# Comparison of the Metropolis-Hastings and Wolff algorithms for Monte-Carlo simulation of phase transitions in the 2D Ising ferromagnet

*Robin de Freitas and Lewis Dean*

*10887826 and 10823674*

School of Physics and Astronomy

University of Manchester

April 2023

## Abstract

Phase transitions in the Ising model of a ferromagnet were investigated through Markov chain Monte-Carlo methods. The Metropolis-Hastings and Wolff algorithms were used to generate Markov chains for lattice sizes of 2x2, 4x4, 8x8 & 16x16 and 16x16, 25x25, 50x50 & 75x75 respectively, extracting ensemble averages for energy and magnetisation. These were used to obtain the Curie temperature and calculate the heat capacity and magnetic susceptibility across a range of temperatures. The results were then compared and used to assess the capabilities of these algorithms at different temperatures.

# 1 Introduction

Ferromagnets are materials with internal magnetisation at low temperatures due to a tendency for their atoms to align magnetic moments. They undergo a phase transition at the Curie temperature, above which they shift from exhibiting ferromagnetic behaviour to paramagnetic behaviour. The Ising model, in conjunction with Monte-Carlo (MC) methods, provides perhaps the simplest approach to calculating canonical ensemble averages for a ferromagnetic lattice and probing the behaviour of such about the magnetic phase transition. Crucially, the study of ferromagnetic phase transitions can apply to other physical systems through the notions of critical exponents and universality in the renormalisation group method [1]. Moreover, the Ising model sees application in other disciplines such as neuroscience; within that context, whether a site in a uniaxial ferromagnetic lattice is spin-up or spin-down corresponds to whether a given neuron in a neural network has fired or not [2].

This report strives to compare two Markov chain MC methods for generating accurate data from a 2D Ising model. These are the Metropolis-Hastings algorithm (hereon referred to as the Metropolis algorithm), which updates lattices locally, and the Wolff algorithm, a type of cluster algorithm. First we analyse differences between the thermal and magnetic responses in which said algorithms predict, followed by a comparison of Curie temperature accuracy.

# 2 Modelling ferromagnetism

A simple model for representing the atomic spins in an $n$ dimensional uniaxial lattice is a rank $n$ tensor containing $\pm 1$ values that reflect the corresponding spin orientations. This is the lattice devised by Wilhelm Lenz and later solved in one dimension by Ernst Ising. In the case of no external magnetic field, the Ising model has lattice Hamiltonian ($H$) calculated via [3],

$$H = -\frac{J}{2} \sum_{\langle i,j \rangle} \sigma_i \sigma_j. \tag{1}$$

Here $\sigma_k$ represents the direction of the $\mathrm{k}^{\mathrm{th}}$ magnetic moment which can take the values $\pm 1$, labelled spin-up and spin-down respectively. $J$ denotes a positive ferromagnetic coupling constant describing the exchange interaction. It must be positive to ensure that aligned moments contribute negatively to the lattice energy. A negative $J$ manifests antiferromagnetic behaviour. In this analysis, $J$ is set to 1 for simplicity. Summing over $\langle i, j \rangle$ refers to the sum of nearest neighbour contributions for all spin sites and the factor of a half is essential to avoid double counting during summation. The 2D Ising model has 4 nearest neighbours per spin site and this condition is maintained for sites along a lattice edge through the application of toroidal periodic boundary conditions, as shown in Figure 1.

