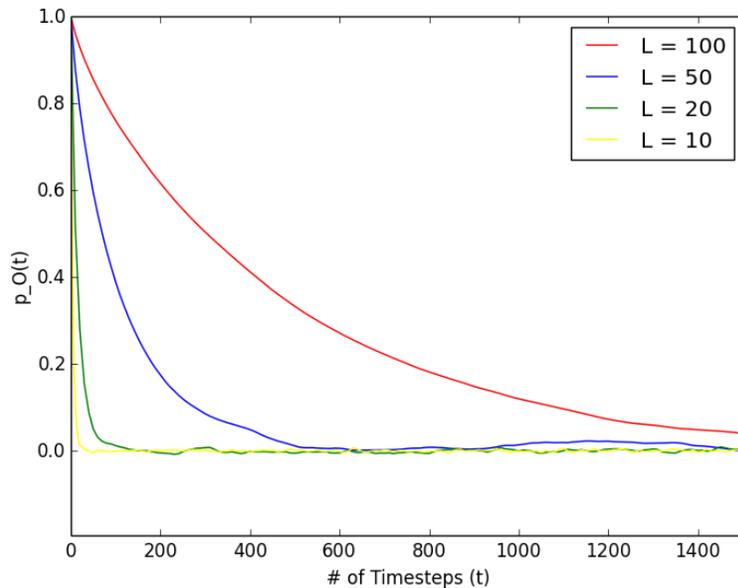


Figure 5.10: Autocorrelation functions for lattice sizes $L=10$; $L=20$; $L=50$; $L=100$ for the Metropolis algorithm. Sitting on the critical temperature $\beta = 0.440687$. Note $p(0) = 1$ on all cases.

In figure 5.10 we have plotted the autocorrelation function for the observable magnetisation in the critical (or *pseudo critical*) region for the Metropolis algorithm. This was done for 1,000,000 timesteps as we require $t_{max} \gg \tau$ as discussed in section 3.4. It can be clearly seen that there are much larger autocorrelation times for larger lattice sizes. According to

²Where it is relevant only in the pseudocritical region $\beta_c(L)$ where we have $\xi \sim L$.

equation (3.33) we can predict that the difference in these autocorrelation times should be given by

$$\frac{\tau_{L'}}{\tau_L} = a^z \quad (5.5)$$

where $L' = L \cdot a$, this can easily be derived from equation (3.33). First we must determine τ from the graph³.

Having calculated $p(t)$ to produce the graph we may calculate τ using equation (3.20). Where τ is the decay constant and its expectation value can be derived from the average of these computed values. This relation clearly does not hold for the entire region shown in figure 5.10 however. But we expect an exponential decay until at least $t = 2\tau$. This corresponds to $p(t) = e^{-2}$ as discussed in section 3.4.

A linear fit of the log of the autocorrelation function, figure 5.11, for $p_m(t) > e^{-2} \approx 0.14$ confirms that we may take equation (3.20) to be satisfied in this region. We must still note that this is an approximation. It can be seen from the figure that there is a small curvature for a low number of timesteps. If we recall our discussion in section 3.1.3, we argued that the eigenvalues of the Markov Matrix (P) fall off as we take P to higher powers. Hence the difference between 1 and the next highest eigenvalue gives the exponential autocorrelation time. This is neglecting the contribution from the third; fourth; etcetera largest eigenvalues, the contribution of which would be most significant for a small number of timesteps. This is a source of systematic error.

The values of τ obtained from the data is shown in table 5.2. Using the values from table 5.2 and equation (3.20) we find that

$$z_{Metropolis} = 2.11 \pm 0.073 \quad (5.6)$$

where the error was found from the standard deviation of the 24 values of z . Six of which were obtained from the τ s of the four different lattice sizes, and the similar data created in

³Technically we should be using τ_m and $p_m(t)$ in this as, magnetisation was the observable used. However, we shall refer to these as τ and $p(t)$ for convenience as only one observables needs to be considered while calculating z .

