

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA DE SÃO CARLOS

RELATÓRIO CIENTÍFICO FINAL DE INICIAÇÃO CIENTÍFICA
FOMENTADO PELA FUNDAÇÃO DE AMPARO À PESQUISA DO
ESTADO DE SÃO PAULO

Monte Carlo techniques applied to phase transitions in magnetic
systems

Processo FAPESP 2024/18683-1 com vigência de 01/01/2025 a 31/12/2025
Relatório referente ao período de 01/01/2025 a 31/12/2025

Bolsista: Tony Maciel Alexander

Orientador: Prof. Dr. José Abel Hoyos Neto

SÃO CARLOS - SP
DEZEMBRO DE 2025

SUMMARY

This research project employs basic concepts of phase transitions and Monte Carlo techniques to the q -state clock model and the 2D XY $J_1 - J_2$ model on the square lattice with frustrated nearest and next-nearest-neighbor interactions. First, the Weiss molecular field theory and Ginzburg-Landau mean-field theory will be used to try to quantify these models' critical exponents. Then, the Metropolis and heat bath single-spin-flip algorithms will be introduced and proven, alongside the Wolff and Swendsen-Wang cluster-flip algorithms. Afterwards, multiple different variables will be introduced, explained and compared to study the phase transitions that occur in these models, such as various Binder cumulants and the helicity modulus. In doing so, the physics of these models will be elucidated, such as the emergence of a quasi-long-range order phase for the q -state clock model and the order-by-disorder mechanism at play in the 2D XY $J_1 - J_2$ model. Finally, finite-size scaling theory will be discussed as a way to estimate the critical exponents and critical temperatures for both models studied. With these methods: the critical exponents of the 2D Ising model (including the dynamical critical exponents for the Metropolis and Wolff algorithms, to verify the critical slowing down phenomenon) will be calculated and compared with their known analytical results (except for the dynamical critical exponents, which don't have exact results), a phase diagram will be made for the q -state clock model and the 2D XY $J_1 - J_2$ model, the Metropolis, heat bath and Wolff algorithms will be compared to see which one can simulate the 2D XY $J_1 - J_2$ model more efficiently, and the universality class of the phase transition that occurs for the 2D XY $J_1 - J_2$ model in the frustrating regime ($J_2/J_1 > 0.5$) will be numerically verified to belong to the 2D Ising universality class.

CONTENTS

Contents	3	
1	INTRODUCTION	4
1.1	Mean-field theory	5
1.1.1	Weiss molecular field theory	5
1.1.2	Ginzburg-Landau mean-field theory	10
2	METHODS	12
2.1	Monte Carlo algorithms	12
2.1.1	Metropolis	14
2.1.2	Heat bath	14
2.1.3	Wolff	16
2.1.4	Swendsen-Wang	17
2.1.5	Choosing an algorithm	18
2.1.6	Equilibration	19
2.2	Relevant variables	20
2.2.1	Magnetizations and energy	20
2.2.2	Specific heat and magnetic susceptibility	21
2.2.3	Correlation functions and critical slowing down	21
2.2.4	Binder cumulants and the BKT transition	23
2.2.5	Helicity modulus (or spin stiffness)	27
2.2.6	Nematic order parameter and order-by-disorder	30
2.3	Finite-size scaling	36
3	Practical considerations	38
4	RESULTS	41
4.1	q-state clock model	41
4.1.1	q = 2 (Ising)	41
4.1.2	q ≥ 3	46
4.2	2D XY J ₁ – J ₂ model	52
5	CONCLUSION	60
6	REFERENCES	60

1 INTRODUCTION

The Hamiltonian of the q -state clock model is given by Eq. 1:

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} \vec{S}_i \cdot \vec{S}_j = -J \sum_{\langle i,j \rangle} \cos\left(\frac{2\pi}{q}(p_i - p_j)\right) , \quad J \in \mathcal{R}_+^* , \quad (1)$$

where $\vec{S}_i \equiv (\cos(\theta_i), \sin(\theta_i))$ ($\theta_i \equiv p_i 2\pi/q$ and $p_i \equiv 0, 1, \dots, q-1$) is one of the q possible states of site i . It will be assumed that the lattice is rectangular, two dimensional, has periodic boundary conditions and every site is a localized particle that only interacts with its four nearest neighbors. If $q = 2$, this reduces to the 2D Ising model. If $q \rightarrow \infty$, this becomes the 2D XY model.

The Hamiltonian of the 2D XY $J_1 - J_2$ model is given by Eq. 2:

$$\mathcal{H} = J_1 \sum_{\langle i,j \rangle} \vec{S}_i \cdot \vec{S}_j + J_2 \sum_{\langle\langle i,j \rangle\rangle} \vec{S}_i \cdot \vec{S}_j , \quad J_1, J_2 \in \mathcal{R}_+^* , \quad (2)$$

where $\vec{S}_i \equiv (\cos(\theta_i), \sin(\theta_i))$ ($\theta_i \in [0, 2\pi)$) is a 2D vector on the unit circle. Again, periodic boundary conditions and a rectangular lattice will be assumed. Each site interacts with its nearest neighbors (with coupling energy J_1) and next-nearest-neighbors (with coupling energy J_2).

Both of these models will be studied to reproduce various results in the literature. The following sections present: an analytical approach with mean-field theory, the basics of Monte Carlo simulations of systems in thermal equilibrium (elucidating various algorithms and proving detailed balance for each one), an exposition of relevant observables that will aid in describing the physics of each model, finite-size scaling theory to calculate critical exponents and critical temperatures, some practical considerations to bear in mind and a potpourri of results (including phase diagrams, plots of thermodynamic variables, data collapses, tables of calculated critical exponents etc.).

In practice, the importance of these paradigmatic models resides in their simplicity to be studied, while simultaneously being able to exhibit rich physical phenomena. To illustrate this, consider the Mermin-Wagner theorem (1) which states that for systems in 2 dimensions with short-range interactions (such as the ones in Eqs. 1 and 2), continuous symmetries cannot be spontaneously broken at finite temperatures. The 2D XY model is a well known example of a system that satisfies these criteria, but it was discovered by Kosterlitz and Thouless in 1973 (16) that it does in fact have a phase transition: it has a paramagnetic phase with exponentially decaying spin-spin correlation function (finite correlation length) and a quasi-long-range order phase at low temperatures with algebraically decaying spin-spin correlation function (infinite correlation length). Since the 2D Ising model already had an exact solution at that time thanks to Onsager (26), it was known that it had a long range order phase at low temperature and a paramagnetic phase

at high temperature. The phase transition between these two phases spontaneously breaks the discrete Z_2 symmetry. In this light, some very natural questions arise: “What does the phase diagram of the 2D Ising model look like when spins take on more directions and the system starts to resemble the XY model? What happens to the long range order phase? When does the quasi-long-range order phase appear?” The answer to all these questions can be uncovered by studying the q -state clock model. On the other hand, the 2D XY $J_1 - J_2$ model is important to be studied because it is one of the simplest models to exist that allows for magnetic frustration. Moreover, from a computational physics perspective, these models are frequently used to test emergent computational theories, such as novel Monte Carlo methods, tensor network approaches (31), parallelization techniques on the GPU (14, 15) etc. . Nonetheless, before any numerical methods are to be employed, a theoretical inquiry is in order, and to this end, mean-field theory will be a starting point.

1.1 Mean-field theory

Mean-field theory has many different formulations, but all of them essentially produce a framework that aims to simplify the analysis of systems and phase transitions by mathematically treating the constituents as non-interacting (2 - 5). Despite being a major approximation, it’s able to capture the general behavior of phase transitions in a quantitative manner, like, for instance, predicting some critical exponents reasonably well (but, in general, it provides wrong expressions for relevant variables).

Since the critical exponents for the Ising model ($q = 2$) are known exactly, it’s prudent to calculate these same critical exponents with mean-field theory and Monte Carlo methods to corroborate and further scrutinize these robust methods. In table 1 in the “results” section, these results are summarized.

1.1.1 Weiss molecular field theory

To treat the Ising model with Weiss’ molecular field theory (4 , 5), one has to replace the local fluctuating fields that a site experiences with an effective field caused by all the spins on the lattice and neglect second order fluctuations on this effective field. Quantitatively, let the spin on a site be given by $S_i = m + \delta S_i$, where m is the average magnetization of the system (assumed to be constant) and δS_i is the fluctuation that a single spin undergoes about m (assumed to be small). It was assumed that the ground state is ferromagnetic (to minimize Eq. 1, all spins have to be pointing in the same direction), so $m \neq 0$ for a certain temperature interval, and the magnetic moments of the spins of each site are taken along the z axis. So, substituting this into the Ising Hamiltonian given

by Eq. 1 with $q = 2$:

$$\begin{aligned}
\mathcal{H}_{mf} &= -J \sum_{\langle i,j \rangle} (m + \delta S_i)(m + \delta S_j) \\
&\approx -J \sum_{\langle i,j \rangle} (m^2 + m\delta S_i + m\delta S_j) \\
&= -J \sum_{\langle i,j \rangle} (m^2 + m(S_i - m) + m(S_j - m)) \\
&= \frac{Jm^2 z N}{2} - Jmz \sum_{i=1}^N S_i ,
\end{aligned} \tag{3}$$

where N is the total number of sites on the lattice and $z = 4$ is the coordination number (amount of neighbors a single site has)¹. Now, the partition function from mean-field is given by (using the notation $\beta \equiv 1/(k_B T)$):

$$\mathcal{Z}_{mf} = \left(e^{-\beta \frac{Jm^2 z}{2}} 2 \cosh(\beta Jmz) \right)^N . \tag{4}$$

Since the free energy is $\mathcal{F}_{mf} = -k_B T \ln(\mathcal{Z}_{mf})$, the self-consistency condition is given by minimizing this with respect to the magnetization:

$$\frac{\partial \mathcal{F}_{mf}}{\partial m} = 0 \implies m = \tanh(\beta Jmz) . \tag{5}$$

The critical temperature is obtained by numerically solving this transcendental equation. This is done by plotting $y = m$ and $y = \tanh(\beta Jzm)$ (*c.f.* Fig. 1), varying βJz and noting that if $\beta Jz < 1$, then the only solution that minimizes the free energy F is the origin, where $m = 0$. Otherwise, if $\beta Jz > 1$, then two other stable solutions exist: $m = \pm m_0$, corresponding to the spontaneous breaking of the Z_2 symmetry². So the critical temperature must be the limit of these two regimes, hence $\beta Jz = 1 \implies T_c = Jz/k_B$. Since the coordination number is 4, the critical temperature estimated with mean-field theory is $T_c = 4$, in units of J/k_B .

¹ The $\frac{1}{2}$ in the first term in Eq. 3 is present to avoid double counting, since the sum is over all sites and all of their neighbors.

² Since $\pm m_0$ both have the same energy, as given by Eq. 1, but the systems considered are infinitely large, only one of these values must be chosen for $T < T_c$ and the Z_2 symmetry is spontaneously broken.

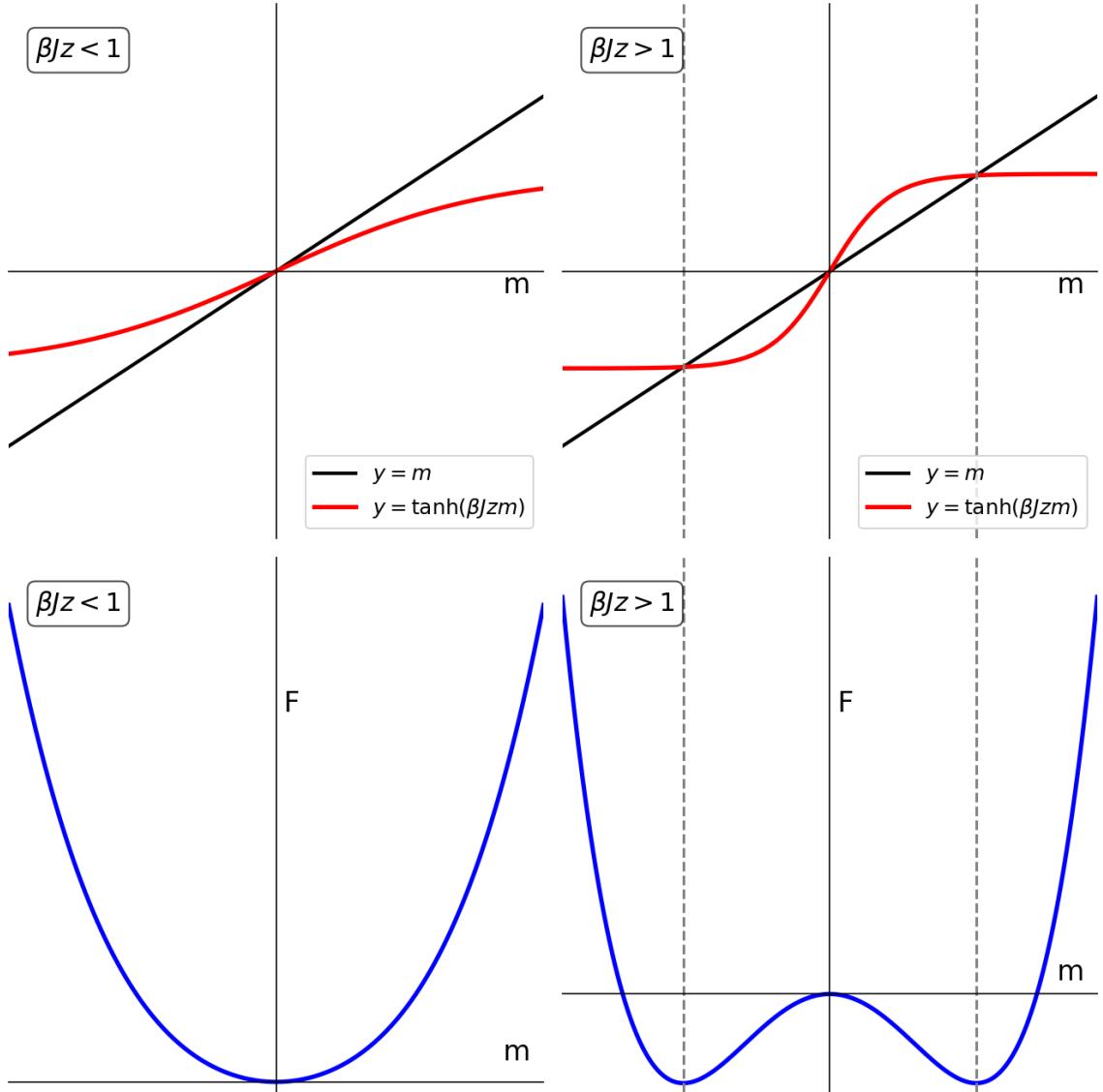


Figure 1 – Numerical solutions to Eq. 5. There are two regimes for βJz : when $\beta Jz < 1$, i.e., $T > T_c$, then the only magnetization which minimizes the free energy F is $m = 0$, when $\beta Jz > 1$, i.e., $T < T_c$, then the magnetizations which minimize the free energy F are $m = \pm m_0 \neq 0$.