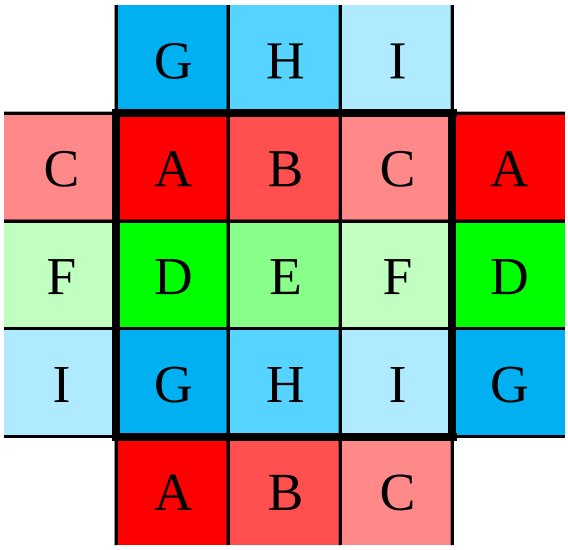


Figure 1: Diagram showing a 3x3 lattice (surrounded by the bold, black line) with the periodic boundary conditions around the outside.

.

Within this model, one can calculate the magnetisation per spin using,

$$m = \frac{1}{N}\sum_{i=1}^{N}\sigma_i, \tag{2}$$

where $N$ is the total number of spin sites. For a square lattice of side length $L$, $N$ is $L^2$.

### 2.1 Markov chain Monte-Carlo sampling

A Markov chain is a process in which the probabilities of subsequent event occurrences depend only on the current event; i.e. they have no memory of any previous events which led them to their current state. Thus, they are perfect for modelling stochastic trajectories within a state space. Furthermore, according to the ergodic hypothesis, the macroscopic properties of a system can be extracted by means of an ensemble average. Hence, MC simulations are often employed since averaging over a Markov chain can estimate said macroscopic properties.

The Markovian master equation for the Ising model [4] is,

$$\frac{\partial P_k(t)}{\partial t} = \sum_{k\neq l} P_l(t)R_{l\to k} - P_k(t)R_{k\to l}, \tag{3}$$

wherein $P_m(t)$ gives the probability of the lattice being in microstate $m$ at time $t$ and $R_{x\to y}$ is a conditional probability encapsulating the transition rate from state $x$ to $y$. Within a discrete-time Markov chain, as to be generated here, a $t$ step corresponds to a single transition between states.

Note that a system must have reached thermal equilibrium for the methods of statistical mechanics to apply. Such equilibria are characterised by detailed balance in which the net probability fluxes between all states in the state space are zero. This condition is met when the Markovian master equation is set to zero, which requires,

$$\frac{P_k}{p_l} = \frac{R_{l\to k}}{R_{k\to l}}. \tag{4}$$

Considering that our state space is a canonical ensemble, which has been allowed to thermalise, probabilities are given by the Boltzmann distribution. The probability then of the lattice possessing energy $\epsilon_k$ is,

$$P(\epsilon = \epsilon_k) = \frac{1}{Z} e^{-\epsilon_k \beta}, \tag{5}$$

with $\beta = 1/k_B T$. $Z$ is the partition function, $k_B$ the Boltzmann constant, which is set to 1 for convenience, and $T$ the temperature of the lattice's heat reservoir. Therefore, to meet the condition of detailed balance, the ratio of transition rates must satisfy,

$$\frac{R_{l\to k}}{R_{k\to l}} = e^{-\frac{\Delta\epsilon}{T}}, \tag{6}$$

in which $\Delta\epsilon$ is $\epsilon_k - \epsilon_l$. It is easy to see that when focusing on a single spin site $i$, flipping its spin would result in the energy change,

$$\Delta\epsilon = -2\epsilon_i, \tag{7}$$

in which $\epsilon_i$ is the original energy of site $i$. The transition rates chosen to satisfy detailed balance vary with the Markov chain MC algorithm implemented. Regardless of approach, the important result is that averaging over such a Markov chain gives canonical ensemble averages which may be used, for example, to calculate the magnetic contribution to heat capacity or magnetic susceptibility. Generally, an observable's MC average ($\langle O \rangle$) is calculated through,

$$\langle O \rangle = \frac{1}{N_{MC}} \sum_{i=1}^{N_{MC}} O_i, \tag{8}$$

for $N_{MC}$ samples of observable $O$ from a Markov chain.

