Monte Carlo Simulations of Phase Transitions

Zhiyu Liu 10999337 and Benjamin Bottomley 10635454
Department of Physics and Astronomy, University of Manchester

The phase transition of the Ising model can be simulated by Monte Carlo algorithms, including single flip algorithms such as the Metropolis algorithm and more advanced cluster algorithms such as the Wolff algorithm. The advantages of cluster algorithms discussed in the report focus around their ability to converge to equilibrium much faster than single flip algorithms especially when around criticality. Errors in results can be decreased by reducing the correlation of Markov Chains. At different temperatures, observables such as magnetization per spin, energy per spin, specific heat, and magnetic susceptibility are measured and used to describe the ferromagnetic-paramagnetic phase transition. For finite size lattices, observables satisfy finite-size scaling functions in the vicinity of transition point. The critical temperature can be determined by the Binder cumulant.

I. INTRODUCTION

Lenz began investigating why magnets lose their magnetism when heated above the Curie temperature in the 1920s. Ising reduced the magnetic material to a one-dimensional chain of arrows with up or down spins. He demonstrated that this one-dimensional model could not maintain magnetism at any temperature and generalised his findings to 2D and 3D cases[1]. In 1944, Onsager solved the two-dimensional Ising model analytically and described the phase transition using a series of critical exponents.[2] The Ising model is powerful because it demonstrates that when a series of unrelated substances approaches a phase transition, they all have the same critical exponent. This is referred to as the universality of the Ising model, and it was mathematically demonstrated by Kenneth Wilson, earning him the Nobel Prize[3]. The systems macroscopic properties have been decoupled from its microscopic details. Any system with the same dimension and symmetry, regardless of whether it is composed of iron atoms, water molecules, or small arrowheads, will experience the same phase transitions.

The Metropolis-Hastings algorithm is the most frequently used Monte Carlo method for estimating the Ising model[4]. The objective is to generate representative samples in order to obtain thermodynamic quantities without solving the system analytically. When simulating, the time required for equilibration and the time required to eliminate the majority of correlation between states in the Markov Chain, referred to as the equilibration time and correlation time, respectively, are crucial, as they affect both the accuracy of the results and the amount of time consumed. A distinguishable feature of the Metropolis-Hastings algorithm is that only one spin is flipped along the Markov Chain. Modern cluster algorithms have replaced the single flip algorithms due to the significantly faster speed to reach equilibrium around criticality. The primary factor affecting time consumption and accuracy is the spatial correlation length. Spins mostly point in the same direction on this length scale, resulting in the critical slowing down of single flip algorithms[5]. Cluster algorithms flip a collection of spins together, resulting in a more rapid convergence to equilibrium. The autocorrelation functions of the Metropolis and Wolff algorithms have been compared, demonstrating that the Cluster algorithms produce results that are less correlated and thus more accurate.

The simulation is examined by demonstrating how observables change as it progresses through different phases. The magnetization per spin, energy per spin, specific heat, and magnetization susceptibility are typically chosen as the observables. By reading the plots of observables as functions of temperature, we can locate the regions of ferromagnetic phase, criticality, and paramagnetic phase.

Due to the fact that the system we simulate is constrained to a finite size, the results are size dependent, which is referred to as the finite-size effect. This can be expressed in terms of finite-size scaling relations, from which we can also obtain the Binder cumulant. The Binder cumulant can be used to locate the transition point. The value of the Binder cumulant at the transition point has been accurately determined using Monte Carlo methods[6].

II. THE ISING MODEL

The Ising model is the simplest model to describe the ferromagnetic-paramagnetic phase transitions. Consider a square lattice of size $L \times L = N$, on each site there is a spin that can take values ± 1. Suppose only neighbouring sites can interact, the Hamiltonian of the Ising model is:

$$E = -J \sum_{\langle i,j \rangle} \sigma_1 \sigma_2 \quad \sigma_i = \pm 1. \tag{1}$$

The sum $\langle i, j \rangle$ is over all neighbours of a site and the exchange energy J is set to 1 by convention for ferromagnetic interactions.

We study the properties of the material at a certain temperature, so the system can be considered as a canonical ensemble. Thus we have:

$$Z = \sum_{s_1=\pm 1} \sum_{s_2=\pm 1} \cdots \sum_{s_N=\pm 1} e^{-\beta E(s_1, s_2, \cdots, s_N)}. \tag{2}$$

When studying the thermal properties of the system, it is more common to consider intensive properties. We introduce the magnetisation per spin m, energy per spin E, specific heat C_V and magnetic susceptibility χ. If we generate samples of the configurations of the system according to the Boltzmann distribution, we get the expected value of energy per spin as:

$$\langle E\rangle = \sum_i P(E_i)E_i \approx \sum_i \frac{N(E_i)}{N}E_i = \frac{1}{N}\sum_j E_j, \quad (3)$$

where there are $N(E_i)$ samples of energy per spin E_i in the total N samples, and $\sum_j$ sums over all samples. There must be a sufficient number of samples to ensure that the simulation is ergodic.

If the system has energy E_0 and m_0, we flip every spin of the system and will get a system with energy E_0 and $-m_0$. These two systems have the same Boltzmann factor and thus has equal probability of being in that state at any temperature. We will regard the magnetization per spin as the absolute value $|m_0|$ since the direction of the magnetization is not what are interested about. Theoretically, under ergodicity the system goes over all possible configurations and the expected value of magnetization will be zero if we do not consider the absolute value. However at low temperature, the energy barrier for the system to flip all spins from m_0 to $-m_0$ is very huge. Phenomenologically the magnetization per spin of the system will persist in m_0 for a very long time. But near criticality the system travels from m_0 to $-m_0$ with a higher frequency so we must use the absolute value. This phenomenon is shown in the fourth subplot in figure (10).

$$\langle|m|\rangle \approx \sum_i \frac{N(|m_i|)}{N}|m_i| = \frac{1}{N}\sum_j |m_j|. \quad (4)$$

By taking the partial derivatives of $\langle E\rangle$ and $\langle M\rangle$ with respect to T, we derive the specific heat and susceptibility per spin, which can be related to the variance of E and $|m|$ as following:

$$C_V = \frac{\beta^2}{N}\big(\langle E^2\rangle - \langle E\rangle^2\big), \quad (5)$$

$$\chi = \frac{\beta}{N^2}\big(\langle M^2\rangle - \langle|M|\rangle^2\big) = \beta N^2\Big(\langle|m|^2\rangle - \langle|m|\rangle^2\Big), \quad (6)$$

where m is the magnetization per spin and $M = Nm$ is the magnetization of the system.