Substituting Jz/k_B for T_c , defining the reduced temperature $t := (1 - T_c/T)$ and expanding the magnetization near T_c with Taylor series, from Eq. 5:

$$m = \tanh\left(m \frac{T_c}{T}\right) \approx \left(m \frac{T_c}{T}\right) - \frac{1}{3} \left(m \frac{T_c}{T}\right)^3 \implies m \sim \begin{cases} \sqrt{3(-t)}, & T \lesssim T_c \\ 0, & T > T_c \end{cases}. \quad (6)$$

The magnetization critical exponent β is defined by the relation near T_c : $m \sim | - t |^\beta$. Therefore, $\beta = 1/2$.

Analogously, the magnetic susceptibility is given by the derivative of the magnetization (given by Eq. 5) with respect to a magnetic field B that is initially added, but is

later set to zero:

$$\chi = \frac{\partial}{\partial B} \tanh \left(m \frac{T_c}{T} + \frac{B}{k_B T} \right) \Big|_{B=0} \implies \chi k_B T_c \sim \begin{cases} 4 \exp(-2T_c/T), & T \ll T_c \\ \frac{1}{2|t|}, & T \lesssim T_c \\ \frac{1}{|t|}, & T > T_c \end{cases}. \quad (7)$$

Near T_c , the magnetic susceptibility was given by its Taylor series, while, for $T \ll T_c$, the definition of the hyperbolic tangent was used: $\tanh(x) \triangleq (1 - \exp(-2x))/(1 + \exp(-2x))$, with $x = T_c/T$. The magnetic susceptibility critical exponent γ is defined by the relation near T_c : $\chi \sim |t|^{-\gamma}$. Therefore, $\gamma = 1$.

Finally, the specific heat is given the derivative of the internal energy with respect to the temperature, which is precisely:

$$C_v = k_B \beta^2 \frac{\partial^2}{\partial \beta^2} \ln(\mathcal{Z}_{mf}) \implies C_v/(k_B N) \sim \begin{cases} 4(T_c/T)^2 \exp(-2T_c/T), & T \ll T_c \\ \frac{3}{2}, & T \lesssim T_c \\ 0, & T > T_c \end{cases}. \quad (8)$$

The specific heat critical exponent α is defined by the relation near T_c : $C_v \sim |t|^{-\alpha}$. Therefore $\alpha = 0$, *i.e.*, mean field theory predicts that the specific heat is discontinuous at T_c . Even though this is the only critical exponent that mean field theory correctly predicts the value of for the 2D Ising model, in the exact solution, the specific heat actually diverges logarithmically.

Diagrammatically, these observables are shown in Fig. 2. The magnetization was actually obtained by numerically solving Eq. 5:

$$m = \tanh(\beta J m z) = \tanh \left(m \frac{T_c}{T} \right) \implies \frac{T}{T_c} = \frac{m}{\operatorname{arctanh}(m)}. \quad (9)$$

From Eq. 9, T/T_c was calculated for a grid of evenly spaced m in the interval $(0, 1)$. Then, the values T/T_c were interpolated (because now the temperature grid is irregularly spaced) and the inverse relation was plotted: m versus T/T_c .

In Fig. 2, the linear behavior of the specific heat for $T < T_c$, but not $T \ll T_c$, stems from expanding $\tanh(x) \approx x - x^3/3 + 2x^5/15$ in the magnetization, then substituting this in the internal energy. The first two terms give the constant behavior of C_v very close to T_c . But, as T ventures below T_c even more, this approximation isn't as good, and the third term needs to be incorporated, which gives the linear behavior of C_v (until a certain point. Then the exponential character takes over).

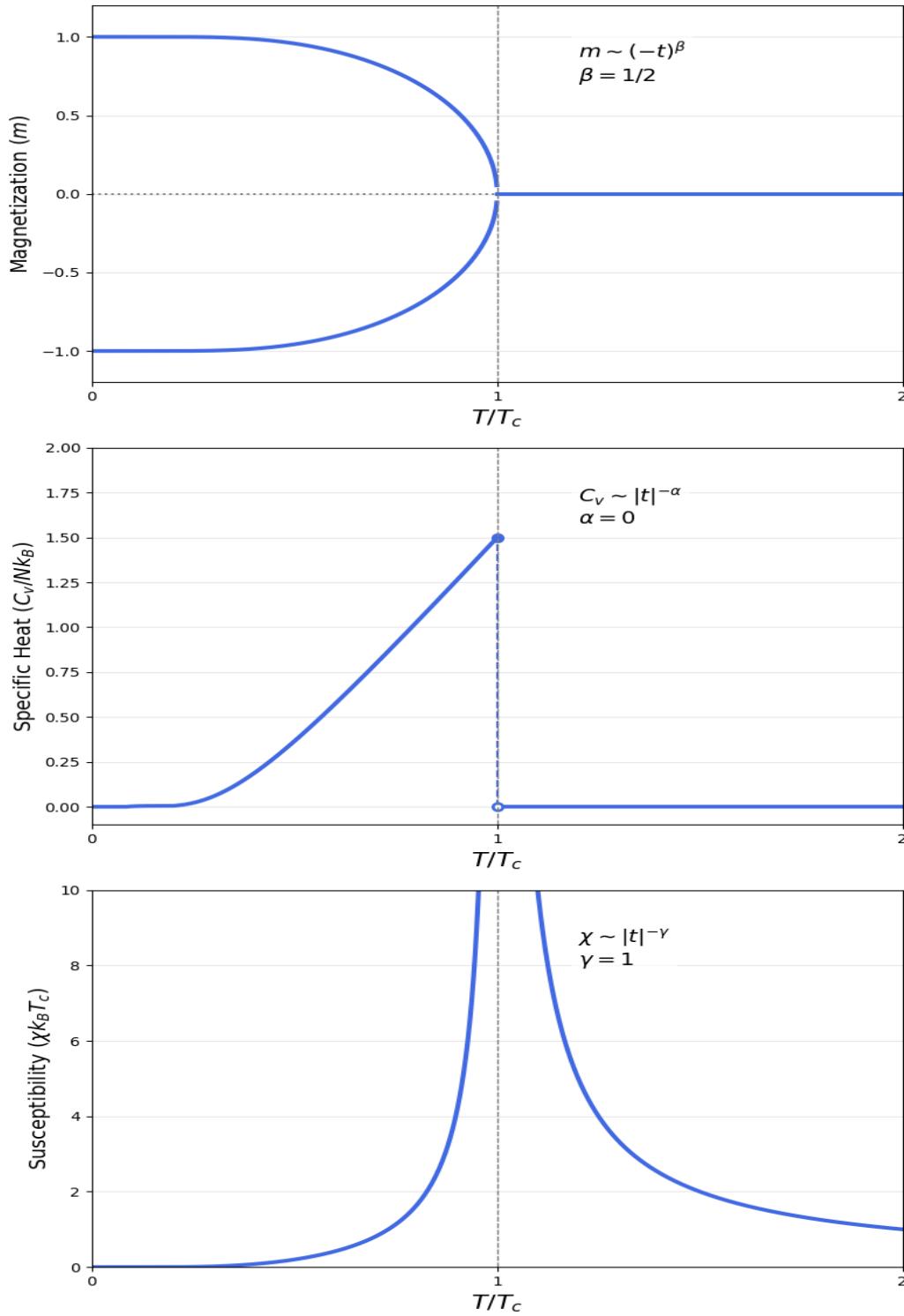


Figure 2 – Magnetization, specific heat and magnetic susceptibility, respectively, for the 2D Ising model as predicted by mean-field theory.

To calculate the correlation length critical exponent ν and anomalous critical exponent η , it's convenient to work with the free energy. Instead of using the free energy from \mathcal{Z}_{mf} directly, now would be a good time to turn to the Ginzburg-Landau theory, which already starts with an expression for the free energy.

1.1.2 Ginzburg-Landau mean-field theory

A generalization of Weiss' molecular field theory is given by Landau's ansatz, Eq. 10, which assumes that the functional form of the free energy $\mathcal{F}_{mf} = \mathcal{F}_{mf}(\vec{\phi}, T, B, N, \dots)$ can be expanded in powers of the order parameter $\vec{\phi}(\vec{x})$ near the critical temperature (2, 3):

$$\mathcal{F}_{mf} = \int_{\text{space}} d\vec{x} \left[\vec{B} \cdot \vec{\phi}(\vec{x}) + Ct \left(\vec{\phi}(\vec{x}) \cdot \vec{\phi}(\vec{x}) \right) + E \left(\vec{\phi}(\vec{x}) \cdot \vec{\phi}(\vec{x}) \right)^2 + g \left| \nabla \vec{\phi}(\vec{x}) \right|^2 + \dots \right], \quad (10)$$

where \vec{B} is the external magnetic field, $C = C(T, B, N, \dots)$, $E = E(T, B, N, \dots)$, $g = g(T, B, N, \dots)$ are all functions of the temperature, external field, number of particles etc. The order parameter is defined as an observable specific to a certain system that is identically zero for $T > T_c$ and nonzero for $T < T_c$. In Landau's theory, the order parameter is assumed to be such that it minimizes the free energy functional.

The integral in Eq. 10 is to take into account the fluctuations of the order parameter in space. The constant term in the expansion is already assumed to be zero without loss of generality (energy will be measured with respect to this constant). Apart from the first term owing to the external magnetic field, there are only even terms in Eq. 10 to respect the inversion symmetry of the Hamiltonians used in Eqs. 1 and 2: $(\vec{\phi} \rightarrow -\vec{\phi}) \implies (\mathcal{F}_{mf} \rightarrow \mathcal{F}_{mf})$. This occurs because of the dot product in both hamiltonians, which cancels two negative signs. In general however, the terms allowed in Eq. 10 depends on the symmetries of the system and those broken by the order parameter. Finally, the gradient term is the simplest term that can be added to take into account the energy differences that can occur due to fluctuations in space of the order parameter (it's squared to respect the aforementioned symmetries).

Eq. 10 also encompasses the results of Weiss' molecular field theory because if the free energy is assumed to depend on only the average value of the order parameter ($\vec{\phi}(\vec{x}) \rightarrow \vec{\phi}$), then the integral in Eq. 10 is absolved, along with any terms containing derivatives of the order parameter in space. Identifying ϕ with the magnetization m and letting the magnetic field point in the same direction as m , the first few terms in the free energy then become: $\mathcal{F}_{mf} = B(T)m + C(T)tm^2 + E(T)m^4$. The dependence of B, C, E on variables other than T have been omitted for brevity. Minimizing this free energy with respect to m :

$$\frac{\partial \mathcal{F}_{mf}}{\partial m} = 0 \implies B(T) + 2C(T)tm + 4E(T)m^3 = 0. \quad (11)$$

Let the magnetic field $B(T)$ be zero. Isolating m yields the magnetization critical exponent for $T \lesssim T_c$ ($t < 0$):

$$2C(T)tm + 4E(T)m^3 = 0 \implies m = 0, m = \pm \sqrt{\frac{-C(T)t}{2E(T)}} \sim (-t)^{1/2} \therefore \beta = 1/2. \quad (12)$$

Isolating the magnetic field $B(T)$ in Eq. 11 and expanding $C(T)$ in a power series near T_c as $C(T) = \sum_{k \geq 0} C_k t^k$, then the magnetic susceptibility for $T \gtrsim T_c$ will be (using $m = 0$):

$$\chi := \frac{\partial m}{\partial B} \Big|_{B=0} = \frac{1}{\frac{\partial B}{\partial m}} \Big|_{B=0} = \frac{-1}{2C(T)t + 12E(T)m^2} \approx \frac{-1}{2C_0 t} \sim \frac{1}{|t|} \quad \therefore \gamma = 1 . \quad (13)$$

For $T \lesssim T_c$, using Eq. 12 and expanding $E(T)$ in a power series near T_c too:

$$\chi = \frac{-1}{2C(T)t + 12E(T)m^2} \approx \frac{-1}{2C_0 t + 12E_0(t)} \sim \frac{1}{|t|} \quad \therefore \gamma = 1 . \quad (14)$$

The specific heat critical exponent is given by (using the fact that near T_c , $\partial T \approx T_c \partial t$ ³):

$$C_v \triangleq T \frac{\partial S}{\partial T} = -T \frac{\partial^2 \mathcal{F}_{mf}}{\partial T^2} \approx -\frac{T}{T_c^2} \frac{\partial^2 \mathcal{F}_{mf}}{\partial t^2} \approx \begin{cases} (\text{const1}) T, & t \gtrsim 0 \\ (\text{const2}) T, & t \lesssim 0 \end{cases} \quad \therefore \alpha = 0 \text{ (discontinuity)}, \quad (15)$$

where (const1) and (const2) are different constants.