### 2.2 Thermodynamic quantities

The partition function of the 2D Ising ferromagnet was solved for analytically by Onsager [5] but its use in the calculation of magnetic heat capacity ($C$) and susceptibility ($\chi$) can be bypassed if their equations are rewritten to contain the variance in energy and magnetisation, respectively. These latter statistics can be computed through MC simulation. Therefore,

$$C = \frac{\partial \langle E \rangle}{\partial T} = \frac{(\Delta E)^2}{T^2} = \frac{\langle E^2 \rangle - \langle E \rangle^2}{T^2}, \tag{9}$$

$$\chi = \frac{\partial \langle M \rangle}{\partial T} = \frac{(\Delta M)^2}{T} = \frac{\langle M^2 \rangle - \langle |M| \rangle^2}{T}, \tag{10}$$

where angular brackets symbolise thermal averages, $E$ represents the lattice energy calculated through the Ising Hamiltonian and $M$ is the total lattice magnetisation, $Nm$. Note the use of absolute magnetisation within the equation for $\chi$ is essential for producing plots representative of those in the thermodynamic limit. For finite size lattices, there is a finite probability of spontaneous magnetisation reversal which is completely unphysical for macroscopic lattices due to its infinitesimal probability of occurrence. This use of $|M|$ however leads to reasonable finite size scaling [6].

### 2.3 Ferromagnetic-paramagnetic phase transition

Within ferromagnetic materials, a continuous phase transition is observed. Such transitions have discontinuous second derivatives in free energy at their critical point with respect to some control parameter (temperature here). Given that response functions, such as heat capacity and magnetic susceptibility, are proportional to second derivatives in Helmholtz free energy, they too should be singular at the Curie temperature, $T_c$, the critical point for ferromagnetic phenomena.

At this critical point, behaviour characteristic of both the ferromagnetic and paramagnetic phases can be observed. Domain nucleation signifies the close proximity to low temperatures and yet the typical high temperature fluctuations are present, only at a larger scale. These conflicting behaviours mark the border between prioritising the minimisation of internal energy through spin alignment and the maximisation of thermodynamic entropy through random spin orientations. This priority is designated so as to best minimise the lattice's Helmholtz free energy.

Systems which exhibit phase transitions are said to possess an order parameter. These quantify the extent of transition between system phases across the critical point. A useful order parameter for ferromagnets is magnetisation because above the Curie temperature, we expect the average microstate magnetisation to be zero, reflecting the lattice's absence of order. On the contrary, the magnetisation below the Curie temperature increases with decreasing temperature, reflecting the increase in spin correlation due to exchange interaction. Pivotally, the Binder parameter [7] is the reduced fourth order cumulant, $U_4$, of the order parameter which can be expressed in the form,

$$U_4 = 1 - \frac{\langle M^4 \rangle}{3\langle M^2 \rangle^2}. \tag{11}$$

This is of great importance as it acts as a universality class whose value at the critical point converges on a scale invariant limit. As a consequence of this property, the intersection point of two Binder parameter curves as a function of temperature gives an estimate of the true Curie temperature in the thermodynamic limit.

# 3 Markov chain Monte-Carlo methods

## 3.1 Metropolis algorithm

The Metropolis algorithm satisfies the ratio of transition probabilities by locally updating a spin site each iteration. A spin site is selected randomly and a transition probability assigned dependent on the associated energy change. If the proposed flip decreases the site's energy, the transition automatically occurs (a probability of unity). Otherwise, the spin is flipped with a probability of $e^{-\Delta\epsilon\beta}$. Referring back to 6, with the latter probability here being $R_{l\to k}$, it is clear that the condition for detailed balance is met.

Prior to taking data from the Metropolis algorithm's implementation, the lattice's initial spin configuration must be allowed to reach equilibrium. This was considered satisfied when $m$ was seen to fluctuate about 0 above $T_c$ or 1 below $T_c$. By applying the algorithm over a transient period without sampling, the lattice is able to converge to thermal equilibrium for that temperature. Then energy and magnetisation samples from the current state are taken every $N$ Metropolis iterations.