For finite size systems we need to consider boundary conditions to avoid disturbance from boundaries. Usually we employ periodic boundary conditions (PBC). Thus, the sites at the top and bottom, left and right, are considered neighbours.

III. METROPOLIS ALGORITHM

We expect to get samples that follow the Boltzmann distribution at a given temperature, this can be achieved by adopting Markov Chain Monte Carlo (MCMC) algorithms. For a Markov Chain with transfer matrix $\mathcal{P}$, the probability of transfer from state i to state j is $\mathcal{P}(i,j)$. The stationary distribution $\boldsymbol{\pi} = (\pi(1), \cdots)$, which satisfies $\boldsymbol{\pi}\mathcal{P} = \boldsymbol{\pi}$. $\boldsymbol{\pi}$ is independent of the initial state and is uniquely determined by $\mathcal{P}$. So we intend to find a transfer matrix which leads to the Boltzmann distribution $\pi(i) = \frac{1}{Z}e^{-\beta E(i)}$ at any initial state. A sufficient condition for a given $\boldsymbol{\pi} = (\cdots, \pi(i), \cdots)$ to be the stationary distribution, with $\mathcal{P}$ being the transfer matrix, is the detailed balanced condition:

$$\pi(i)P(i \to j) = \pi(j)P(j \to i). \quad (7)$$

If we choose

$$P(i \to j) = \min\left[1, \frac{\pi(j)}{\pi(i)}\right] = \min\left[1, e^{-\beta\Delta E}\right], \quad (8)$$

then (7) is satisfied and the stationary distribution will be the Boltzmann distribution. This is called the Metropolis algorithm. Its implementation on the Ising model is illustrated in algorithm (1), where the evolution of the system at temperature T corresponds to a Markov Chain. Within each procedure, a random site is picked. Whether the spin at that site will be flipped is determined by the acceptance rate shown in (8), which requires the calculation of energy difference between two states.

Algorithm 1 Metropolis Algorithm

Initialize: $L \times L$ LAT $= [[-1, \cdots, -1], \cdots, [-1, \cdots, -1]]$
procedure METROPOLIS SWEEP(T)
  for $i \leftarrow 1$ to L^2 **do**
    $x, y \leftarrow$ `nran(0,L-1)`
    $\Delta E \leftarrow$ Energy change of flipping LAT(x, y)
    if `ran(0,1)` $< e^{-\Delta E/T}$ **then**
      LAT$(x, y) \leftarrow -$LAT(x, y)
    end if
  end for
  Take measurements
end procedure

Algorithm (1) implements L^2 trials of flipping, intending to run over all L^2 sites. Such L^2 runs is called a Monte Carlo Sweep (MCS). When simulating, we first run sufficient MCS to achieve equilibrium, at which point the system corresponds to samples following the Boltzmann distribution, and then take energy and magnetization measurements following some specific MCS.

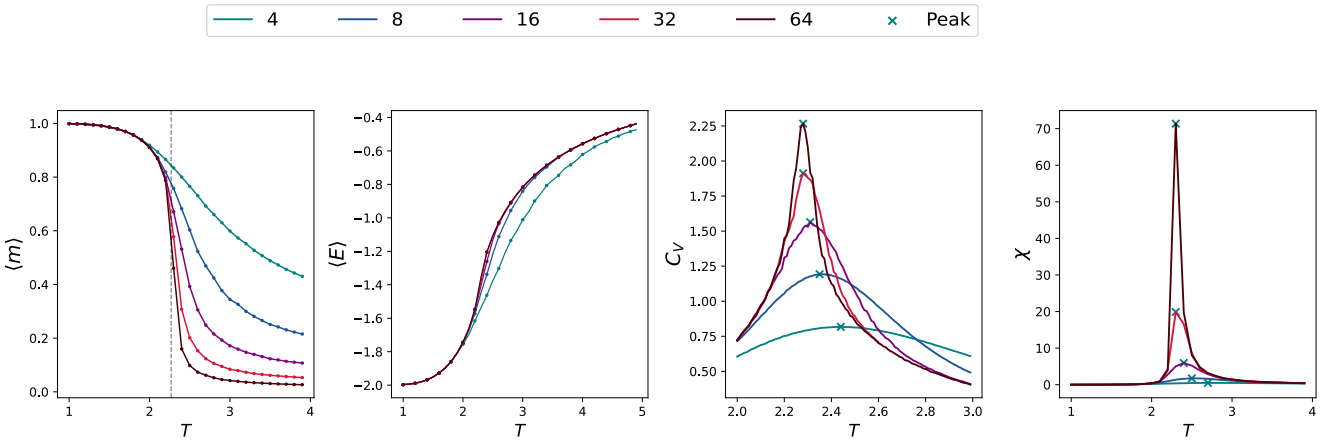

FIG. 1 The changes of Intensive properties with temperature compared with different sizes. The gradient of magnetization near phase transition becomes steeper as the lattice size gets larger(*left*). The curves of energy starts to converge for small sizes and no obvious transition point can be spotted(*middle*). The singularity of C_V, which is the phase transition point, moves leftwards and converges to T_C eventually. The movement of singularity is a result of finite-size scaling.

IV. GENERAL PROPERTIES

The system goes through three regions during the simulation as shown in Table (I).

Temperature regime	Phase	Correlation length
$T < T_C$	ferromagnetic	$\xi < L$
$T \sim T_C$	critical	$\xi \sim L$
$T > T_C$	paramagnetic	$\xi < L$

TABLE I: Phases of the system

The microscopic mechanism of phase transition is determined by the interaction of the microscopic particles that comprise the system and their random thermal motion, with temperature acting as an external parameter adjusting the strength of the competition between the two. This is exactly what the Ising model is simulating.