When the fluctuations are taken into account, then it is possible to calculate the spatial spin correlation function $C(\vec{x})$, given by Eq. 16:

$$C(\vec{x}) := \langle \delta \vec{S}_{\vec{x}_0} \cdot \delta \vec{S}_{\vec{x}_0 + \vec{x}} \rangle . \quad (16)$$

Let $\vec{\phi}(\vec{x}) = \bar{\phi} + \delta \vec{\phi}(\vec{x})$, where $\bar{\phi}$ is the mean value of the order parameter and $\delta \vec{\phi}(\vec{x})$ is the fluctuation of this order parameter (assumed to be small). Now set $B = 0$ (no external field), $T > T_c$ (so $\bar{\phi} = 0$), neglect terms on the order of $\delta \phi^4$ (and higher) and replace the order parameter with its Fourier series in equation 10:

$$\delta \vec{\phi}(\vec{x}) = \sum_{\vec{k}} e^{i\vec{k} \cdot \vec{r}} \delta \phi_{\vec{k}} \implies F \approx \sum_{\vec{k}} [C_0 t - gk^2] (\delta \vec{\phi}_{\vec{k}})^2 . \quad (17)$$

This last implication used the fact that $\delta(k) \propto \int e^{ikx} dx$. From the equipartition theorem, each term in the sum is an independent quadratic degree of freedom, so $\langle (\delta \vec{\phi}_{\vec{k}})^2 \rangle \sim 1/(C_0 t - gk^2)$. However, $\langle (\delta \vec{\phi}_{\vec{k}})^2 \rangle$ is precisely equal to the Fourier transform of $C(\vec{x})$. Since at $T \neq T_c$, the Fourier transform of $C(\vec{x}) \sim e^{-x/\xi}$ (definition of correlation length ξ) is $1/(\xi^{-2} + k^2)$, then, by comparison,

$$1/(C_0 t - gk^2) \sim 1/(\xi^{-2} + k^2) , \quad \xi \sim |t|^{-\nu} \quad \therefore \nu = 1/2 , \quad (18)$$

and at $T = T_c$, the Fourier transform of $C(\vec{x}) \sim 1/x^{d-2+\eta}$ (η is the anomalous critical exponent and d is the dimension of the system. This is the definition of η) is $1/k^{2-\eta}$, then,

$$1/(0 - gk^2) \sim 1/k^{2-\eta} \quad \therefore \eta = 0 . \quad (19)$$

³ This is because $t = 1 - T_c/T = T_c/T(T - T_c)/T_c \approx (T - T_c)/T_c$.

It should be worth noting that even though the 2D Ising model was explicitly solved with mean-field theory, the q -state clock model and 2D XY $J_1 - J_2$ model would have the exact same critical exponents when calculated with mean-field theory. This is because of the fact that their symmetries produce a free energy functional with the same dependences on the order parameter, such that the mathematical treatment given beforehand for each critical exponent is valid for each of these systems.

With all the relevant critical exponents for this project calculated through mean-field theory, now they should be compared with their exact results for the Ising model. In doing so, it will be shown that these critical exponents don't agree well with their known analytical results (*c.f.* table 1). In fact, mean-field theory can't even predict the new quasi-long-range order phase that emerges from $q \geq 5$ in the q -state clock model. This happens because near phase transitions (and inside the quasi-long-range order phase), the main assumption in mean-field theory breaks down: fluctuations in spins become increasingly more correlated. So second order fluctuations between two spins isn't approximately zero.

However, despite these shortcomings, it is known that in some conditions, mean-field theory can give exact results. The Ginzburg criterion defines an upper critical dimension for which mean-field treatment of a model is exact when ascertaining its critical exponents. For the Ising model, this upper critical dimension is $d_c = 4$. So if $d \geq d_c$, then $k_B T \chi \ll \langle m \rangle^2$, *i.e.*, fluctuations can be neglected.

Although mean-field treatment is very easy to work with and can describe some basic properties of phase transitions, clearly it doesn't paint the whole picture. Therefore, better methods have to be used to obtain critical exponents of systems below their upper critical dimension. One versatile and putative way is through Monte Carlo methods.

2 METHODS

2.1 Monte Carlo algorithms

The objective of the following Monte Carlo algorithms is to simulate the random thermal fluctuation of the system from state to state over the course of an experiment. This is done by choosing the rates for transitions from one state to another, such that the Boltzmann distribution is satisfied. The Boltzmann distribution is given in Eq. 20. It states that for a system in thermal equilibrium with a reservoir at temperature T , the equilibrium occupation probability of a microstate μ is P_μ .

$$P_\mu := \frac{e^{-\frac{E_\mu}{k_B T}}}{Z}, \quad Z := \sum_\mu e^{-\frac{E_\mu}{k_B T}}. \quad (20)$$

In practice, the most common way to satisfy the Boltzmann distribution is through Markov processes. A Markov process is a sequence of possible events, where the probability of each event occurring is time-independent and depends only on the immediately preceding

event. Moreover, a Markov process is chosen specially so that when it is run for long enough (starting from any state), it will eventually produce states which appear with probability given by the Boltzmann distribution. In order to satisfy this, three conditions must be met:

1. **Ergodicity:** In principle, the Markov process should be able to reach any state from any other one given enough time ⁴.
2. **Detailed balance:** $\frac{P(\mu \rightarrow \nu)}{P(\nu \rightarrow \mu)} = \frac{P_\nu}{P_\mu} = e^{-\beta(E_\nu - E_\mu)}$, where $P(\mu \rightarrow \nu)$ is the transition probability from microstate μ to microstate ν . The first equality states that on average, the system should go from state ν to μ just as often as it goes from state μ to ν ⁵. The second equality simply imposes a Boltzmann distribution by using equation 20.
3. **Normalization:** $\sum_\nu P(\mu \rightarrow \nu) = 1$, i.e., given that the system is in the state μ , it either has to go somewhere else, or stay where it is.

Instead of defining the transition probabilities per se, it's common to let $P(\mu \rightarrow \nu) := g(\mu \rightarrow \nu)A(\mu \rightarrow \nu)$, where $g(\mu \rightarrow \nu)$ is the probability of generating state ν given state μ ("selection probability") and $A(\mu \rightarrow \nu)$ is the probability of accepting state ν given that the algorithm was in state μ and generated state ν ("acceptance ratio"). For example, given that the system is in microstate μ , the probability that it will go to ν is calculated in the following algorithms by first generating the microstate ν with a certain probability $g(\mu \rightarrow \nu)$, then it will actually change the system to be in ν with probability $A(\mu \rightarrow \nu)$.

A pragmatic Monte Carlo algorithm will always aim to make the selection probabilities as close to the actual transition probabilities and the acceptance ratio as close to unity as possible. The latter can be explained by considering the unfathomably large number of microstates a commonly simulated system can have. For example, a 10 by 10 system with the Ising Hamiltonian will have roughly ten million more microstates in its phase space than there are known stars in the observable universe! If the q -state clock model and/or larger systems were considered, there would be exponentially more microstates. So, since there's no shortage of accessible microstates, it wouldn't be very efficient if an algorithm decides to just stay put in only a couple of them: ideally it would be constantly moving through phase space to get the best possible estimates. This way, the configuration space is traversed efficiently and correctly.

⁴ This is because even though the probability P_μ of a microstate is exponentially small, it's never zero.

⁵ This doesn't mean that $P(\mu \rightarrow \nu) = P(\nu \rightarrow \mu)$, because the rate that the system goes from microstate μ to ν is the probability of the system first **being** in microstate μ (i.e., P_μ) times the probability of it going to ν ($P(\mu \rightarrow \nu)$), given that it is currently in μ .

2.1.1 Metropolis

The Metropolis algorithm (8) works in the following way: given a system in any initial state, a spin is chosen randomly and a new state for it is chosen (different from the one it's in now). If this new state decreases the energy of the system, then the spin is put into this new state. If this new state increases the energy of the system, then the spin is put into this new state with probability equal to $e^{-\beta\Delta E}$, where ΔE is the positive energy difference that the system would incur if the spin were changed to the new state. If there are N total spins on the lattice, then this whole process is normally performed N times, which would constitute a single “Monte Carlo step” (or “MC step”).

In general, single-spin-flip dynamics (such as the Metropolis algorithm) ensure ergodicity is satisfied because any configuration of spins can reach any other one by just successively flipping spins one by one. Furthermore, the normalization condition is also satisfied because some state has to be chosen for the next step (even if it's the same state). Finally, detailed balance is satisfied because:

$$\frac{P(\mu \rightarrow \nu)}{P(\nu \rightarrow \mu)} = \frac{g(\mu \rightarrow \nu)A(\mu \rightarrow \nu)}{g(\nu \rightarrow \mu)A(\nu \rightarrow \mu)} = \frac{A(\mu \rightarrow \nu)}{A(\nu \rightarrow \mu)} = \frac{e^{-\beta(E_\nu - E_\mu)}}{1} = e^{-\beta(E_\nu - E_\mu)}. \quad (21)$$

The selection probabilities cancel because each configuration has the same probability of being selected (namely, $1/N$ if it differs from the current configuration by one spin, or 0 if it doesn't). Also, it was assumed that state μ has less energy than state ν , but if it were the other way around, the last equality would read $1/\exp(-\beta(E_\mu - E_\nu)) = \exp(-\beta(E_\nu - E_\mu))$.

2.1.2 Heat bath

The heat bath algorithm (9) works by randomly choosing a spin i and letting it take on a new value, say s_n , with probability $p_n \propto \exp(-\beta E_n)$, where E_n is the energy of site i with its neighbors when spin i is in state s_n . The proportionality constant is the reciprocal of the sum (or integral, if the spins are continuous) of the total number of possibilities $\exp(-\beta E_n)$ can have when the neighbors of site i are fixed.

If the q -state clock model is considered, this proportionality constant, say \mathcal{Z}_i^{-1} , can be calculated exactly for a site i :

$$\mathcal{Z}_i^{-1} = \sum_{j=0}^{q-1} \exp(-\beta E_{i,j}) \quad , \quad (22)$$

where $E_{i,j}$ is the energy of site i with its neighbors given that site i has spin pointing in the direction j (out of all q possible directions it could point in). The subscript i for \mathcal{Z}_i^{-1} is necessary because it changes value depending on the configuration of the neighbors of site i .

If the XY $J_1 - J_2$ model is considered, then, since the spins are continuous variables, this proportionality constant can't be calculated exactly, since the sum in Eq. 22 will have infinitely many terms. To proceed, from Eq. 2, consider expressing the energy of a site i with its neighbors by the dot product of its spin \vec{S}_i with the effective magnetic field \vec{h}_i that it feels because of its neighbors:

$$\mathcal{H}_i = J_1 \sum_{j \in \text{NN}} \vec{S}_i \cdot \vec{S}_j + J_2 \sum_{j \in \text{NNN}} \vec{S}_i \cdot \vec{S}_j = \vec{S}_i \cdot \left(J_1 \sum_{j \in \text{NN}} \vec{S}_j + J_2 \sum_{j \in \text{NNN}} \vec{S}_j \right) \equiv \vec{S}_i \cdot \vec{h}_i , \quad (23)$$

where $\sum_{j \in \text{NN}}$ and $\sum_{j \in \text{NNN}}$ are supposed to represent the sums over only nearest and next-nearest neighbors of site i . This is to differentiate it from the notation used in Eq. 2 which contains the sums over all sites **and** over all their nearest and next-nearest neighbors.

Since $\vec{S}_i = (\cos(\theta_i), \sin(\theta_i))$, then $\vec{h}_i = (h_x, h_y)$, where:

$$\begin{cases} h_x = J_1 \sum_{j \in \text{NN}} \cos(\theta_j) + J_2 \sum_{j \in \text{NNN}} \cos(\theta_j) , \\ h_y = J_1 \sum_{j \in \text{NN}} \sin(\theta_j) + J_2 \sum_{j \in \text{NNN}} \sin(\theta_j) . \end{cases} \iff \begin{cases} h_i = \sqrt{h_x^2 + h_y^2} , \\ \phi = \arctan\left(\frac{h_y}{h_x}\right) . \end{cases} \quad (24)$$

The desired distribution to sample is:

$$p_n \propto \exp(-\beta E_n) = \exp(-\beta \vec{S}_i \cdot \vec{h}_i) = \exp(\beta h_i \cos(\theta_i - \phi)) . \quad (25)$$

In analogy to Eq. 22, the proportionality constant is $\mathcal{Z}_i^{-1} = \int_0^{2\pi} \exp(-\beta h_i \cos(\theta_i - \phi)) d\theta_i$. This integral is the modified Bessel function of the first kind of order zero (10). There is no closed form expression for this integral in terms of elementary functions⁶, so it is much more practical to use rejection sampling on Eq. 25 directly (22, p. 365). This can be done by taking the ratio of p_n with the maximum value of p_n allowed (this happens whenever $\cos(\theta_i - \phi) = \pm 1$). Calculating this ratio $W(\theta_i)$:

$$W(\theta_i) = \frac{p_n}{p_{\max}} = \left(\frac{e^{-\beta h_i \cos(\theta_i - \phi)}}{\mathcal{Z}_i} \right) \left(\frac{e^{-\beta h_i (-1)}}{\mathcal{Z}_i} \right)^{-1} = \exp(-\beta h_i (\cos(\theta_i - \phi) + 1)) . \quad (26)$$

To sample this distribution, a random angle θ between 0 and 2π should be generated. Then a random number r in the range $[0, 1)$ should be generated. If $r \leq W(\theta)$, then spin i is set to have the angle θ , else a new random angle is chosen along with another random number r and this whole process is repeated until spin i is set to some new state.