A total of 1 million samples, per temperature, were gathered for lattices of side length 2, 4, 8 and 16 between the temperatures of 0.5 and 5. The relevant powers of energy and magnetisation could then be averaged for each temperature and used to produce response function plots as a function of temperature. Finally, the Metropolis algorithm was used to identify the Curie temperature of the 2D Ising ferromagnet using the concept of cumulant intersection discussed in Section 2.3. Again, 1 million samples were taken for each temperature for a temperature range of 2.23 to 2.28 with 0.01 steps.

## 3.2 Wolff algorithm

Unlike the Metropolis algorithm, the Wolff algorithm forms a cluster of spins around an initial spin site and flips the entire cluster at once. The initial spin site is randomly selected and its neighbouring spins are added to the cluster only if they are equal in spin and with a probability of $1-e^{-2\beta}$. This satisfies the detailed balance condition [8]. Each new point added has its neighbouring points which are not already in the cluster tested in the same way with this continuing until the cluster does not grow any further. The cluster is then flipped, energy and magnetisation recorded and a new initial point selected. The implementation of this is shown in Figure 2.

As with the Metropolis method, a wide sweep over temperatures from 0.5 - 5 (of step 0.1) was taken but now for lattices of side length 16, 25, 50 and 75. MC ensemble averages here used 100,000 samples per temperature. Again, an initial transient period was excluded from the data taken, allowing for equilibration of the lattice. Data to identify $T_c$ was also gathered in the temperature range 2.23 - 2.28, with a step of 0.01, for lattices of side length 25 and 50.