At $T \gg T_C$, the system is at paramagnetic phase, thermal fluctuations prevent the system to be magnetized, resulting in the expected absolute value of magnetisation per spin $\langle |m| \rangle$ to be 0. The system macrostate is thus symmetric, in accordance with its symmetric Hamiltonion. As the system cools down while still above the critical temperature, order begins to emerge and the absolute value of magnetisation increases. At $T < T_C$, the system is at ferromagnetic phase, the spins tend to align parallel, resulting a nonzero $\langle |m| \rangle$. When $B \neq 0$, since the symmetry breaking is caused by interactions within the system itself, it is called spontaneous symmetry breaking. This is illustrated in the first plot of figure (1), where second order phase transition can be seen clearly. As the system size grows, the curves get closer to the theoretical curve. The curve and its first order derivative should be continuous while its second order derivative is not continuous at the critical point.

Now consider the energy change during the simulation, at $T \ll T_C$, $e^{-\beta \Delta E} \to 0$ if $\Delta E > 0$, only flips that reduce the energy will be accepted, the system will stay in the lowest state at equilibrium. At $T \gg T_C$, $e^{-\beta \Delta E} \to 1$, almost every flip will be accepted, giving an almost zero $\langle E \rangle$. The second plot of figure (1) illustrates the expected mean energy per spin as a function of temperature, but the curves are not as distinct as the plot for magnetisation per spin.

The third subplot of figure (1) shows the specific heat per spin as a function of temperature. We expect C_V to diverge at T_C theoretically, so the potential critical point can be approximated as the maximum value as cross-marked in the figure. C_V tends to 0 for regions away from T_C, this is because changes in temperature have little effect on the energy. At low temperature, there is not enough energy to excite the transitions. At high temperature, all states are equally populated, almost no more energy can be absorbed. The rightward moving peak is a result of the finite size of the system that will be discussed later.

Magnetic susceptibility per spin as a function of temperature is shown in the last subplot of figure (1). A sharp peak can be easily observed, indicating that the transition is a second-order phase transition..

The correlation length ξ quantifies how strongly two spins are correlated, On this length scale, the majority of spins point in the same direction and cluster together. The net magnetization of clusters of size ξ contributes to the magnetization independently. At regions $T \ll T_C$ or $T \gg T_C$, $\xi \ll L$ so that we can assume every spin contributes to the magnetization independently. Near the critical point ($|T - T_C|/T_C \ll 1$) it is described by a critical exponent ν in the form $\xi \sim |T - T_C|^{-\nu}$[7] so it diverges at T_C. However due to the finite size of the system, the maximum of ξ can only be L.

Snapshots of the Monte Carlo simulation of the Ising model, with L equal to 256, are shown in figure (2). As the temperature gets higher from left to right and from top to bottom, the system goes through ordered ferromagnetic phase, criticality and disordered paramagnetic phase. The correlation length can be observed directly. Near the critical region, large clusters can be found depicting $\xi \sim L$. Clusters disintegrate as the temperature increases, finally reaching $\xi \ll L$ at disordered paramagnetic phase.

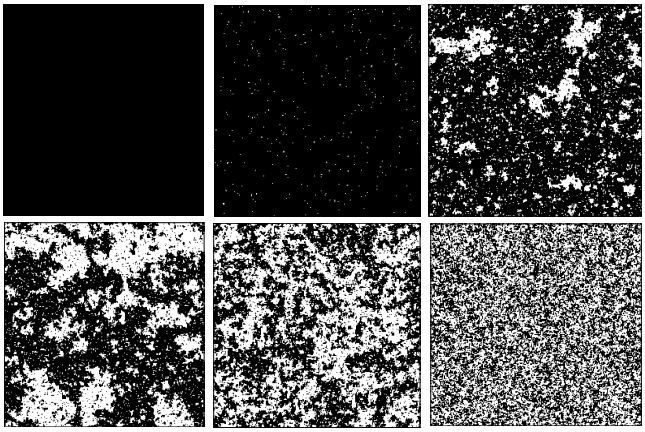

FIG. 2 Snapshots of Monte Carlo simulation of Ising model of size 256. The temperature regions are $T_1 < T_2 < T_3 < T_4 < T_5 < T_6$ from left to right, top to bottom, where $T_1 \sim 0$, $T_3 \sim T_C$, $T_6 \gg T_C$. Spins are coloured black/white for +1/-1 state.

V. DETAILED ANALYSIS

A. Behavior away from critical region

For N random variables x_i with the same distribution, define the random variable Y as following:

$$Y = \frac{\sum_i^N x_i - N\langle x\rangle}{\sqrt{N}}. \tag{9}$$

The central limit theory states that for $N \to \infty$ the distribution of Y tends to a Gaussian distribution with mean value $\langle y\rangle = 0$ and variance $\sigma^2(y) = \sigma^2(x)$.

In the paramagnetic phase, $\xi \ll L$ so each spin can be considered independent. Consider the set of independent variables $\left\{x_i : \frac{S_i}{N}\right\}$ and define the random variable

$$y = \frac{\sum_{i=1}^N x_i - N\langle x\rangle}{\sqrt{N}} = \frac{\sum_{i=1}^N \left(\frac{S_i}{N}\right) - N\langle \frac{S}{N}\rangle}{\sqrt{N}} = \frac{m - \langle m\rangle}{\sqrt{N}}. \tag{10}$$

According the the central limit theorem, y follows a Gaussian distribution with $\langle y\rangle = 0$, $\sigma^2(y) = \sigma^2(x) = \frac{1}{N^2}\sigma^2(s_i)$.

$$p(y) = \frac{N}{\sqrt{2\pi\sigma^2(s_z)}} \exp\left(-\frac{(m - \langle m\rangle)^2 N}{2\sigma^2(s_i)}\right). \tag{11}$$

Note that $\sigma^2(m) = \frac{1}{N}\sigma^2(s_i)$ for independent $\{s_i\}$. Thus we expect the width of the Gaussian $\Gamma \propto \sigma(m) \propto \frac{1}{\sqrt{N}}$. The distribution of m can be written as:

$$p(m) = p(y)\frac{\mathrm{d}p}{\mathrm{d}m} = \frac{1}{\sqrt{2\pi\sigma^2(m)}} \exp\left(-\frac{(m - \langle m\rangle)^2}{2\sigma^2(m)}\right). \tag{12}$$

For $T = T_c$, $\langle m\rangle \approx 0$, the normalized distribution functions are shown in figure (4). The y axis is normalized to $\frac{N(m)}{N_{\max}}$. It's clear that the widths are getting narrower as the theory prescribes. It can be speculated that for infinite large size $p(m = 0) = 1$. We can again verify this by plotting the magnetization per spin at each MCS shown in figure (3).