The heat bath algorithm is also a single-spin-flip algorithm, so, analogously to the arguments made for the Metropolis algorithm, the ergodicity principle is satisfied. Since some state is chosen as the next configuration too, the normalization condition is also satisfied. As for the detailed balance condition, let the state μ of the system have energy E_μ and a site i be in the spin state s_μ . Analogously, let the state ν have energy E_ν and the site i be in the spin state s_ν . Then, because the energy of the system E_μ (or E_ν) can

⁶ The only representations known for this function are in terms of infinite series, such as: $\mathcal{Z}_i^{-1} = \sum_{k \geq 0} \left(\frac{x^k}{2^k (k!)} \right)^2$.

always be written as the sum of the energy between a site and its neighbors E_N and the energy of the rest of the system E_{rest} : $E_\mu = E_N + E_{\text{rest}}$,

$$\frac{P(\mu \rightarrow \nu)}{P(\nu \rightarrow \mu)} = \frac{A(\mu \rightarrow \nu)}{A(\nu \rightarrow \mu)} = \left(\frac{e^{-\beta(E_\nu - E_{\text{rest}})}}{\mathcal{Z}_i} \right) \left(\frac{e^{-\beta(E_\mu - E_{\text{rest}})}}{\mathcal{Z}_i} \right)^{-1} = e^{-\beta(E_\nu - E_\mu)}. \quad (27)$$

Regardless of the Hamiltonian used (continuous or not), the proportionality factors always cancel, such that the Boltzmann distribution is correctly sampled.

2.1.3 Wolff

The Wolff algorithm (11) works as follows: a spin is randomly chosen at first and is added to the stack and cluster data structures. Then a random rotation amount $\delta\theta \in [0, 2\pi)$ is selected ⁷. While the stack isn't empty, pop a site and add its neighbors to the stack and cluster, if it isn't already in the cluster, with probability $P_{\text{add}} = \max(0, 1 - \exp(-\beta\Delta E_{ij}))$, where ΔE_{ij} is the energy in the bond between site i and site j after site i is rotated minus the energy in this same bond before site i is rotated. Finally, rotate every spin in the cluster by $\delta\theta$. This whole process will be defined as a MC step once it has been performed N times, where N is the total number of sites in the system.

Ergodicity is satisfied because there is always a finite chance that a single spin could constitute a cluster and thus get rotated. Therefore, analogously for single-spin-flip algorithms, any other configuration could be arrived at by successively rotating single spin clusters. Normalization is also clearly satisfied too, because some state will always be chosen for the next step.

Finally, to check the detailed balance, suppose that there exist two microstates μ and ν that differ from each other only by a rotation of a cluster of spins. Let the difference in energy of these microstates be $(E_\nu - E_\mu) \equiv \Delta E$. From the Hamiltonian considered (equation 1 or 2), clearly this energy difference resides on the border ∂C of the cluster of rotated spins (all spins inside the cluster maintain their same relative angle to their neighbors, and all the spins beyond nearest neighbors, or next-nearest-neighbors, to the border of the cluster aren't affected by the rotation). Let i be a spin in ∂C and j be its neighbor that is not a part of the cluster. In going from μ to ν , j had to have been considered for addition to the cluster, but rejected (possibly many times). So, define ΔE_{ij} as the energy difference the system accrued in going from state μ to state ν owing only to the bond between site i and site j . Then, by definition, $\Delta E = \sum_{\partial C} \Delta E_{ij}$. If $\Delta E_{ij} > 0$, then the rejection probability from μ to ν is $1 - P_{\text{add}} = \exp(-\beta\Delta E_{ij})$, while the rejection probability from ν to μ is unitary (because $P_{\text{add}} = 0$). Thus, for these two sites only, using the fact that the acceptance probability is always one (a new configuration is always

⁷ For the q -state clock model, this rotation is a multiple of $2\pi/q$, which is the smallest angle difference two distinct states can have. For the XY $J_1 - J_2$ model, this rotation can be made arbitrarily small.

chosen because at least one spin is always rotated),

$$\frac{P_k(\mu \rightarrow \nu)}{P_k(\nu \rightarrow \mu)} = \left(\frac{g_k(\mu \rightarrow \nu)}{g_k(\nu \rightarrow \mu)} \right) \left(\frac{A_k(\mu \rightarrow \nu)}{A_k(\nu \rightarrow \mu)} \right) = \left(\frac{\exp(-\beta(\Delta E_{ij}))}{1} \right) \left(\frac{1}{1} \right) = e^{-\beta \Delta E_{ij}}. \quad (28)$$

Analogously, if $\Delta E_{ij} \leq 0$, then $P_{\text{add}} = 0$ and the rejection probability is unitary in going from μ to ν , while the rejection probability from ν to μ is $1 - P_{\text{add}} = \exp(-\beta(-\Delta E_{ij}))$. Thus:

$$\frac{P_k(\mu \rightarrow \nu)}{P_k(\nu \rightarrow \mu)} = \left(\frac{g_k(\mu \rightarrow \nu)}{g_k(\nu \rightarrow \mu)} \right) \left(\frac{A_k(\mu \rightarrow \nu)}{A_k(\nu \rightarrow \mu)} \right) = \left(\frac{1}{\exp(-\beta(-\Delta E_{ij}))} \right) \left(\frac{1}{1} \right) = e^{-\beta \Delta E_{ij}}. \quad (29)$$

So, regardless of the sign of ΔE_{ij} , the ratios of the transition probabilities between sites i and j will have the same form (an exponential of minus beta times something). Therefore, detailed balance is satisfied by considering the product of the ratios of the transition probabilities for each spin on the border of the cluster:

$$\frac{P(\mu \rightarrow \nu)}{P(\nu \rightarrow \mu)} = \prod_k \frac{P_k(\mu \rightarrow \nu)}{P_k(\nu \rightarrow \mu)} = \prod_k \frac{g_k(\mu \rightarrow \nu)}{g_k(\nu \rightarrow \mu)} = \prod_{\partial C} e^{-\beta \Delta E_{ij}} = e^{-\beta \sum_{\partial C} \Delta E_{ij}} = e^{-\beta(E_\nu - E_\mu)}. \quad (30)$$

This proof is valid because the probability of generating all the spins to be inside of the cluster is exactly equal to the probability of rejecting all the spins on the border of said cluster. Hence **rejection** probabilities can be used in place of the **generation** probabilities. A useful analogy would be to consider the occurrence of a hypothetical event A . Say event A has a 20% chance of occurring. So instead of directly using this 20% chance of occurrence, someone could just as well say that event A has an 80% chance of **not** occurring. This is convenient for the Wolff algorithm only because the energy difference between two successively generated microstates resides on the border of the clusters of spins formed.

2.1.4 Swendsen-Wang

The Swendsen-Wang algorithm (12) is very similar to the Wolff algorithm: the entire lattice is divided into clusters (each with their own random rotation amount), which are formed following the same procedure as in the Wolff algorithm⁸ until every site is part of a cluster. Then each cluster is considered for rotation by their respective random angle with probability 1/2.

Ergodicity is satisfied because there is a finite chance of forming only single-spin clusters and then flipping just one of them in the whole lattice. This process can be repeated to arrive at any state. Also, normalization is clearly satisfied for the same reasons as before: a state will always be chosen for a step.

⁸ A spin is added to the stack, then, while the stack isn't empty, a spin is popped from the stack and its neighbors are added to the stack and cluster with probability $P_{\text{add}} = 1 - \exp(-\beta \Delta E_{ij})$ if they aren't already part of any other cluster on the lattice.

To check detailed balance, for each cluster of spins on the lattice, define an initial microstate μ' and a final microstate ν' ⁹ which differ from each other by only the potential rotation of only that one cluster. Then, applying the same arguments that prove detailed balance for the Wolff algorithm, consider the ratio of the transition probabilities for spins along the border of this one cluster:

$$\frac{P(\mu' \rightarrow \nu')}{P(\nu' \rightarrow \mu')} = \prod_k \left(\frac{g_k(\mu' \rightarrow \nu')}{g_k(\nu' \rightarrow \mu')} \right) \left(\frac{A_k(\mu' \rightarrow \nu')}{A_k(\nu' \rightarrow \mu')} \right) = \prod_{\partial C} \left(e^{-\beta \Delta E_{ij}} \right) \left(\frac{1/2}{1/2} \right) = e^{-\beta(E'_\nu - E'_\mu)}. \quad (31)$$

Since the lattice is divided into multiple clusters, consider an initial microstate α_0 where no cluster is rotated, then apply this argument to a single cluster to potentially arrive at a new microstate α_1 . Then consider another cluster and repeat this logic until every cluster has been considered for rotation and a (possibly new) microstate α_N is the end result after applying the Swendsen-Wang algorithm once. Then, since at each step detailed balance is satisfied, for the entire algorithm, the ratio of transition probabilities is just the product of the ratio of the transition probabilities at each step:

$$\frac{P(\alpha_0 \rightarrow \alpha_N)}{P(\alpha_N \rightarrow \alpha_0)} = \prod_{i=0}^{N-1} \frac{P(\alpha_i \rightarrow \alpha_{i+1})}{P(\alpha_{i+1} \rightarrow \alpha_i)} = \prod_{i=0}^{N-1} e^{-\beta(E_{\alpha_{i+1}} - E_{\alpha_i})} = e^{-\beta(E_{\alpha_N} - E_{\alpha_0})} \equiv e^{-\beta(E_\nu - E_\mu)}. \quad (32)$$

2.1.5 Choosing an algorithm

For most Hamiltonians, the best choice of an algorithm isn't obvious at all. Each one has their strengths and weaknesses. For example, it will be shown that for the q -state clock model, the Wolff algorithm is best, but for the XY $J_1 - J_2$ model, it's actually the Metropolis and heat bath algorithms. The ideal way to compare these algorithms is to actually implement each one and see which one takes the least amount of time to run and decorrelate successively generated microstates the best.

In general, owing to the complexity in the implementation of the Swendsen-Wang algorithm, the Wolff algorithm tends to be faster, making it the cluster-flip algorithm of choice. Analogously, since the Metropolis algorithm is significantly easier to implement than the heat bath algorithm, it usually ends up being faster and more widely used. Because of these reasons, the Metropolis and Wolff algorithms were the single-spin-flip and cluster-flip algorithms chosen to have their dynamical critical exponents calculated in simulating the Ising model.

However, the heat bath and Swendsen-Wang algorithms have their place. For instance, in some models, the heat bath algorithm tends to decorrelate the system better over the course of an MC step than the Metropolis algorithm. So even though it takes longer to run, it normally generates more independent samples. Of course now the question

⁹ which is precisely the initial microstate μ' half of the time, because now, a cluster is rotated with probability 1/2.

becomes whether this extra time per MC step that the heat bath algorithm takes makes up for these better samples. This question depends entirely on the model and implementation details. Some authors have discovered that a combination of the heat bath algorithm with others is actually the best compromise (13). Meanwhile, the Swendsen-Wang algorithm really shines in the context of parallel Monte Carlo algorithms. It has been shown in the literature that the Swendsen-Wang algorithm is easier to parallelize and more efficient than the Wolff algorithm (14) ¹⁰. Moreover, when using GPU calculation with the common unified device architecture (CUDA) to simulate the q -state clock model, the Swendsen-Wang algorithm appears to be the gold standard (15).

2.1.6 Equilibration

For all of the aforementioned algorithms, the correct way to use them is to: choose appropriate parameters for the system (such as the temperature, value of q if the q -state clock model is used or the ratio J_2/J_1 if the XY $J_1 - J_2$ model is used), impose any initial state for the spins, simulate said system for many Monte Carlo steps until equilibrium is reached and record any variable of interest over the course of many more Monte Carlo steps.

However, how can one determine if equilibrium has actually been reached? Equilibration times depend on a lot of factors, such as the: Hamiltonian, algorithm used, size of the system, definition of MC step, temperature etc. The traditional way to see if equilibrium has been reached is to just plot a desired observable over the course of many MC steps and see when a plateau has been established. To illustrate this, consider Fig. 3 which uses the heat bath algorithm to simulate the 2D XY $J_1 - J_2$ model and plots the nematic order parameter (Eq. 52) over the course of 3000 MC steps starting from an initial configuration with spins pointing in random directions. The parameters used were: $k_B T/J_1 = 0.65$, $J_2/J_1 = 0.8$ and $L = 32, 64$. It should be noted that eventually equilibrium is in fact reached, and larger systems require longer equilibration times. This latter point is normally what prohibits Monte Carlo simulations from simulating very large systems: equilibrium is very difficult to reach, so data collection will be precipitous and yield specious results.

¹⁰ In short, this is because the Swendsen-Wang algorithm can identify all clusters on the lattice simultaneously, while the Wolff algorithm has to grow a single cluster outward, assigning the parallel work to only the border of this cluster which makes attempts at parallelization awkward.