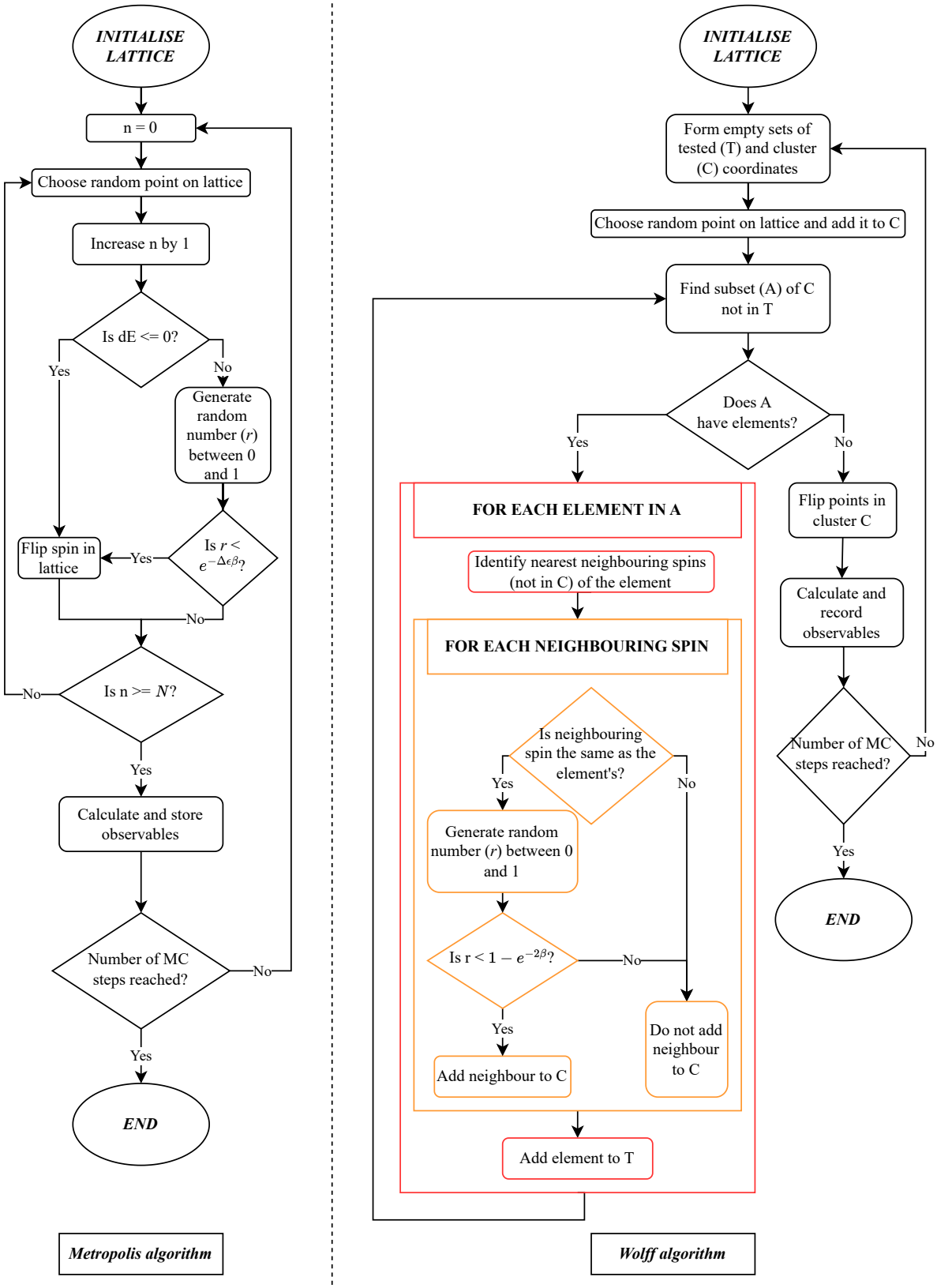


Figure 2: Flowcharts showing the steps of the Metropolis (left) and Wolff (right) algorithms in their implementation for a given temperature. The number of MC steps indicates the number of samples taken from the Markov chain.

# 4 Results

## 4.1 Heat capacity

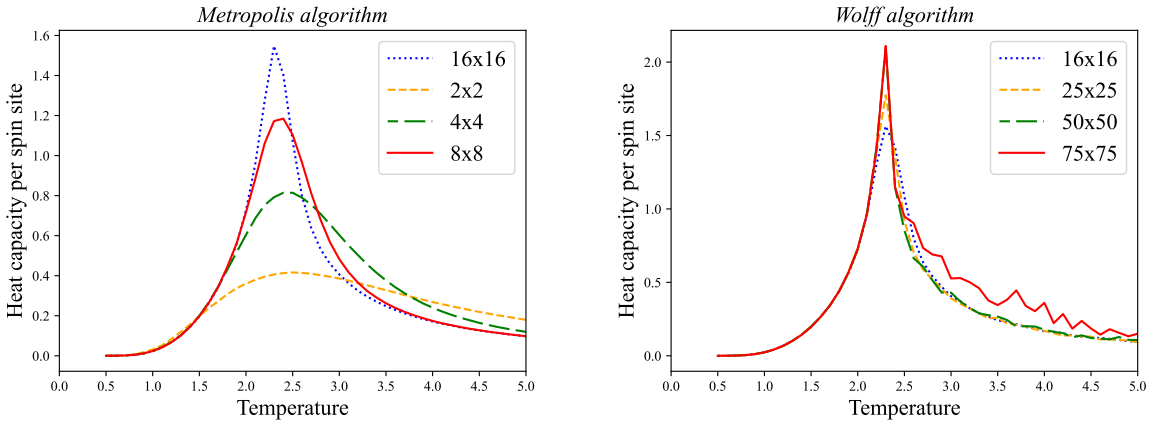


Figure 3: Plots of normalised heat capacities, as a function of lattice size and temperature, when the Metropolis algorithm was implemented (left) and the Wolff algorithm (right).

It can be seen in Figure 2, that both magnetic heat capacities have the appropriate functional form; most notably, they share an indication of divergence at $T_c$ if taken in the thermodynamic limit. This is much more pronounced in the Wolff data as it can be seen to reach a sharp peak with the step size used. In comparison, the Metropolis data is clearly continuous but indicates a strong progression towards a discontinuity with the amplitude of the peak increasing with lattice size. The Wolff algorithm shows an identically shaped plot for lattice size 16 to that of Metropolis indicating that the superior display of a discontinuity comes from its ability to generate large lattice Markov chains. The result is a better representation of the thermodynamic limit and hence a more accurate reflection of a real ferromagnet.