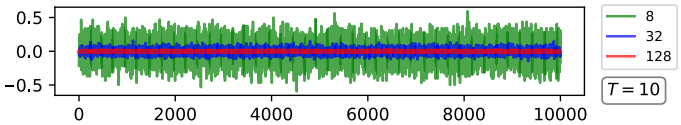

FIG. 3 Magnetization per spin at each MCS for size 8,16,32,128 at $T = 10$.

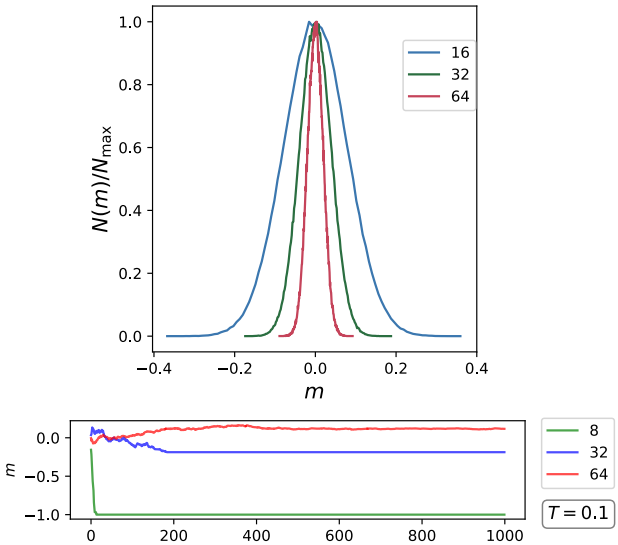

FIG. 4 Normalized distribution of m in varied sizes at $T = 10$.(*upper*) Trajectory of magnetization per spin as the simulation proceeds, showing spontaneous magnetization at ferromagnetic phase for varies sizes.(*lower*)

In the ferromagnetic phase where the system gets spontaneously magnetized it may stay in any stable or metastable states, as shown in figure (4). Under ergodicity the system will leave the metastable state and reach the stable state, there will be a pair of spontaneous magnetization m_{sp} and $-m_{\text{sp}}$. The distribution can be written as:

$$p(m) = \frac{1}{2}\left[g_m(+m_{\text{sp}}, \sigma^2(m)) + g_m(-m_{\text{sp}}, \sigma^2(m))\right], \tag{13}$$

where g_m stands for the Guassian distribution for m.

B. Finite-size scaling

Continuous phase transitions can usually be determined using power laws in the vicinity of the transition point. For the Ising model with infinite size:

$$\xi(T) \propto |T - T_c|^{-\nu}, \tag{14}$$
$$M(T) \propto |T - T_c|^{\beta}, \tag{15}$$
$$\chi(T) \propto |T - T_c|^{-\gamma}, \tag{16}$$
$$C_V(T) \propto |T - T_c|^{-\alpha}, \tag{17}$$

where $\alpha, \beta, \gamma, \nu$ are critical exponents

For finite large lattice sizes, the observables close to the transition point are given by[8][9]:

$$\langle m(T, L)\rangle = L^{-\beta/\nu}\mathcal{M}\Big[(T - T_C)L^{1/\nu}\Big], \tag{18}$$

$$\chi(T, L) = L^{\gamma/\nu}\mathcal{X}\Big[(T - T_C)L^{1/\nu}\Big], \tag{19}$$

where $\mathcal{M}, \mathcal{X}$ are scaling functions. Earlier discussion shows the difficulty of locating T_C simply by plotting the observables in figure (1), since the transition points are affected by the finite size of the system. Equations (18) and (19) show that at $T = T_C$, simulations for different sizes $m(T_C, L_1)L_1^{\beta/\nu}$, $m(T_C, L_2)L_2^{\beta/\nu}, \cdots$ have the same value $\mathcal{M}(0)$, from which we can determine T_C by finding the intersection of curves $m(T, L_i)L_i^{\beta/\nu}$. But this can not be performed since the critical exponents are unknown. However, we remove the critical exponents by division. Consider the Binder cumulant given by:

$$U(T, L) = 1 - \frac{\langle m^4(T, L)\rangle}{3\langle m^2(T, L)\rangle^2}. \tag{20}$$

Since $\langle m^k\rangle \sim L^{-k\beta/\nu}$, we have $U \sim \mathcal{U}\big[L^{1/\nu}(T - T_C)\big]$. Consequently we can determine T_C by locating the intersection of $U(T, L_1)$, $U(T, L_2), \cdots$, as shown in figure (5). Errors are estimated by calculating the square root of the sum of the diagonal elements of the covariance matrix, which can be simply obtained using the `scipy` package. The goodness of fit can be examined using Eq.(1) .The results of the fitted functions are shown in Table (II).

size	fitted function	errors
8	U_L = -0.264T+1.213	[0.0056, 0.013]
16	U_L = -0.521T+1.793	[0.0054, 0.012]
32	U_L = -1.008T+ 2.898	[0.0160, 0.037]

TABLE II Fitted function of cumulant for sizes of 8,16,32 along with the errors of the two parameters.

The data is not accurate enough to result a unique intersection point. Three intersection points are shown in Table (III). It's clear that the results are more accurate when analyzing larger sizes.

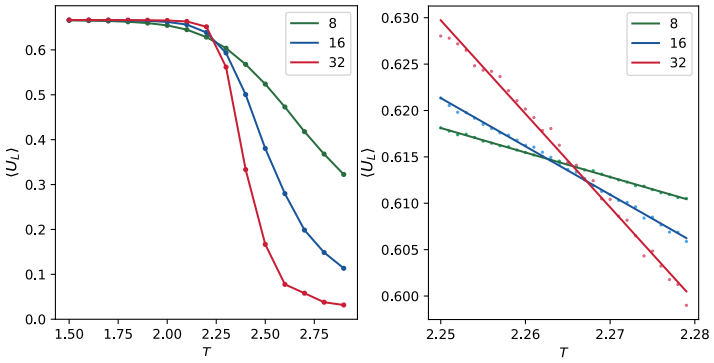

FIG. 5 Cumulant as a function of temperature for different sizes(*left*). Fitted linear function in the vicinity of the intersection point(*right*).

	intersection	$\left(\frac{\lvert T-T_C\rvert}{T_C}\right)$
(8,16)	2.2626	0.29%
(8,32)	2.2656	0.16%
(16,32)	2.2672	0.09%

TABLE III Intersection points and the accuracy compared to T_C.