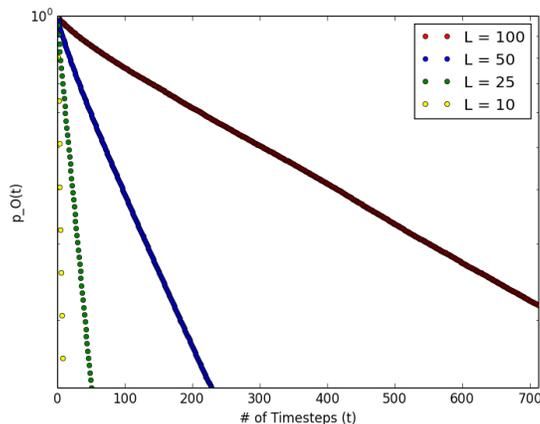


Figure 5.11: Plot of $\log(p_m(t))$ against t for $p_m(t) > 0.14$ and lattice sizes $L=25$; $L=50$; $L=100$. Sitting on the critical temperature $\beta = 0.440687$.

the same way using three different random number seeds than the one used to create the data in table 5.2. The same process was used to calculate $z_{cluster}$, except we had 40 estimates of z due to our extra simulation of the 100^2 system.

Considering that for $L = 100$ we only get an independent measurement every $2\tau \approx 866$ sweeps of the lattice, it is very costly to examine the critical region for larger systems where the autocorrelation time will more than quadruple each time we double L . Ie. for an $L=200$, we'd expect that we require greater than 40,000 sweeps for each independent measurement. Considering that each sweep will also require 4 times more calculations due to our definition of a step, this value of z makes it impractical to study large systems using the Metropolis algorithm. This is the effect of severe critical slowing down.

Table 5.2: Autocorrelation Times (β_c) for magnetization.

L	$\tau_{Metropolis}$	$\tau_{Cluster}$
200	—	5.7 ± 0.74
100	434 ± 10	4.2 ± 0.52
50	104 ± 4	3.2 ± 0.38
20	15.6 ± 0.6	2.2 ± 0.19
10	4.46 ± 0.04	1.61 ± 0.12

Note: Computing τ_m for the Metropolis for the 200^2 lattice would have required not only long simulation time per Monte Carlo step but also a very large sample as we recall that we require $t_{max} \gg \tau$ however we can see from the trend that τ here will be $> 1,600$.

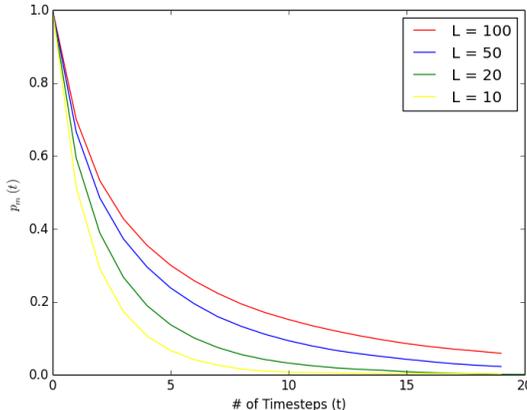


Figure 5.12: Autocorrelation functions for lattice sizes $L=10$; $L=20$; $L=50$; $L=100$ for the Metropolis algorithm. Sitting on the critical temperature $\beta = 0.440687$.

The autocorrelation function for the same lattice sizes, using data generated by the cluster algorithm is shown in figure 5.12. The autocorrelation times are also shown in table 5.2. Using these results we find

$$z_{cluster} = 0.47 \pm 0.051 \tag{5.7}$$

This can be compared to the binning method. We recall that when we calculate the error using equation (3.35) we underestimate the error since we have used states that are not statistically independent. We overcame this in our measurements by putting the measurements in bins and increasing our bins until the error stopped increasing which meant we had an independent sample. We also recall that the error increase began to plateau when number of elements in each bin (bin size) $= 2\tau$. Thus we can compare the plateau expected when binning, shown in figure 5.13. We can see that for both Metropolis 5.13b and Cluster 5.13a algorithms the error does indeed begin to plateau at 2τ by considering the values in table 5.2. It is arguable where this does 'plateaus', but if we recall that at $t = 2\tau$ we still

have an autocorrelation of ≈ 0.14 and thus we expect that the error will still raise for a while when are bins are larger than 2τ .