For $T < T_c$, the algorithms both display realistic thermal response. However, the Wolff algorithm forms very large clusters at these temperatures and can be expected to update approximately $N$ spin sites per implementation. In our implementation, this resulted in, the Wolff algorithm being slower per MC step than the Metropolis algorithm.

Where $T > T_c$, the Metropolis algorithm produces a much smoother heat capacity curve. This is likely because at high temperatures the Metropolis algorithm will visit each spin site on average once per MC step whereas the Wolff algorithm forms only very small clusters resulting in far less sites being updated in a MC step, leading to high autocorrelation (high similarity between the previous and current state). This is an issue in the application of the algorithm as such behaviour reduces the effective sample size for MC simulation. This minimal variation between adjacent samples limits the potential to manifest ergodicity, a process essential for state space

samples to generate physically meaningful ensemble averages.

### 4.2 Magnetic susceptibility

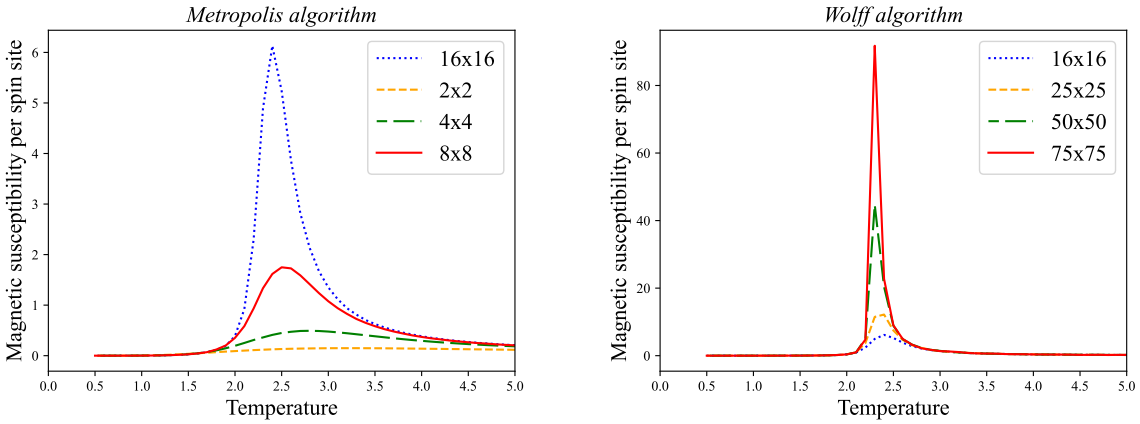


Figure 4: Plots of normalised magnetic susceptibilities, as a function of lattice size and temperature, when the Metropolis algorithm was implemented (left) and the Wolff algorithm (right).

The plots in Figure 3 allude to the expected form for magnetic susceptibility in the thermodynamic limit. The approach towards singularity centered on the Curie temperature marks the infinite derivative of the step-functional transition of magnetisation between the zero field paramagnetic and ferromagnetic phases. The two algorithms exhibit susceptibility differences similar to those for the heat capacities but those belonging to the 16x16 lattice, common to both data sets, have consistent forms as expected.

### 4.3 Curie temperature

It was apparent that a linear least squares fitting procedure would under-fit our data and thus a quadratic term was added as a first order correction for curvature. An $R^2$ statistic (supported for quadratic regression in [9]) was calculated for all regressions with none falling below 0.98, giving validity to the fit quality of fits shown in Figure 5. The exact Curie temperature of the 2D Ising model [10] is,

$$T_c = \frac{2}{\ln\left(1+\sqrt{2}\right)} \approx 2.269, \tag{12}$$

and so the Wolff algorithm has ascertained a more accurate Curie temperature of 2.268. The Metropolis algorithm gave an intersection at $T_c = 2.264$. The discrepancy between algorithms is to be expected given that the lattices in which this computation was feasible for Metropolis implementation contained an order of magnitude less in spin sites. The Wolff algorithm allows for Binder parameter calculations for large lattices which will inevitably converge on a temperature closer to that in the thermodynamic limit.