C. Critical temperature for finite size system

Assume $\xi = L$ in the neighbourhood of critical temperature. For a system of size L^2, assume the critical temperature is $t_C(L)$. Then equation (14) indicates $L \propto |T_C - t_C(L)|^{-1/\nu}$, which yields:

$$t_C(L) = CL^{-1/\nu} + T_C. \tag{21}$$

Theoretically $\nu = 1$, so as the system size increases, $t_C(L)$ decreases, and converges to T_C eventually. This explains the rightwards moving of peaks of χ and C_V in figure (1).

VI. DISCUSSION OF ALGORITHMS

A. Equilibration time

Only after the system has reached equilibrium can measurements be taken; the time required τ_{eq} is measured in MCS. Table (IV) shows τ_{eq} at sizes of 8, 64, 128 using m as the criterion of equilibration. At T, the system is considered to be in equilibrium if $\Delta\sigma_m$ spends 99% of time in a certain range, as demonstrated in figure (3). It is clear to see that τ_{eq} has a positive correlation with system size and we can take $\tau_{\text{eq}} < 10000$ if simulating a system with size less than 128. Additionally, as the temperature increases, the equilibration time decreases, as it should with the increased possibility of flipping.

The equilibration time is also unpredictable. qFigure (6) shows the changes of magnetization per spin at each sweep. Lattice size of 128 reached equilibrium before size 32. In fact, the latter is struggling in a metastable state, which is common at low temperatures.

size	T	$\Delta\sigma_m$	$\min(\tau_{\rm eq})$	$\max(\tau_{\rm eq})$	$\overline{\tau_{\rm eq}}$	$\sigma(\tau_{\rm eq})$
16	0.1	0.0000	11	96	37.35	18.75
16	1	0.0156	9	80	29.24	11.02
64	0.1	0.0000	113	5363	721.43	808.00
64	1	0.0008	123	2400	478.50	407.25
128	0.1	0.0000	611	3072	1572.70	753.95
128	1	0.0005	294	3952	1055.40	765.60

TABLE IV : Equilibration time for different sizes measured at T. The maximum, minimum, average value

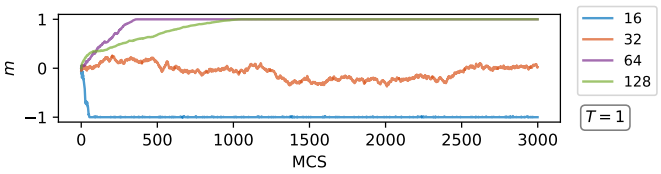

FIG. 6 Magnetization to MCS for size 8,16,32,128 at $T = 1$.

B. correlation time

Another important feature is the correlation time $\tau_{\rm cor}$, which is the time required for systems to be significantly uncorrelated. We consider the time-displaced autocorrelation function(ACF) $\Theta_m(t)$ of m, giving a measure of the correlation at two distinct times [10]:

$$\Theta_m(t) = \int \mathrm{d}t' \left[m(t')\, m(t'+t) - \langle m \rangle^2 \right], \tag{22}$$

where the time interval t is of unit MCS. It's clear to see that if $m(t')$ and $m(t+t')$ has only slightly difference, then $m(t')m(t+t')$ will always yield a positive value and contribute a large number to the integral. However in the uncorrelated situation $m(t')$ and $m(t+t')$ tend to take a random value according to their distribution and the integral tends to 0. The discrete form in which we will calculate is as follows[10]:

$$\begin{aligned}\Theta_m(t) \approx & \frac{1}{t_M - t} \sum_{t'=0}^{t_M - t} m(t') m(t'+t) - \\ & \frac{1}{t_M - t} \sum_{t'=0}^{t_M - t} m(t') \times \frac{1}{t_M - t} \sum_{t'=0}^{t_M - t} m(t'+t)\end{aligned} \tag{23}$$

where t_M is the total MCS in consideration, $\langle m \rangle$ is approximated as $\frac{1}{t_M - t} \sum_{t'=0}^{t_M - t} m(t')$ and $\frac{1}{t_M - t} \sum_{t'=0}^{t_M - t} m(t'+t)$.

The ACF is expected to fall off exponentially[10]:

$$\Theta_m(t) \sim e^{-t/\tau_{\rm cor}}. \tag{24}$$

Normalizing $\Theta_m(t)$ by dividing by $\Theta_m(0)$, constrains its maximum value to 1 and the constants are canceled. Thus the logarithm of $\Theta_m^{\rm norm}(t)$ is expected to have a linear relation with t:

$$\Theta_m^{\rm norm}(t) = e^{-t/\tau_{\rm cor}}, \quad \ln \Theta_m^{\rm norm}(t) = -t/\tau_{\rm cor}. \tag{25}$$

At $t = \tau_{\rm cor}$, $\Theta_m^{\rm norm}(\tau_{\rm cor}) = \frac{1}{e}\Theta_m^{\rm norm}(0)$, in order to obtain independent samples, the measurement time interval is set to be $2\tau_{\rm cor}$. Figure (7)(*left*) shows the magnetization normalized autocorrelation function(NACF) of a 64×64 Ising model at $T = 2.4$. The NACF decreases from one to close to zero thereafter. The embedded plot is semi-logarithmic, confirming the relation in equation (25). The gradient -0.0151±1.8 × 10^{-9} gives $\lceil \tau_{\rm cor} \rceil = 67$. Thus independent samples should be taken every $t > 2\tau_{\rm cor} = 134$ MCS. Considering the acceptance rate is less than 1, the Markov Chain may stay at its current state if the flip fails, to make the samples completely independent one needs more than 134 MCS. The NACF of samples taken every $2\tau_{\rm cor}$ MCS is shown in figure (25)(*right*). The value drops off rapidly at the beginning as expected, representing uncorrelated samples.