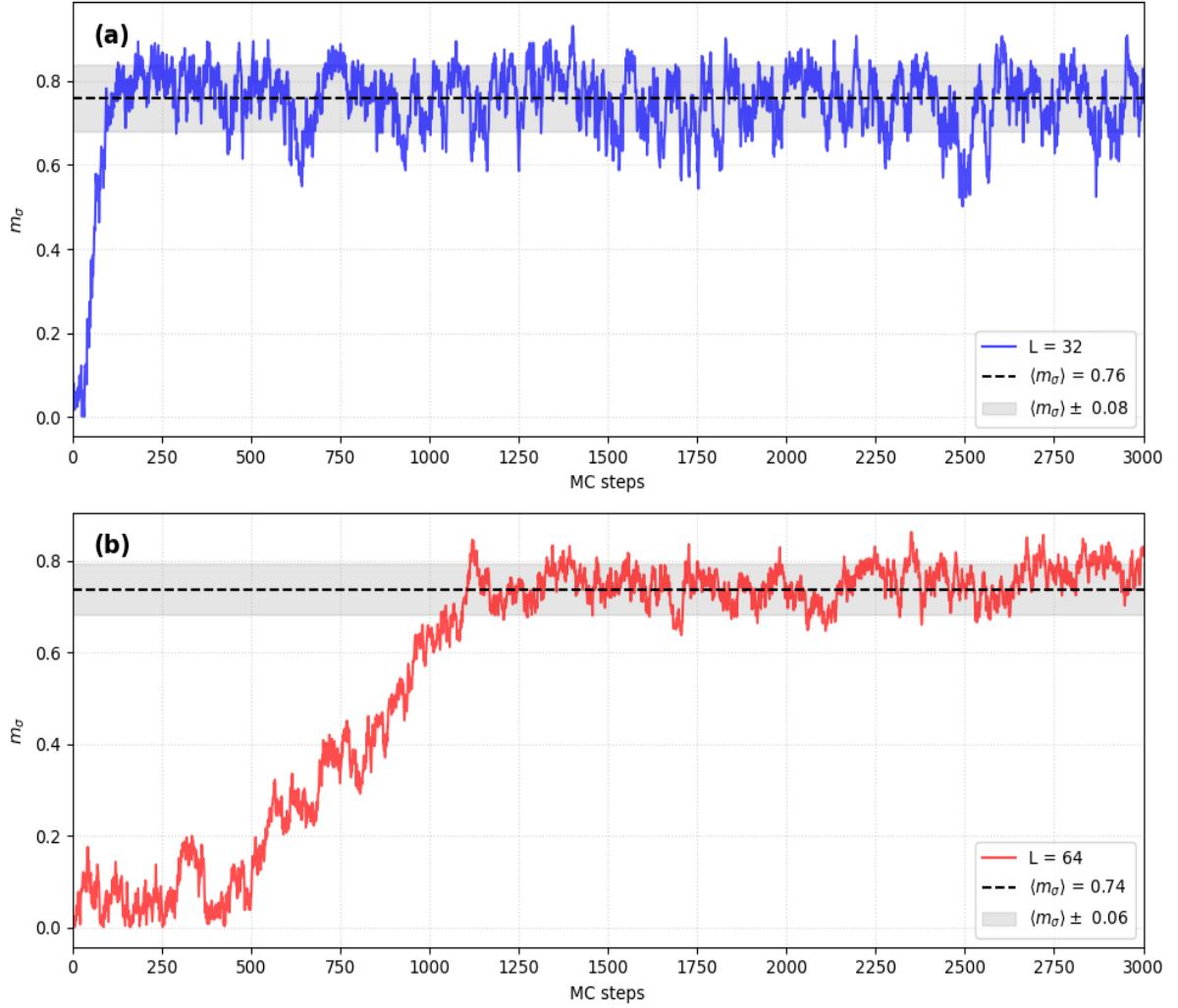


Figure 3 – Nematic order parameter m_σ (Eq. 52) over the course of 3000 MC steps on the 2D XY $J_1 - J_2$ model ($k_B T/J_1 = 0.65$, $J_2/J_1 = 0.8$ and $L = 32, 64$) simulated with the heat bath algorithm on systems of size (a) $L = 32$ and (b) $L = 64$. The system initially had every spin oriented in a random direction. The dotted line and shaded region is the average value of m_σ and its standard deviation calculated after 3000 MC steps up to 10^4 MC steps.

2.2 Relevant variables

2.2.1 Magnetizations and energy

The energy E , magnetization per site m and “angular magnetization” m_ϕ (20) are given by Eqs. 1 (or 2), 33 and 34 respectively (supposing there are N total sites in the system).

$$m = \frac{1}{N} \sqrt{\left(\sum_{i=1}^N \cos \theta_i \right)^2 + \left(\sum_{i=1}^N \sin \theta_i \right)^2}, \quad (33)$$

$$m_\phi = \cos(q\phi) , \quad \phi = \tan^{-1} \left(\frac{\sum_{i=1}^N \sin \theta_i}{\sum_{i=1}^N \cos \theta_i} \right) . \quad (34)$$

The variables m , m_ϕ and E are measured at each Monte Carlo step, and their K th moments are calculated according to Eq. 35 (which assumes an arbitrary variable O_i

measured at each Monte Carlo step and M total Monte Carlo steps taken).

$$\langle O^K \rangle = \frac{1}{M} \sum_{i=1}^M O_i^K . \quad (35)$$

It's worth noting that Eq. 35 actually gives the correct moments of the desired Boltzmann distribution because, even though the Boltzmann weights don't explicitly multiply each term, the appearances of each O_i in the sum occurs with the correct Boltzmann weight (as was shown for each algorithm).

2.2.2 Specific heat and magnetic susceptibility

The specific heat and magnetic susceptibility are given by Eqs. 36 and 37:

$$C_v \triangleq \frac{\partial \langle E \rangle}{\partial T} = k_B \beta^2 \frac{\partial}{\partial \beta} \left(\frac{1}{Z} \frac{\partial Z}{\partial \beta} \right) = k_B \beta^2 \left(\langle E^2 \rangle - \langle E \rangle^2 \right) , \quad (36)$$

$$\chi \triangleq \frac{\partial \langle M \rangle}{\partial B} = \frac{\partial}{\partial B} \left(\frac{1}{\beta Z} \frac{\partial Z}{\partial B} \right) \Big|_{B=0} = \beta \left(\langle M^2 \rangle - \langle M \rangle^2 \right) . \quad (37)$$

For the specific heat, the true average ¹¹ of the energy was used, which is given by: $\langle E \rangle = \sum_r E_r \exp(-\beta E_r)/Z$, where $Z = \sum_r \exp(-\beta E_r)$ and the sum is over all possible microstates. For the magnetic susceptibility, a fictitious magnetic field \vec{B} coupled to the magnetization M was surreptitiously added to the Hamiltonian just to perform the derivative of the true magnetization with respect to the magnetic field, as is required by the definition of the magnetic susceptibility (afterwards it's set to zero). Also, the true mean magnetization was used: $\langle M \rangle = \sum_r M_r \exp(-\beta E_r)/Z$, where $E_r = -\sum_{i,j} J_{i,j} \vec{S}_i \cdot \vec{S}_j - MB$.

2.2.3 Correlation functions and critical slowing down

The spatial spin correlation function is given by Eq. 16, but it can also be written in a more convenient form as in Eq. 38 (the second equality is because of the translational invariance of the system, owing to periodic boundary conditions. It also assumes uniform magnetization).

$$C(r) := \langle \vec{S}_i \cdot \vec{S}_{i+r} \rangle - \langle \vec{S}_i \rangle \cdot \langle \vec{S}_{i+r} \rangle = \langle \vec{S}_i \cdot \vec{S}_{i+r} \rangle - \langle m \rangle^2 . \quad (38)$$

The average is performed over every spin on the lattice (different i) and for every MC step. Also, since calculating this quantity is a costly operation, only two orthogonal directions for each site were considered when getting spins a distance r from each other.

As was hitherto mentioned, the correlation length $\xi(T)$ at $T \neq T_c$ is defined by Eq. 39,

$$C(r) \sim e^{-r/\xi(T)} . \quad (39)$$

¹¹ “true average” should be interpreted as the average one would obtain in the thermodynamic limit after sampling every possible microstate of the system with their correct Boltzmann weights.

Meanwhile, the temporal autocorrelation function for a given observable m measured at each MC step is given by simulating a system and recording m at each MC step for, say, M total MC steps, then calculating Eq. 40 for all available values of t (from $t = 0$ to $t = M - 1$), which corresponds to the time in MC steps.

$$C(t) := \frac{1}{M-t} \sum_{t'=0}^{M-t} m(t')m(t'+t) - \left(\frac{1}{M-t} \sum_{t'=0}^{M-t} m(t') \right) \left(\frac{1}{M-t} \sum_{t'=0}^{M-t} m(t'+t) \right). \quad (40)$$

The correlation time $\tau(T)$ is also defined by Eq. 39, except $C(r)$ becomes $C(t)$ and $\xi(T)$ becomes $\tau(T)$.

The correlation time will serve to calculate the dynamical critical exponent z for the Metropolis and Wolff algorithms, which is given by the relation: $\tau \sim \xi^z \sim |t|^{-\nu z}$ (ν is the critical exponent for the correlation length defined in Eq. 56). This exponent quantifies the critical slowing down effect that occurs near second order phase transitions. Critical slowing down is a property of Monte Carlo algorithms. It occurs when measurements of a given observable over the Monte Carlo steps become very correlated with each other, making it difficult to get independent measurements for the calculation of the moments of their distributions. To counteract this, one should let the simulation run for longer times to get more measurements (or simulate the system more in between each measurement), making the results slower to obtain, hence the name “critical slowing down”.

In general, single-spin-flip algorithms (e.g., Metropolis, heat bath) have a larger dynamical critical exponent than cluster-flip algorithms (e.g., Wolff, Swendsen-Wang), meaning that they take longer to decorrelate successive microstates generated in their Markov chains near the phase transition. Intuitively, this can be understood by considering a Monte Carlo step with a single-spin-flip algorithm near the phase transition of the Ising model: a spin inside a cluster is most likely chosen (because large clusters of spins are present) and not flipped, because it's aligned with all its neighbors, so a flip would increase the system's energy a lot. If a spin on the border of a cluster is chosen, it has a higher chance of getting flipped. Regardless of the spin chosen, the most likely scenario is that very little spins inside each cluster will get flipped and most flips occur on the border of these clusters. So successively generated microstates will look very similar to one another, the only difference being that the clusters of spins will look like they're moving. If the magnetization (or energy) is calculated, it will be practically the same for each microstate. For cluster-flip algorithms, this problem is avoided by flipping precisely these clusters at each Monte Carlo step, decorrelating measurements much faster, which, in turn, yields a smaller dynamical critical exponent.

To see how cluster-flip algorithms have the edge over single-spin-flip algorithms, consider Fig. 4 which shows the correlation time of the magnetization squared calculated for the Wolff and Metropolis algorithms in the 2D Ising model over a temperature range which includes the known critical temperature (vertical dotted line). To obtain the correlation

time, a simple linear regression was performed on the logarithm of the correlation function given by Eq. 40. The values of the dynamical critical exponents in the title were obtained through a finite-size scaling analysis, which will be explained in the “Finite-size scaling” section. Near the critical temperature, for the system of size $L \times L$, with $L = 256$, the Metropolis algorithm has correlation times that are over 10^3 larger than those obtained with the Wolff algorithm.

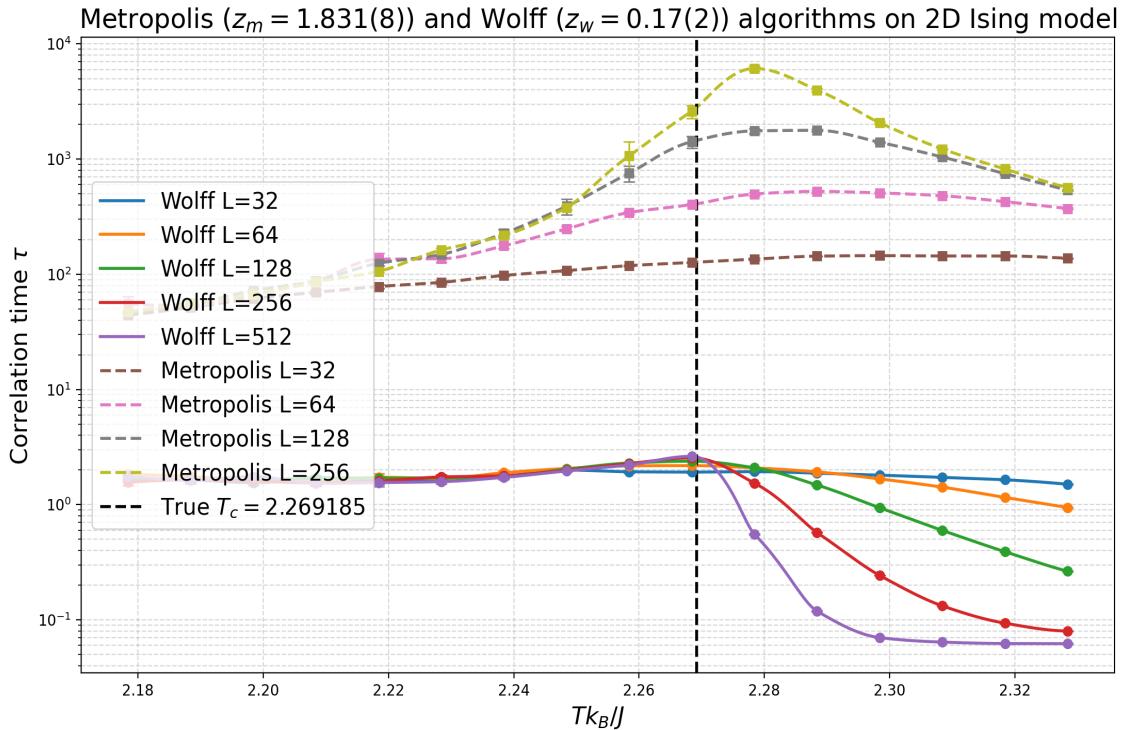


Figure 4 – Correlation time of the magnetization squared calculated for the Metropolis (dotted lines) and Wolff (solid lines) algorithms in the 2D Ising model over a temperature range that includes the known critical temperature (black vertical dotted line). L is the length of one dimension of the system. The dynamical critical exponents in the title were obtained using finite-size scaling.