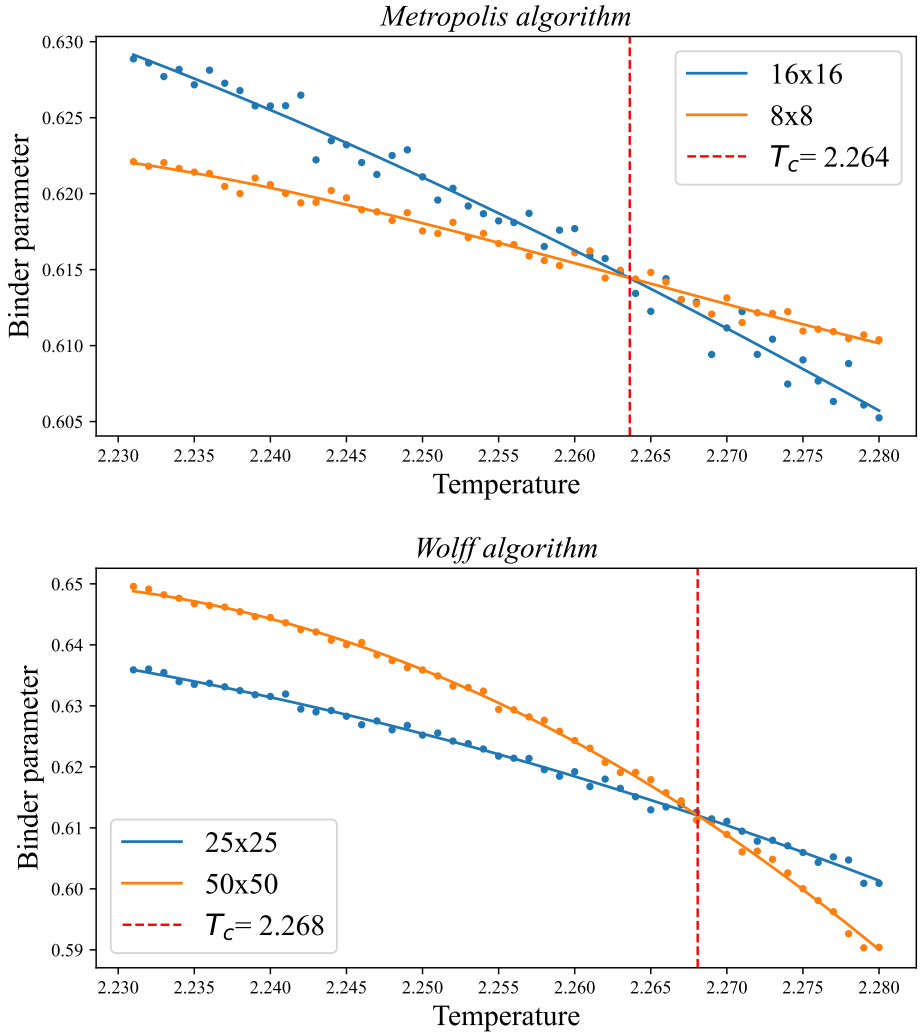


Figure 5: The plots above display the cumulant intersection procedure used to estimate a Curie temperature from both algorithms. The Metropolis plot is shown first with the equivalent for Wolff below. Quadratic least squares regression was used to extract intersections of the data trends. For Metropolis, the fits had $R^2$ statistics of 0.98 and for the Wolff fits the $R^2$ statistics were 0.99. The use of the Metropolis algorithm resulted in a Curie temperature of 2.264 and the Wolff algorithm found the temperature to be 2.268.