To demonstrate that taking samples every $2\tau_{\rm cor}$ does indeed have benefits, we compare the errors of energy simulated in two ways, either measured at every MCS, or measured every $2\tau_{\rm cor}$. To acquire exact solutions, we enumerate every possible configuration, compute the number of configurations $N(E)$ with energy E. Then the exact expected value is given by:

$$\langle E \rangle = \frac{1}{Z} \sum_E E N(E) e^{-\beta E}. \tag{26}$$

It is very hard to compute exact solutions for large systems, for instance a system with size 100 will have 2^{100} different configurations, which is larger than the number of particles in the universe. For simplicity we consider a system of size 6. The enumeration process employs grey code, which differs by only one digit between two adjacent numbers, much like adjacent states of a system differ by only one site. The enumeration results are acquired directly from [5].

Consider a Metropolis simulation of 50000 MCS taken into measurement for a system of size 6. In the first situation we take measurements just at the end of each MCS after $\tau_{\rm eq}$. Another way is to take 50000 measurements every $2\tau_{\rm cor} \approx 20$, which requires 1×10^6 MCS totally. The differences are given by:

$$\eta = \frac{|\langle E \rangle_{\rm Metropolis} - \langle E \rangle_{\rm exact}|}{\langle E \rangle_{\rm exact}} \times 100\%. \tag{27}$$

The results are shown in figure (8). Less correlated samples taken every $2\tau_{\rm cor}$ indeed result in higher accuracy.

C. Improvements of Monte Carlo

While the Metropolis Algorithm produces generally satisfactory results, there are obstacles that result in imprecise results and significant time consumption. Time

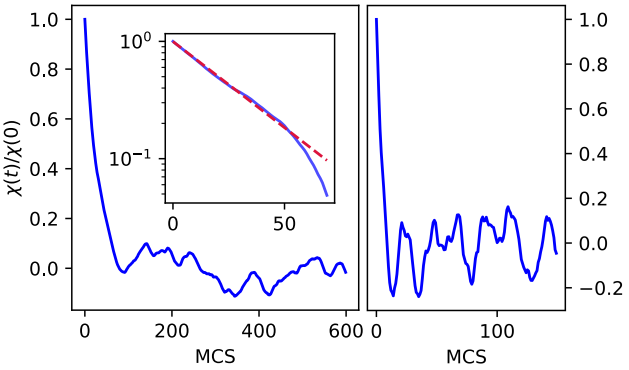

FIG. 7 Normalized autocorrelation function(NACF) of magnetisation per spin for Metropolis algorithm at $T = 2.4$. The embedded semi-logarithm plot confirms the exponential relation, with fitted linear function shows that $\tau_{\text{cor}} \approx 68$(*left*). The NACF for uncorrelated values measured every 70 MCSs drops rapidly.(*right*)

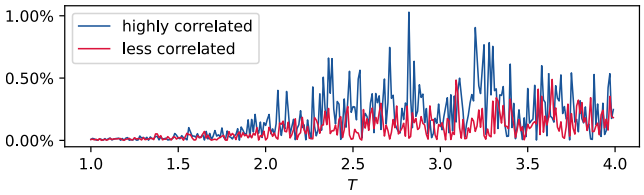

FIG. 8 Energy differences between exact enumeration and Metropolis sampling for highly correlated samples taken every MCS versus less correlated samples taken every τ_{cor}.

consumption is a significant issue for simulation and must be properly addressed prior to delving into research.

First we give some comments about speeding up the procedure shown in algorithm (1). A significant amount of time is consumed in step (4), which is the calculation of energy change, as we must obtain neighbours by performing modulo operation, each of which involves several arithmetic operations. To avoid this, we can keep a table of neighbours and look them up as needed.

Another factor of slowing down is due to exponential operations performed in step (6). Noticing ΔE can only take -8, -4, 0, 4 and 8, we can also store these values before the procedure to avoid doing exponential operations each loop. These two techniques are universal, which means they can be applied to any programming language and significantly increase performance.

Specifically for Python, one can use `Numba` package to speed up loops, which translates Python language to machine language. Its preference for `NumPy` is a significant advantage when performing scientific programming. In fact this package is nearly mandatory when performing simulations, without which the run will take hundreds of times longer.

Although we can speed up the run time using various ways, some problems in this algorithm are insurmountable, which will finally lead us to more advanced algorithms.

The correlation length $\xi \sim L$ causes clusters of scale L to form near T_C, as previously stated, within which spins are highly correlated, making it extremely difficult to flip a spin or a small patch of spins within the clusters. This means that the system tends to stay in its current state longer, slowing down the progress of the Markov Chain. This effect is called critical slowing down, and can be solved by using cluster algorithms.

VII. CLUSTER ALGORITHMS

Instead of trying to flip a single spin, cluster algorithms flip a correlated cluster of spins together. The procedure is as follows: beginning with a random spin, construct the cluster by adding neighbours which have the same spin orientation with probability p. During the construction, we keep a list of sites that have potential to grow new sites. The construction is finished when the list of potential sites is empty, then the all spins in the cluster are flipped. Denote the initial state as i and the final state after flipping as j. The probability $\mathcal{A}(i \to j)$ of constructing from i to j is the *a priori* probability of the transition in this Markov Chain. We have the detailed balanced equation:

$$\pi(i)\mathcal{A}(i \to j)P(i \to j) = \pi(j)\mathcal{A}(j \to i)P(j \to i), \quad (28)$$

where $\pi(i)$ follows the Boltzmann distribution, $P(i \to j)$ is the acceptance probability of the Markov Chain.