2.2.4 Binder cumulants and the BKT transition

The Binder cumulants used for the q -state clock model are defined according to Eq. 41:

$$U_4 := 1 - \frac{\langle m^4 \rangle}{3\langle m^2 \rangle^2}, \quad U_\phi := 1 - \frac{\langle m_\phi^4 \rangle}{2\langle m_\phi^2 \rangle^2}, \quad (41)$$

where m is the magnetization per site and m_ϕ is the “angular magnetization” (in general, these could be any relevant order parameters). U_4 will be used to estimate the critical temperature from the long range order (LRO), or ferromagnetic (FM), phase to the disordered, or paramagnetic (PM), phase for $q = 2, 3, 4$. It will also be used to estimate the critical temperature from the quasi-long-range order (QLRO) phase to the disordered phase for $q \geq 5$. U_ϕ will be used to estimate the critical temperature for the LRO phase to the QLRO phase for $q \geq 5$.

Binder cumulants work as follows: for infinitely large systems, it's a step function with discontinuity exactly at the critical temperature. So, as the simulated system size increases, the calculated Binder cumulants will start to resemble step functions. In this process, they will inevitably intersect one another at the discontinuity, which is precisely the desired critical temperature.

To prove this, consider Figs. 5 - 7, which show the probability distributions of the magnetization vector (with components m_x and m_y) at each of the three phases: LRO, QLRO and disordered, for a given value of q and L (what really matters is the general form of the distributions, not the exact values of q and L themselves). For the arguments that follow, consider that the moments of m are calculated for large systems (not infinitely large¹²) and the moments of m_ϕ are calculated for infinitely large systems, and these moments are further calculated for an ensemble of infinitely many similar systems (at the same temperature), such that the distributions of these calculated moments follow gaussian distributions as is required by the central limit theorem.

Probability distribution of magnetization per site (LRO phase)
($q = 8$, $L = 32$, $T = 0.2 \text{ J}/k_B$)

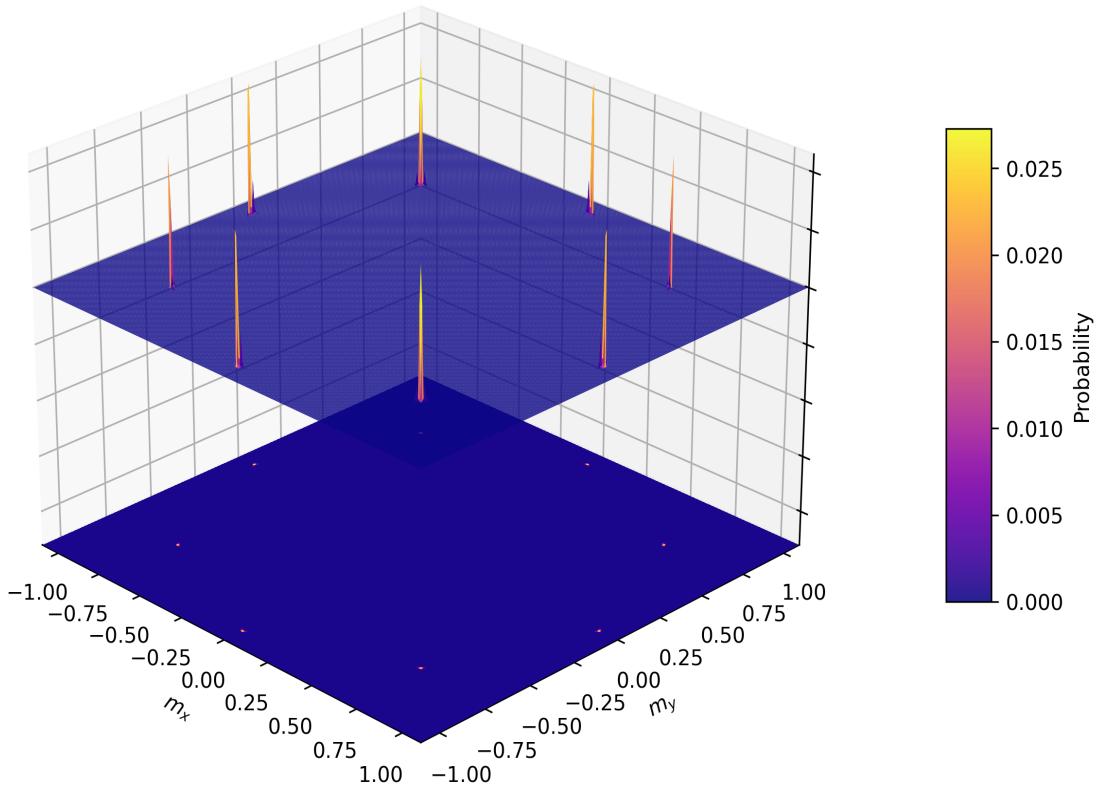


Figure 5 – Probability distribution of the magnetization vector for $q = 8$, $L = 32$, at $T = 0.2 \text{ J}/k_B$ (LRO phase).

¹² This is because the magnetization is identically zero in the QLRO phase for infinitely large systems, but not for finite size systems, so it isn't actually a good choice for an order parameter here. This reason will be explained in more detail in the “Results” section.

In the LRO phase (Fig. 5), for all values of q , the probability distributions of the magnetization vector resembles delta functions at each of the q distinct states as the size of the system increases to infinity (this is because all the spins are roughly aligned with each other). So the moments of the magnetization per site (norm of the magnetization vector) are just constants, which means that $U_4 = 1 - 1/3 = 2/3$. For m_ϕ , the angle of the magnetization vector can only be $\phi = (0)2\pi/q, (1)2\pi/q, \dots$, so $m_\phi = \cos(q\phi) = \cos(2\pi n) = 1$ (n is a natural number) and $U_\phi = 1 - 1/2 = 1/2$.

Probability distribution of magnetization per site (QLRO phase)
($q = 8, L = 32, T = 0.6 J/k_B$)

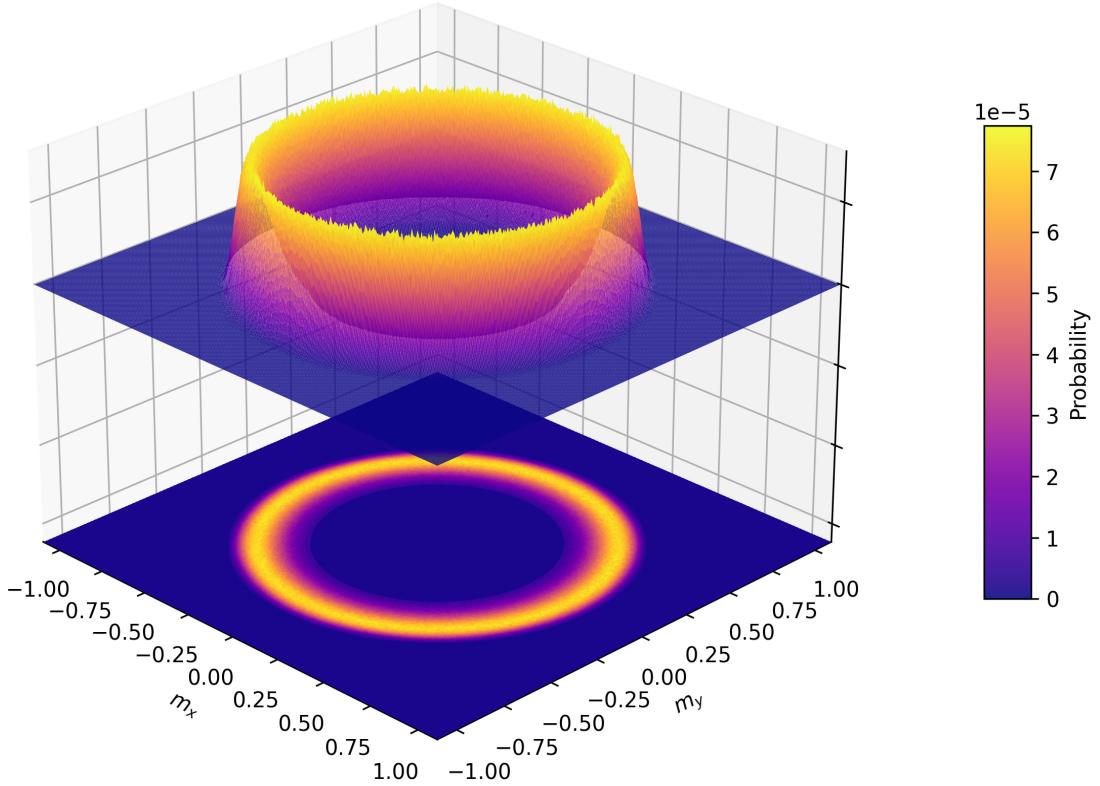


Figure 6 – Probability distribution of the magnetization vector for $q = 8, L = 32$, at $T = 0.6J/k_B$ (QLRO phase).

In the QLRO phase (Fig. 6), the magnetization vector can point in any direction with equal probability, so ϕ is an independent identically distributed (iid) random variable in the interval $[0, 2\pi]$. Therefore m_ϕ can assume any value in the range $[-1, 1]$ with a certain probability. To determine the moments of m_ϕ , from the central limit theorem, one needs to integrate over all values it can assume from the constituents in the ensemble weighted by a gaussian distribution with zero mean and standard deviation σ . This integral will be calculated over the entire XY plane to make calculations easier. This is obviously an approximation, since m_ϕ can only assume values in the unit square, but it's a good one because for very large systems, the gaussian ring in Fig. 6 will become very tall and thin,

i.e., very tightly bound around its mean, so the integrand will be extremely small outside the unit square. Using polar coordinates and the binomial expansion of $\cos(x)$:

$$\begin{aligned}
\langle m_\phi^N \rangle &\triangleq \int_0^{2\pi} \int_0^\infty \cos^N(q\phi) \frac{1}{2\pi\sigma^2} e^{-\frac{r^2}{2\sigma^2}} r dr d\phi \\
&= \frac{1}{2\pi} \int_0^{2\pi} \cos^N(q\phi) d\phi \\
&= \frac{1}{2\pi q} \int_0^{2\pi q} \cos^N(u) du \\
&= \frac{1}{2\pi} \int_0^{2\pi} \cos^N(u) du \\
&= \frac{1}{2\pi} \int_0^{2\pi} (e^{iu}/2 + e^{-iu}/2)^N du \\
&= \frac{1}{2\pi} \int_0^{2\pi} \frac{1}{2^N} \sum_{k=0}^N \binom{N}{k} e^{iu(N-2k)} du = \begin{cases} 0, & N \text{ odd.} \\ \frac{\binom{N}{N/2}}{2^N}, & N \text{ even.} \end{cases}
\end{aligned} \tag{42}$$

Thus, $U_\phi = 1 - \frac{1}{2} \frac{3/8}{(1/2)^2} = 1/4$, i.e., it suffers a discontinuity from the LRO phase ($U_\phi = 1/2$) to the QLRO phase ($U_\phi = 1/4$), making it a suitable candidate to detect the transition temperature between these two regimes for $q \geq 5$.

On the other hand, the moments of m can be obtained in a similar manner: integrate m over all the values it assumes in every ensemble, weighted by a gaussian with mean m_0 (this is roughly the radius of the ring in figure 6) and standard deviation σ :

$$\langle m^N \rangle \triangleq \int_{-\infty}^{\infty} (m - m_0)^N \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(m-m_0)^2}{2\sigma^2}} dm = \sum_{k \geq 0} \binom{N}{2k} m_0^{N-2k} \sigma^{2k} (2k-1)!! . \tag{43}$$

This last equality in Eq. 43 was taken from page 135 of reference 18. Alternatively, integration by parts could have been used to find the results for $N = 2$ and $N = 4$. Explicitly, $\langle m^2 \rangle = m_0^2 + \sigma^2$ and $\langle m^4 \rangle = m_0^4 + 6m_0^2\sigma^2 + 3\sigma^4$. Thus, $U_4 \approx 1 - 1/3 = 2/3$ ($\sigma/m_0 \ll 1$). Because U_4 is the same in the LRO and QLRO phases, it cannot determine the transition temperature therein.

Probability distribution of magnetization per site (disordered phase)
($q = 8$, $L = 32$, $T = 1.5 J/k_B$)

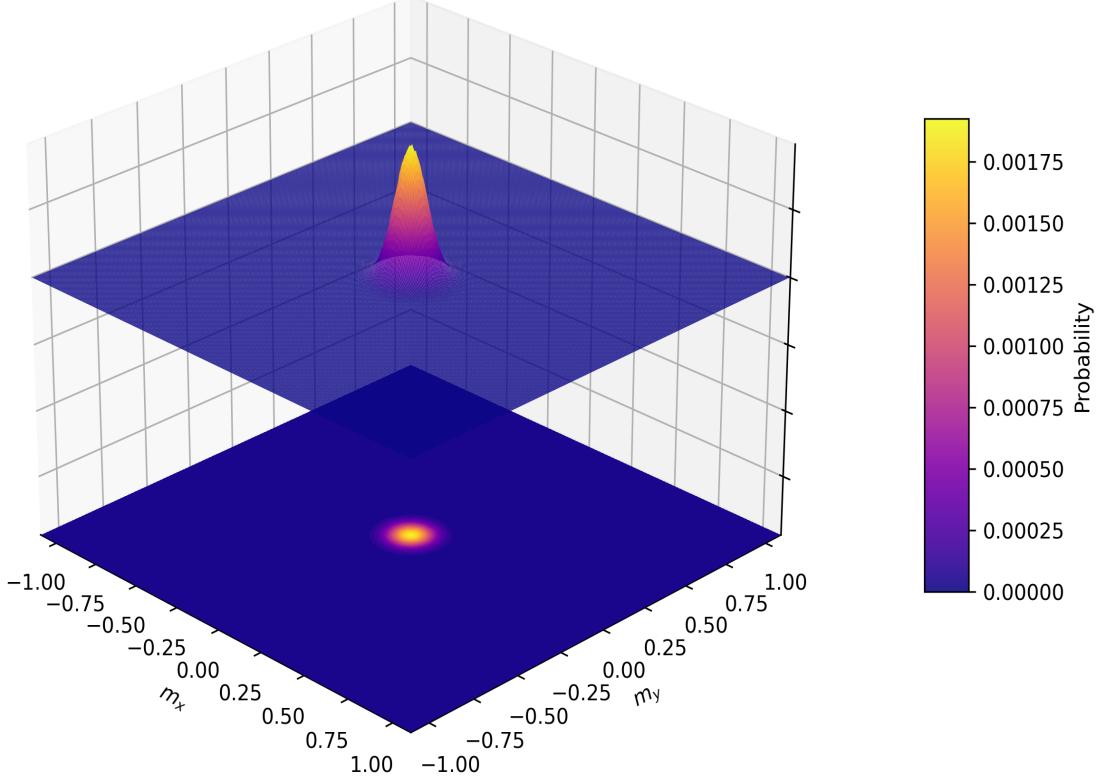


Figure 7 – Probability distribution of the magnetization vector for $q = 8$, $L = 32$, at $T = 1.5 J/k_B$ (disordered phase).