## 5 Assessing algorithmic efficiency near criticality

As has been seen the Wolff algorithm produces far more accurate data around the critical point, with the peaks of the heat capacity and magnetic susceptibility plots being sharper and $T_c$ being closer to the analytical value. The larger lattice size used to yield this improvement is not feasible with the Metropolis algorithm due to the occurrence of critical slowing down [11]. Close to the critical point, domains of like spins form; this means spin sites picked in the centre of the domain

will have a low probability of flipping. The result is a slower variation in states and consequently high autocorrelation in this region, requiring more MC steps in order to sample effectively.

In contrast, for a similar spin site selection at criticality, the Wolff algorithm's clusters allow for large changes in the spin configuration. This corresponds to exploring the state space, in accordance with the ergodicity condition, at a faster rate. Therefore, less MC steps are required for effective sampling.

# 6 Conclusion

To conclude, the application of both the Metropolis and Wolff algorithms produce ensemble-averaged quantities tending towards that seen in a continuous phase transition. This is seen in the $C$ and $\chi$ plots due to the increasingly high variance in energy and magnetisation, at the Curie temperature, with increasing lattice size. The Wolff algorithm is seen to be extremely efficient at producing accurate results in the critical region since it experiences reduced critical slowing down for these larger domains. This is made clear through the higher accuracy of Curie temperature, achieved with the larger lattices. However, at $T < T_c$, our implementation of the Wolff algorithm was often slower than the Metropolis algorithm, for a given lattice size, to complete the same number of MC steps. Moreover, at $T > T_c$ the Wolff algorithm used adjacent state samples that were severely under updated, leading to high autocorrelation and a reduced effective sample size. To offset this, more Wolff iterations could be applied before energies and magnetisations are sampled. In essence, while the Wolff algorithm prevails over the Metropolis close to criticality, it is by no means globally applicable. Prior knowledge of where high autocorrelation will occur would be needed in order to efficiently apply the algorithm to the Ising model, through temperature dependent MC steps.

## References


[1] A. Z. Patashinskiĭ and V. L. Pokrovskiĭ, "The renormalization-group method in the theory of phase transitions," *Soviet Physics Uspekhi*, vol. 20, pp. 31–54, Jan. 1977.

[2] E. Schneidman, M. J. Berry, R. Segev, and W. Bialek, "Weak pairwise correlations imply strongly correlated network states in a neural population," *Nature*, vol. 440, pp. 1007–1012, Apr. 2006.

[3] J. C. Goldstein and M. O. Scully, "Nonequilibrium properties of an ising-model ferromagnet," *Phys. Rev. B*, vol. 7, pp. 1084–1096, Feb 1973.

[4] R. J. Glauber, "Time-dependent statistics of the ising model," *Journal of Mathematical Physics*, vol. 4, pp. 294–307, Feb. 1963.

[5] L. Onsager, "Crystal statistics. i. a two-dimensional model with an order-disorder transition," *Physical Review*, vol. 65, pp. 117–149, Feb. 1944.

[6] J. Kotze, "Introduction to monte carlo methods for an ising model of a ferromagnet," 2008.

[7] K. Binder, "Finite size scaling analysis of ising model block distribution functions," *Zeitschrift fr Physik B Condensed Matter*, vol. 43, pp. 119–140, June 1981.

[8] U. Wolff, "Collective monte carlo updating for spin systems," *Physical review letters*, vol. 62, no. 4, pp. 361–364, 1989.

[9] F. Yao and H.-G. Muller, "Functional quadratic regression," *Biometrika*, vol. 97, pp. 49–64, Jan. 2010.

[10] T. W. Burkhardt, "Interface free energy and critical line for the ising model with nearest and next-nearest-neighbor interactions," *Zeitschrift fr Physik B Condensed Matter and Quanta*, vol. 29, pp. 129–132, June 1978.

[11] H. Gould and J. Tobochnik, "Overcoming critical slowing down," *Computers in Physics*, vol. 3, no. 4, p. 82, 1989.