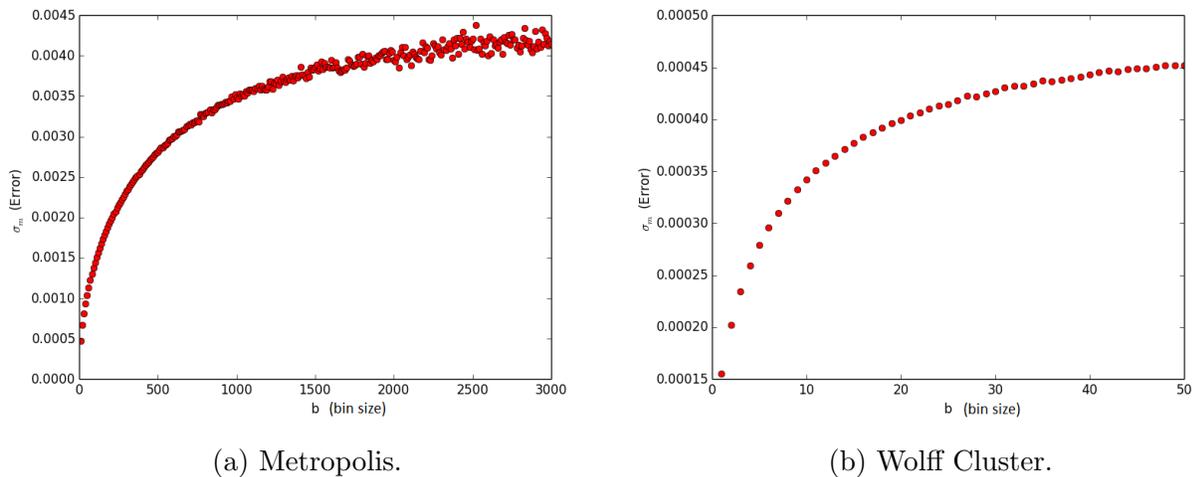


Figure 5.13: $L = 100$ Binning method.

Initial calculations of critical exponents yielded results with error that did not explain the difference between the computed value and the generally accepted ones. It was concluded that a bin size greater than 2τ was desirable and so we considered difference bin sizes in figure 5.13 to be representative of the plateaued region. The highest of these were 10,000 which was the bin size used to produce the figures in table 5.2. The values of τ produced using different random number seeds, as discussed previously, were also calculated for different bin sizes.

Chapter 6

Conclusions

The Ising model in 2-dimensions and zero magnetic field undergoes a second order phase transition from a phase with no magnetisation to one with spontaneous magnetisation. However, for a finite system there is no discontinuity in the observables or their derivatives. Observables such as magnetic susceptibility and heat capacity which diverge in the thermodynamic limit have a finite analytic maximum for the simulations made.

It is possible to compare the exact and statistical solutions for systems that are small enough such that their partition function is directly calculable. For a working algorithm and a sufficient amount of data the exact and statistical solutions should always agree within the estimate of statistical error for any given temperature.

In the cold region thermalisation depends on lattice size for both algorithms. In the hot region lattice size has little or no effect on the thermalisation time of the Metropolis algorithm and in fact there is a also very small effect of system size on autocorrelation times in this region. It generally takes less than 10 Monte Carlo steps to thermalise reatively large lattice sizes. It was clear to us from figure 5.3 that the fantastic efficiencies of the Metropolis algorithm in this region was due to the average fraction of lattice sites flipped per Monte Carlo step. When we compare this to the average cluster size at similar temperatures it is clear why one would prefer to use the Metropolis algorithm to simulate the system at hot temperatures.

While the fraction of spins flipped per Monte Carlo step in the Metropolis algorithm remains reasonably sized when we approach the critical region, it was found that due to the formation of clusters in the lattice and the fact that the local change property of the

Metropolis algorithm causes it to move clusters by predominantly changing their boundaries, there are always many sites that remain the same for a large number of Monte Carlo steps. We found that this was the cause of the extremely large correlation times of the system and hence the reason for its severe critical slowing down. The dynamic critical exponent of the algorithm was estimated to be 2.11 ± 0.073 which agrees well with literature ($z = 2.17$). This value of z implies that we more than quadruple autocorrelation times when we double L . This is clearly a major problem if we want to study large systems.