For the disordered phase (7), $\langle m_\phi^N \rangle$ is given by the same integral in Eq. 42, because ϕ is once again an iid variable. However, $\langle m^N \rangle$ is given by Eq. 43 with $m_0 = 0$, so $U_4 = 0$. Since, U_4 suffers a discontinuity from $U_4 = 2/3$ in the QLRO phase to $U_4 = 0$ in the PM phase, U_4 should be able to predict the critical temperature separating these phases. Also, since this doesn't occur for U_ϕ , it cannot predict this critical temperature.

2.2.5 Helicity modulus (or spin stiffness)

In the “results” section, it will be shown that U_4 isn’t good in practice at detecting the transition temperature from the QLRO phase to the PM phase. To surmount this difficulty, another observable has to be introduced: the helicity modulus, or spin stiffness (19). The helicity modulus is defined by Eq. 44:

$$\Upsilon(T) := \frac{1}{N} \frac{\partial^2 F}{\partial \delta_x^2}, \quad (44)$$

where N is the total number of spins in the system, F is the free energy and δ_x is an infinitesimal twist angle along the x direction. Intuitively, the spin stiffness can be thought

of as an analogue to the shear modulus in classical mechanics, but for magnetic systems. It is the tendency of ordered spins to adapt in response to a “twist” of the direction of the order parameter along a given direction.

To calculate Eq. 44 in practice, suppose an infinitesimal (uniform) twist δ_x is given to spins along the x axis as in Fig. 8. Every spin along the x direction is rotated relative to its left and right neighbors, but not to its up and down neighbors.

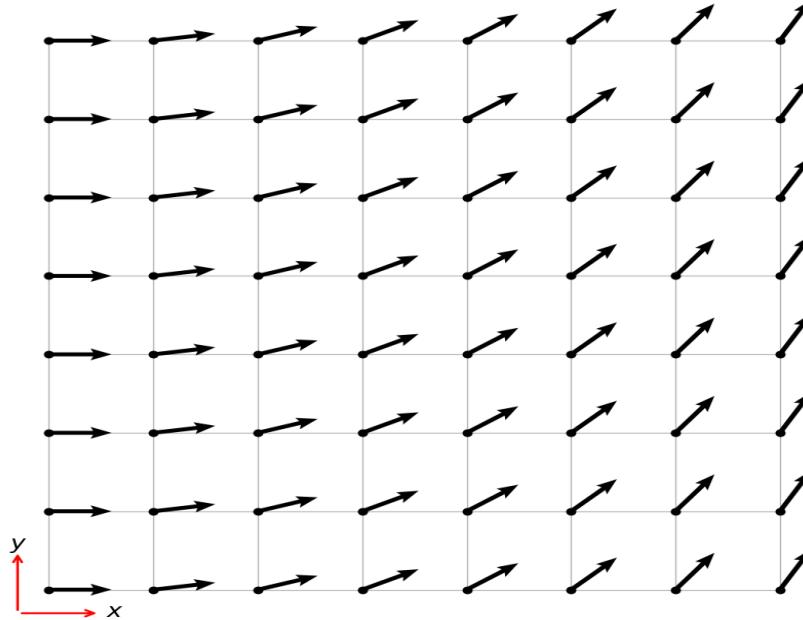


Figure 8 – Infinitesimal twist along every spin in the x direction of a given Hamiltonian.

Suppose a given Hamiltonian has a generic dependence on this infinitesimal twist parameter: $\mathcal{H} = \mathcal{H}(\delta_x)$. The free energy F is given by Eq. 45:

$$F \triangleq \frac{-1}{\beta} \ln \mathcal{Z} = \frac{-1}{\beta} \ln \sum_r e^{-\beta \mathcal{H}(\delta_x)} . \quad (45)$$

Differentiating Eq. 45,

$$\frac{\partial F}{\partial \delta_x} = \frac{-1}{\beta \mathcal{Z}} \sum_r (-\beta) \frac{\partial \mathcal{H}}{\partial \delta_x} e^{-\beta \mathcal{H}} = \left\langle \frac{\partial \mathcal{H}}{\partial \delta_x} \right\rangle . \quad (46)$$

Once more,

$$\frac{\partial^2 F}{\partial \delta_x^2} = \left\langle \frac{\partial^2 \mathcal{H}}{\partial \delta_x^2} \right\rangle + \beta \left(\left\langle \left(\frac{\partial \mathcal{H}}{\partial \delta_x} \right)^2 \right\rangle - \left\langle \left(\frac{\partial \mathcal{H}}{\partial \delta_x} \right) \right\rangle^2 \right) . \quad (47)$$

Eq. 47 can be simplified for the Hamiltonians in Eqs. 1 and 2 because $\left\langle \frac{\partial \mathcal{H}}{\partial \delta_x} \right\rangle = 0$. This is due to the inversion symmetry of these systems: if every spin is reversed $\theta_i \rightarrow -\theta_i$, the energy stays the same, but the derivative of the Hamiltonian changes sign. This happens because when the dot product that appears in the Hamiltonian is differentiated, the $\cos \Delta\theta$ turns into a $\sin \Delta\theta$, which is an odd function. Since the Boltzmann weights of

the configurations of spins that point in the opposite directions are the same, but this derivative changes sign, they perfectly cancel each other in the sum over all microstates. Therefore, Eq. 44 can be written more pragmatically as:

$$\Upsilon(T) = \frac{1}{N} \left(\left\langle \frac{\partial^2 \mathcal{H}}{\partial \delta_x^2} \right\rangle - \beta \left\langle \left(\frac{\partial \mathcal{H}}{\partial \delta_x} \right)^2 \right\rangle \right). \quad (48)$$

For example, in the q -state clock model, applying this infinitesimal twist, the Hamiltonian can be written as:

$$\mathcal{H}(\delta_x) = -J \sum_{i=1}^N (\cos(\theta_i - \theta_{i,\text{up}}) + \cos(\theta_i - (\theta_{i,\text{right}} + \delta_x))) , \quad (49)$$

where $\theta_{i,\text{up}}$ is the angle the spin at the site above ($+y$ direction) the i th one makes with the x axis (analogously for $\theta_{i,\text{right}}$). This way, Eq. 48 becomes:

$$\Upsilon(T) = \frac{1}{N} \left(J \left\langle \sum_{\langle i,j \rangle_x} \cos(\theta_i - \theta_j) \right\rangle - \beta \left\langle \left(J \sum_{\langle i,j \rangle_x} \sin(\theta_i - \theta_j) \right)^2 \right\rangle \right) , \quad (50)$$

where the notation $\sum_{\langle i,j \rangle_x}$ means a sum over all nearest neighbors in the x direction (left and right) for every spin in the lattice (without double counting). A very similar expression can be obtained for the XY $J_1 - J_2$ model using the same logic.

According to the original paper by Nelson and Kosterlitz (17), it can be shown that,

$$\Upsilon(T_{BKT}) = \frac{2T_{BKT}}{\pi} , \quad (51)$$

where T_{BKT} is the transition temperature from the QLRO phase to the PM phase. Because of Eq. 51, the spin stiffness should be able to substitute U_4 in predicting this transition temperature. This can be done by simply plotting $\Upsilon(T)$ and $2T/\pi$ on the same graph: the intersection temperature of these two curves is the transition temperature from the QLRO phase to the PM phase.

The nature of $\Upsilon(T)$ can be understood as follows: in the LRO phase, the system has all the spins locked into place, so any infinitesimal twist will substantially increase the free energy, thus making $\Upsilon(T)$ large. In the QLRO phase (Fig. 9), the spins gain enough energy to be avoid being globally ordered. Now, if you zoom in to a region in the system, for example, the top left corner of Fig. 9, the spins will all appear to be ordered. However, if you move to another region in the system, like the bottom left corner of Fig. 9, the spins will be approximately ordered too, but they will most likely be pointing in another direction. So across the entire system, the spins will be rotating slowly but surely, forming pairs of vortices (16). The defining feature of the QLRO phase is that the spatial spin correlation function decays like a power law, and not exponentially, making it so that the correlation length diverges in every single temperature inside the QLRO phase

($\xi \sim \exp(a/\sqrt{T - T_{\text{BKT}}})$, where a is a constant). Because of this divergence of correlation length, an infinitesimal twist in spins along a direction will be felt by the entire system, since every spin can still “see” each other. Hence, $\Upsilon(T)$ will be nonzero in the QLRO phase too. Finally, in the PM phase, the spins have enough thermal energy to be approximately independent of one another, such that an infinitesimal twist barely alters the free energy, making $\Upsilon(T)$ very small in this phase.

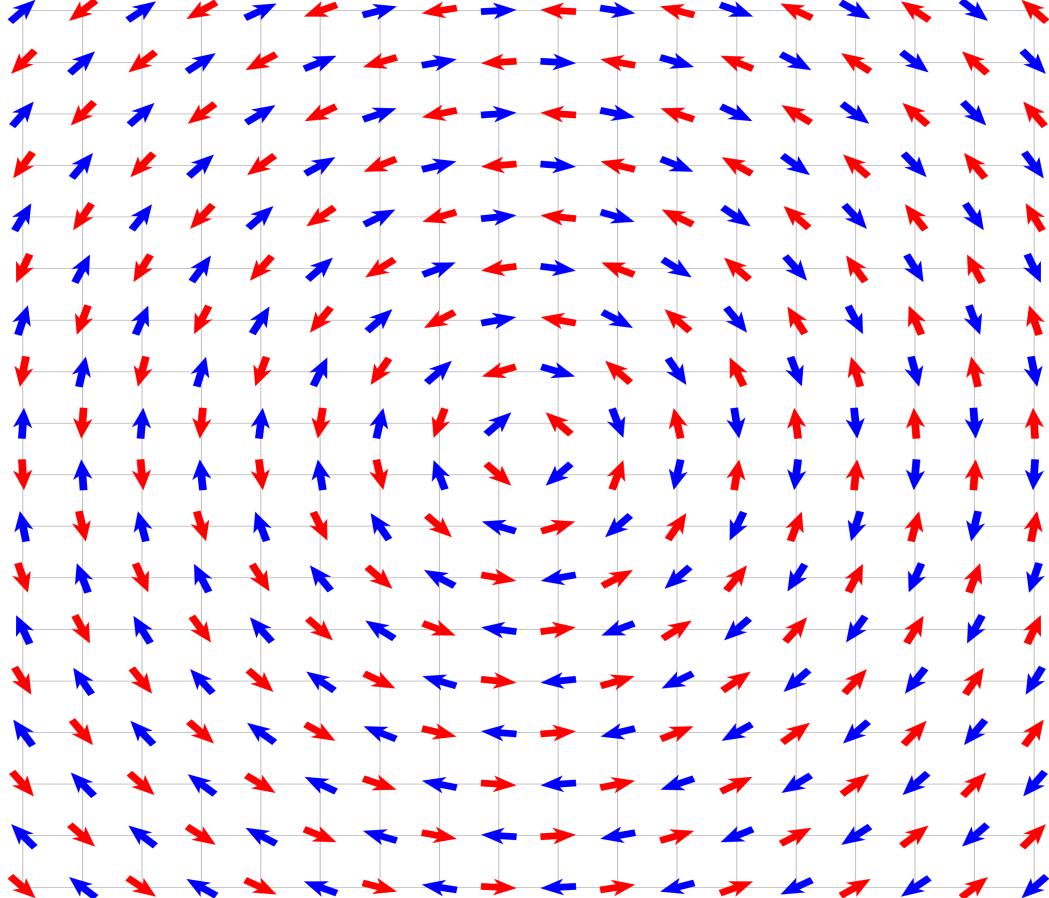


Figure 9 – Schematic representation of spins in the QLRO phase for the antiferromagnetic XY model. Spins in different sublattices have opposite spins.

2.2.6 Nematic order parameter and order-by-disorder

The last of the variables that will be analyzed is the nematic order parameter m_σ , useful for determining the transition temperature from the nematic phase to the PM phase in the XY $J_1 - J_2$ model (see Fig. 29). It is defined by Eq. 52 (13):

$$m_\sigma := \left\langle \frac{1}{4N} \sum_{i=1}^N (\vec{S}_i - \vec{S}_{i+\hat{x}+\hat{y}}) \cdot (\vec{S}_{i+\hat{x}} - \vec{S}_{i+\hat{y}}) \right\rangle , \quad (52)$$

where the spins $\vec{S}_i, \vec{S}_{i+\hat{x}}, \vec{S}_{i+\hat{x}+\hat{y}}, \vec{S}_{i+\hat{y}}$ are the spins in the plaquettes in Fig. 10 starting from the bottom left spin and going counter-clockwise. The term in the sum in Eq. 52 is

calculated for every plaquette and averaged for the whole system in each MC step, then averaged across all MC steps.