We focus on the cluster's edge because it indicates when the construction is complete. Assume all spins in the cluster are positive. At the boundary, there are two types of edges $\begin{Bmatrix} + & - & n_1 \\ + & + & n_2 \end{Bmatrix}$ with energy $E = n_1 - n_2$. Thus

$$\mathcal{A}(i \to j) = \mathcal{A}_{\text{in}}(1-p)^{n_2}, \quad (29)$$

$$\pi(i) = \pi_{\text{in}}\pi_{\text{out}}e^{-\beta(n_2-n_1)}, \quad (30)$$

where $\mathcal{A}_{\text{in}}$ stands for the probability of constructing the interior of the cluster. $\pi_{\text{in}}, \pi_{\text{out}}$ stands for the Boltzmann factor of the interior, exterior of the cluster, respectively. Therefore equation (28) can be rewritten as:

$$e^{-\beta(n_1-n_2)}(1-p)^{n_2}P(i \to j) = e^{-\beta(n_2-n_1)}(1-p)^{n_1}P(j \to i). \quad (31)$$

The acceptance probability is

$$\begin{aligned} P(i \to j) &= \min\left[1, \frac{e^{-\beta(n_2-n_1)}(1-p)^{n_1}}{e^{-\beta(n_1-n_2)}(1-p)^{n_2}}\right] \\ &= \min\left[1, \left(\frac{e^{-2\beta}}{1-p}\right)^{n_2}\left(\frac{1-p}{e^{-2\beta}}\right)^{n_1}\right]. \end{aligned} \quad (32)$$

If we take $p = 1 - e^{-2\beta}$, then $P(i \to j) = 1$, every constructed clusters will be flipped. The procedure is shown in algorithm (2)

Algorithm 2 Wolff Algorithm

Initialize: $N \times N$ LAT $= [[-1, \cdots, -1], \cdots, [-1, \cdots, -1]]$

procedure CLUSTER ISING(T)
  $x, y \leftarrow$ `nran(0,N-1)`
  Cluster $\leftarrow (x, y)$, Potential $\leftarrow (x, y)$
  while Pocket ! $= \varnothing$ **do**
    $(x_k, y_k) \leftarrow$ any element in Potential
    for $\forall (x_l, y_l) \notin$ Cluster and (x_l, y_l) being neighbor of (x_k, y_k) and LAT(x_k, y_k)=LAT(x_l, y_l) **do**
      if `ran(0,1)` $< p$ **then**
        Potential $\leftarrow$ Potential $\cup$ (x_l, y_l)
        Cluster $\leftarrow$ Cluster $\cup$ (x_l, y_l)
      end if
    end for
    Potential $\leftarrow$ Potential $\setminus$ (x_k, y_k)
  end while
  for $\forall (x_k, y_k) \in$ Cluster **do**
    LAT$(x_k, y_k) = -$LAT(x_k, y_k)
  end for
end procedure

The Wolff algorithm increases speed around T_C significantly as clusters that are highly correlated are broken which leads to equilibrium rapidly. Here we compare the snapshots of configuration using Metropolis and Wolff at $T = 2.26$ as shown in figure (9). Figures (1) and (2) of (9) are snapshots of Metropolis simulations, the vast majority of the sites have the same orientation of spin, which will make the system get stuck in a regime of correlated states. Figures (3) and (4) of (9) are snapshots of Wolff simulations, the independent cluster is coloured differently. It's clear that the huge clusters are destroyed, making the system more likely to have transitions between less correlated states. This can be better understood by observing the magnetization per spin at each MCS, as shown in figure (10). The first two figures compare the magnetization per spin of Wolff simulation and Metropolis simulation. Wolff simulation transit from two magnetization peaks with much higher frequency than Metropolis simulation. Some hundreds of Wolff procedures can do the work of tens of thousands of Metropolis sweeps. The same can be deduced by comparing the last two figures, which are simulated at $T = 2.4$. The values spread more evenly at $T = 2.4$ compared to the first two figures and the transition between two peaks are more frequent in Metropolis simulation. Still, hundreds or even dozens of Wolff procedures can do the work of tens of thousands of Metropolis sweeps.

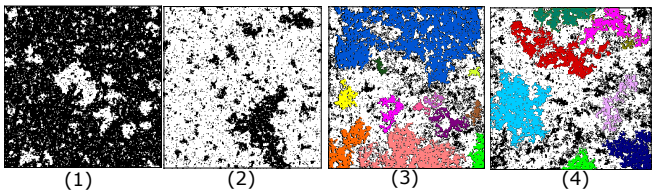

FIG. 9 Snapshots of the system in the middle of simulation at $T = 2.26$. (1) and (2) are results from Metropolis simulation. (3) and (4) are from Wolff simulation.

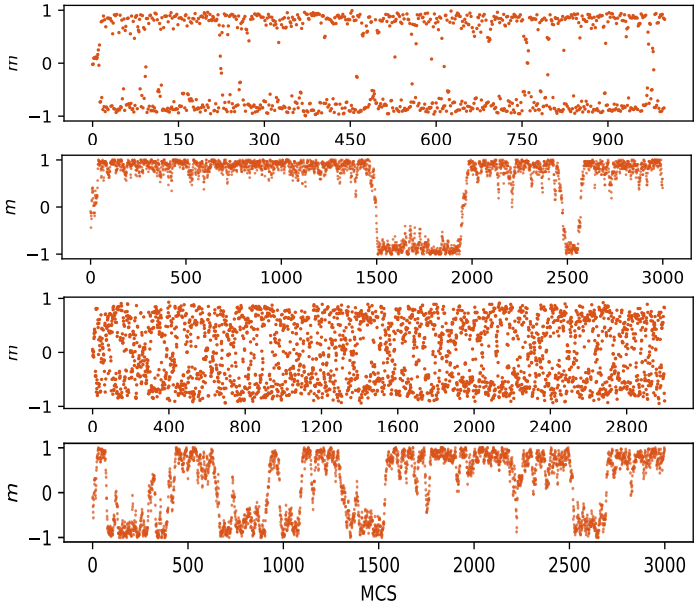

FIG. 10 Magnetization per spin of system with size 8 at each MCS . The figures are numbered (1)∼(4) from top to bottom.(1) Wolff simulation at $T = 2.2$. (2) Metropolis simulation at $T = 2.2$. (3) Wolff simulation at $T = 2.4$. (4) Metropolis simulation at $T = 2.4$.

Finally we consider the autocorrelation function of samples from Wolff simulation, as shown in figure (11). The curve drops off instantly and stays in a region close to 0. The embedded semi-logarithm plot has gradient $\frac{1}{\tau_{\text{cor}}} = 0.178 \pm 3E-5$, giving $\lceil \tau_{\text{cor}} \rceil = 6$.