We can see from figure 5.3 that cluster size is quite large in the critical region. This combined with the fact that the Wolff Cluster algorithm is able to flip large cluster of spins at a time, rather than slowly moving their boundaries, makes it much more efficient at handling critical slowing down. Considering that the Wolff Cluster uses a randomly chosen initial seed it is easy to see that there shall be no sites that will be made 'difficult' for it to reach as is the case at the centre of cluster for the Metropolis algorithm. The dynamic critical exponent of this algorithm was estimated to be 0.47 ± 0.051 which agrees well with literature ($z = 0.52$).

Critical exponents which are properties of observables in the model and independent of the algorithm used to calculate them were calculated for magnetic susceptibility and heat capacity. We clearly have a very fast divergence of magnetic susceptibility (χ) which was calculated to have critical exponent $\gamma = 1.76 \pm 0.023$. This compared very well with the literature ($\gamma = 7/4$). It was clear from our graphs of heat capacity (C) as a function of inverse temperature (figure 5.8) diverges in the thermodynamic limit, however, at a much slower rate than the magnetic susceptibility. Thus our calculated value of $\alpha = 0.29 \pm 1.5$ seems reasonable. However it is agreed in literature that this critical exponent is 0. It is thought that this is due to the fact that the heat capacity has a logarithmic divergence. This may account that for relatively small systems α appears to be non-zero. It is necessary to consider how one can determine *alpha* accurately.

6.1 Acknowledgement

I would like to extend my gratitude to my project supervisor: Professor Stefan Sint for many useful meetings and emails throughout the year.

Bibliography

- [1] Konstantinos N. Anagnostopoulos, Computational Physics (National Technical University of Athens, 2014).
<http://www.physics.ntua.gr/~konstant/ComputationalPhysics/Book/ComputationalPhysicsKNA.pdf>
- [2] Stephen G. Brush, History of the Lenz-Ising Model, Review of Modern Physics, Volume 39 Number 4 (October 1967).
- [3] Chris Blair, The Ising Model: Monte Carlo Methods and Renormalisation, Maths Project, Trinity College Dublin (2010).
- [4] I. Sachs, S.Sen, J.C. Sexton, Statistical Mechanics, Cambridge University Press (2006).
- [5] Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. and Teller, E., J. Chem. Phys. 21 (1953) 1087.
- [6] Pillai, S.U., The Perron-Frobenius theorem: some of its applications, Dept. of Electr. Eng., Polytech. Univ. Brooklyn, NY, USA, March 2005.
- [7] Ulli Wolff, Comparison Between Cluster Monte Carlo Algorithms in the Ising Model, Institut für Theoretische Physik, Universität Kiel, Phys. Lett. B, **228**, 379 (1989).
- [8] Martin Lüscher, User's guide for ranlxs and ranlxd v3.2, (December 2005).
<http://luscher.web.cern.ch/luscher/ranlux/guide.pdf>
- [9] Joel Heinrich, Detecting a Bad Random Number Generator, University of Pennsylvania January 22, 2004.
- [10] Robert H. Swendsen, Jian-Sheng Wang, Nonuniversal Critical Dynamics in Monte Carlo Simulations, Physical Review Letters, **58**, 86 (1987).

- [11] Ulli Wolff, Collective Monte Carlo Updating for Spin Systems, *Physical Review Letters* **62**, 361 (1989).
- [12] Werner Krauth, *Cluster Monte Carlo Algorithms*, CNRS-Laboratoire de Physique Statistique Ecole Normale Sup?erieure, Paris (2008).
- [13] Martin Hasenbusch, *Monte Carlo Simulations in Statistical Physics and Quantum Field Theory*, Lecture notes, Humboldt University of Berlin, (last update July 2005).
<http://www-com.physik.hu-berlin.de/hasenbus/lecture.html>
- [14] Ulli Wolff, Monte Carlo errors with less errors, Humboldt University of Berlin, *Comput. Phys. Commun.* **156**, 143 (2004).