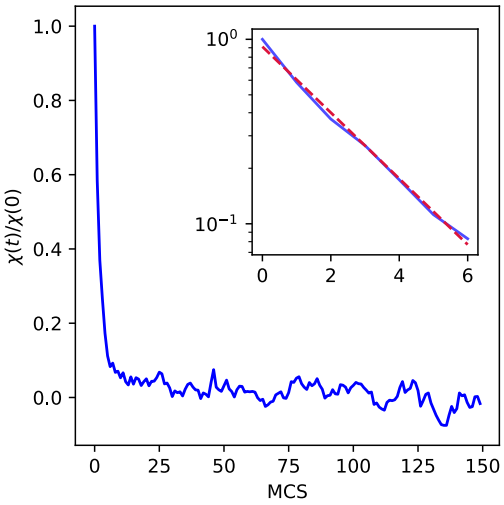

FIG. 11 Normalized autocorrelation function of magnetization per spin for Wolff algorithm at $T = 2.4$. The embedded semi-logarithm plot confirms the exponential relation, with fitted linear function giving $\tau_{\text{cor}} \approx 6$.

VIII. INTRODUCTION TO MORE MODELS

Ferromagnetic materials are described using the most appropriate models. The Ising model describes materials with strong uniaxial anisotropy. Australian mathematician Renfrey B. Potts relaxes the number of spin states in the Ising model from two to any number giving the Potts model:

$$\mathcal{H}_{\text{Potts}} = -J \sum_{\langle i,j \rangle} \delta(\sigma_i, \sigma_j), \tag{33}$$

where σ_i can take values from $\{1, 2, \cdots, p\}$. A more general XY model is given when the spins are vectors corresponding to any angle θ:

$$\mathcal{H}_{\text{XY}} = -J \sum_{\langle i,j \rangle} \boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_j = -J \sum_{\langle i,j \rangle} \left(\sigma_i^x \sigma_j^x + \sigma_i^y \sigma_j^y \right). \tag{34}$$

This model is considered when the material has planar anisotropy, which means that the spins are restricted in the xy plane. For fully isotropic materials we use the Heisenberg model:

$$\mathcal{H}_{\text{Heisenberg}} = -J \sum_{\langle i,j \rangle} \boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_j = -J \sum_{i,j} \begin{pmatrix} \sigma_i^x \\ \sigma_i^y \\ \sigma_i^z \end{pmatrix} \cdot \begin{pmatrix} \sigma_j^x \\ \sigma_j^y \\ \sigma_j^z \end{pmatrix}. \tag{35}$$

Now we provide some details about Potts model in comparison with the Ising model. Suppose there are four different values of σ, $(1, 2, 3, 4)$ correspond to red, yellow, blue and green. Then the 5×5 system shown in figure(12), where the outmost red sites are boundary conditions, has energy -18 with $J = 1$.

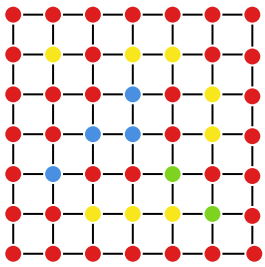

FIG. 12: Example of the Potts model

c

Kawasaki dynamics consider systems with a conservation law, which fixes the concentration of up spins and down spins. We can implement the same Metropolis algorithm as the Ising model, but instead of flipping a spin, we exchange spins.

IX. CONCLUSION

Finally, the Metropolis algorithm is the most frequently used method for simulating the Ising model. The ferromagnetic-paramagnetic transition's properties are illustrated clearly. The system transitions from a disordered paramagnetic phase to an ordered ferromagnetic phase as it cools from a high temperature. Magnetisation per spin and energy per spin are both continuously changing quantities, as are their first order derivatives. The variance of energy per spin and magnetization, respectively, have a singularity at the critical temperature, indicating a peak for finite systems. Scaling functions and critical exponents are used to describe finite-size effects. We can extract the Binder cumulant from these functions and use it to locate the transition point, the result we obtain is $T_C = 2.2672$. Two critical characteristics of Monte Carlo algorithms are their equilibration and correlation times. The equilibration time has a positive dependence on the system size and has a negative dependence on the temperature. The correlation time indicates when two states have lost the majority of their correlation, indicating how frequently we should take independent samples. We demonstrated that sampling at each correlation time interval does indeed result in results with smaller errors. Due to the divergence of correlation length, the Metropolis algorithm suffers the critical slowing down around criticality. As a result, we use the Wolff algorithm, which converges significantly faster than the Metropolis algorithm, as well as resulting in a shorter correlation time. Finally, we give a brief introduction about the other various models used to describe different types of ferromagnetic materials.

[1] Martin Niss. History of the lenz-ising model 1920–1950: From ferromagnetic to cooperative phenomena. *Archive for History of Exact Sciences*, 59(3):267–318, Mar 2005.

[2] Lars Onsager. Crystal statistics. i. a two-dimensional model with an order-disorder transition. *Phys. Rev.*, 65:117–149, Feb 1944.

[3] Kenneth G. Wilson. Renormalization group and critical phenomena. i. renormalization group and the kadanoff scaling picture. *Phys. Rev. B*, 4:3174–3183, Nov 1971.

[4] W. K. Hastings. Monte Carlo Sampling Methods using Markov Chains and their Applications. *Biometrika*, 57(1):97–109, April 1970.

[5] Werner Krauth. *Monte Carlo Simulation in Statistical Physics*. Oxford University Press, 2006.

[6] D Nicolaides and A D Bruce. Universal configurational structure in two-dimensional scalar models. *Journal of Physics A: Mathematical and General*, 21(1):233–243, jan 1988.

[7] Dieter W. Heermann Kurt Binder. *Monte Carlo Simulation in Statistical Physics*. Springer, 2019.

[8] K. Binder. Critical properties from monte carlo coarse graining and renormalization. *Phys. Rev. Lett.*, 47:693–696, Aug 1981.

[9] Michael E. Fisher and Michael N. Barber. Scaling theory for finite-size effects in the critical region. *Phys. Rev. Lett.*, 28:1516–1519, Jun 1972.

[10] Helmut G. Katzgraber. Introduction to monte carlo methods, 2009.