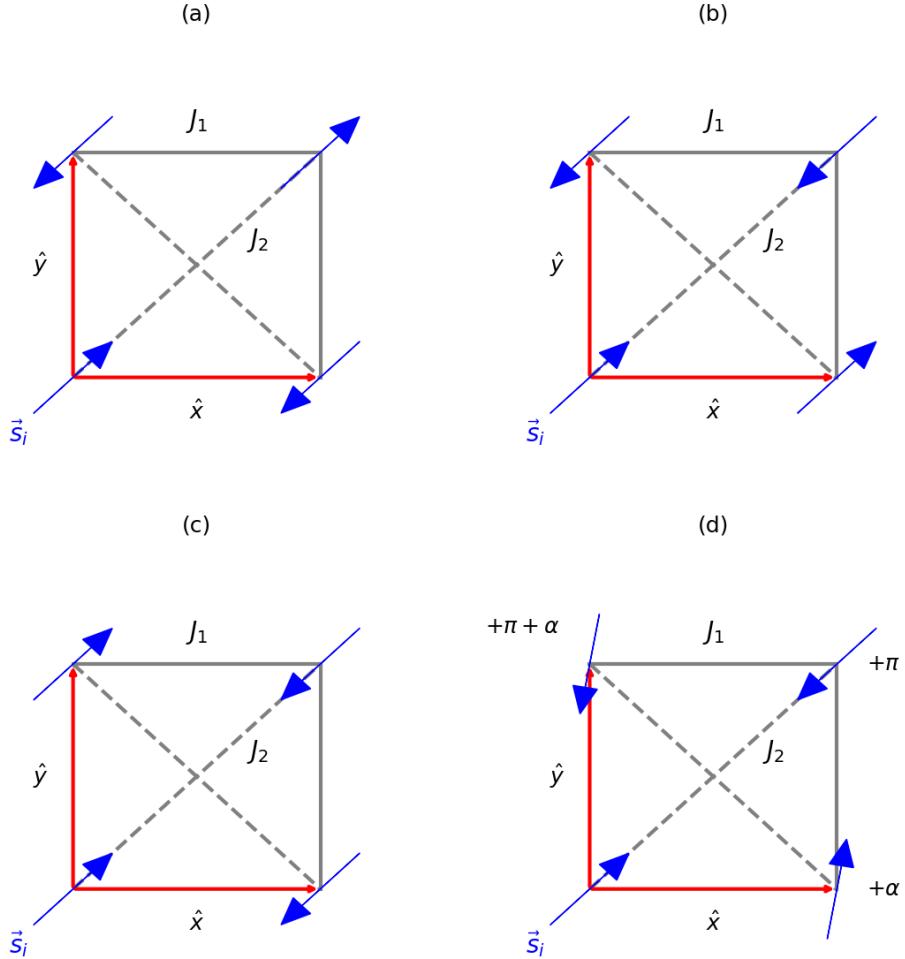


Figure 10 – Plaquettes found in different phases of the 2D XY $J_1 - J_2$ model. Energies in: (a) $E = -2N(J_1 - J_2)$ and (b)-(d) $E = -2NJ_2$ (degenerate).

The nematic order parameter m_σ can be better understood by considering snapshots of configurations in each distinct phase in the 2D XY $J_1 - J_2$ model, given by Figs. 11 - 14.

But first, consider the ground state energy E_{GS} (at $T = 0$) of this system when $J_2/J_1 \ll 1$. This is the case where every plaquette in the system is exactly that of Fig. 10 (a) (nearest neighbor interactions dominate). From Eq. 2, $E_{GS} = -2N(J_1 - J_2)$. Meanwhile, the ground state energy is $E_{GS} = -2NJ_2$ when $J_2/J_1 \gg 1$ (Fig. 10 (b) or (c)). Actually, it will be shown later on in this section that these aren't the only ground state configurations, because the J_1 couplings cancel due to frustration. Clearly there has to be a transition from one ground state to the other by varying J_2/J_1 . It's known in the literature (29) that this occurs when $J_2/J_1 = 0.5$, so in the following sections, the regimes where $J_2/J_1 < 0.5$ and $J_2/J_1 > 0.5$ will be discussed separately.

At low temperatures and for $J_2/J_1 < 0.5$, the system will be in the QLRO 1 phase (Fig. 11 or Fig. 10 (a)), characterized by a quasi-checkerboard pattern. This is because the J_1 coupling dominates the energy contributions, so the system will try to minimize it by keeping nearest neighbor spins antiparallel (remember $J_1, J_2 > 0$). In doing so, the next-nearest-neighbors will have to be parallel, causing magnetic frustration. Increasing the temperature but maintaining the same ratio of J_2/J_1 , the PM phase will be reached (Fig. 12). In both these phases, $m_\sigma = 0$. This can be easily seen by applying the dot product in Eq. 52 to Fig. 10 (a). Thus, m_σ won't be a relevant observable for $J_2/J_1 < 0.5$. To detect this phase transition, using the same logic as in section 2.2.5, we use the helicity modulus.

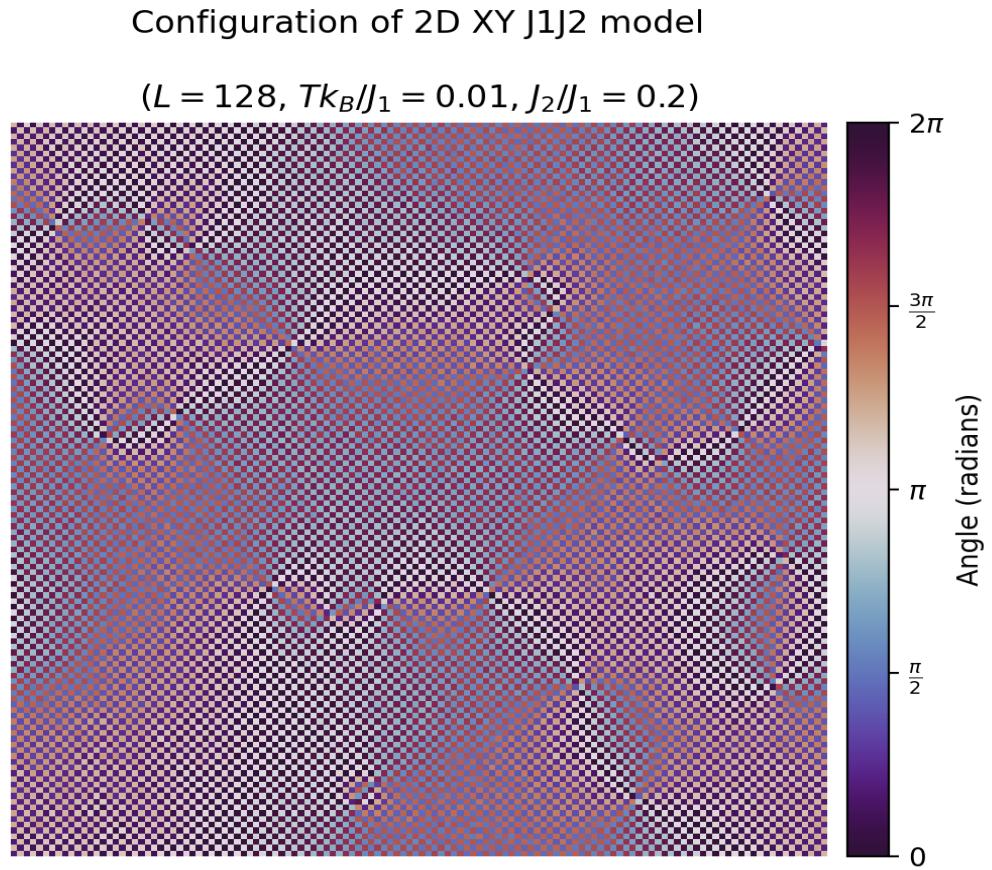


Figure 11 – Snapshot of the spin configuration for the 2D XY $J_1 - J_2$ model in the antiferromagnetic quasi-long-range ordered phase ($T < T_{\text{BKT}}$).

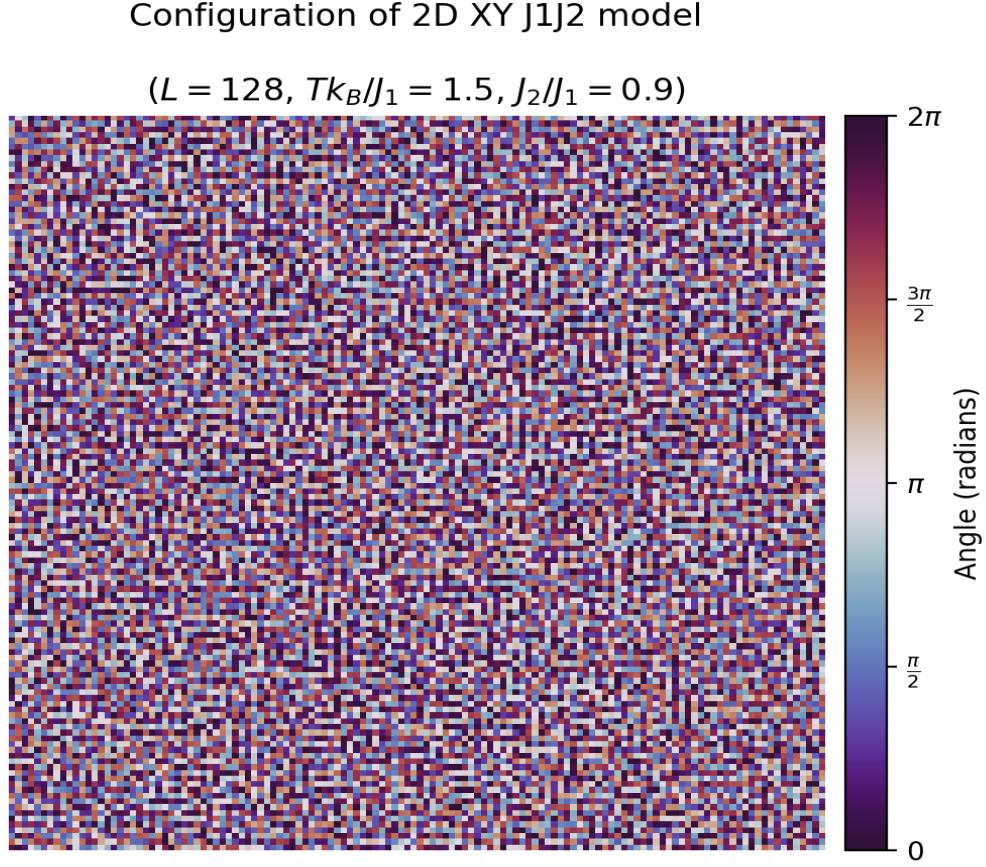


Figure 12 – Snapshot of the spin configuration for the 2D XY $J_1 - J_2$ model in the PM phase ($T > T_{\text{BKT}}$).

As was stated before, the ground state for $J_2/J_1 \gg 1$ is $E_{\text{GS}} = -2NJ_2$. Interestingly, the ground state in this regime has a nontrivial degeneracy. To show this, consider Fig. 10 (d). The bottom right spin is rotated by an arbitrary angle α relative to the bottom left spin \vec{S}_i , the top right spin is still antiparallel to \vec{S}_i (rotated π radians) so as to minimize the energy and the top left spin is rotated by an angle $\pi + \alpha$ relative to \vec{S}_i . The energy in the configuration of this single plaquette is given by:

$$E_{\text{plaquette}} = J_1 \cos(\alpha) + J_1 \cos(\pi + \alpha) + J_2 \cos(\pi) + J_2 \cos(\pi) = -2J_2 . \quad (53)$$

Since there are N such plaquettes in the system (because of periodic boundary conditions), the energy of this configuration is $E_{\text{GS}} = -2NJ_2$, *i.e.*, the ground state for the QLRO 2 phase has a nontrivial degeneracy when $J_2/J_1 \gg 1$.

It should be noted that because of the geometry of the lattice, this same logic can't be applied to the QLRO 1 phase in the $J_2/J_1 \ll 1$ regime. If the top right spin in Fig. 10 (a) is rotated by α radians relative to the bottom left spin, then, in order to minimize the energy (since $T = 0$), the top left and bottom right spins would each have to choose between a rotation of π (satisfying the bond with the bottom left spin) or $\pi + \alpha$ (satisfying the bond with the top right spin). But if one of these is chosen, then one of the nearest

neighbor bonds won't have the smallest energy possible, meaning that this isn't a ground state.

If the temperature is increased from the ground state of the QLRO 2 phase, it can be shown (29, 30) that, by expanding the free energy up to Gaussian fluctuations, the correction in the free energy is proportional to $-\cos^2(\alpha)$, where α is the same parameter as before: a rotation in one of the two J_2 sublattices. Therefore, the free energy is minimized whenever $\alpha = 0, \pi$, which corresponds to the horizontal and vertical stripe plaquettes in Figs. 10 (b) and 10 (c), respectively. In other words, the nontrivial degeneracy is partially lifted due to thermal fluctuations. This entropic selection is known as "order-by-disorder".

At low temperatures and for $J_2/J_1 > 0.5$, the system will be in the QLRO 2 phase (Fig. 13), characterized by a "striped" pattern that could be either horizontal (similar to Fig. 10 (b), but with some thermal fluctuations) or vertical (similar to Fig. 10 (c)). This is because now the J_2 coupling dominates the energy contributions, so the system will try to minimize it by keeping next-nearest-neighbor spins antiparallel, creating frustrated parallel nearest neighbors. This time, $m_\sigma \approx \pm 1$, as can once again be seen from the dot product in Eq. 52, but applied to Figs. 10 (b) ($m_\sigma \approx +1$) and 10 (c) ($m_\sigma \approx -1$). Because of these two possible values of m_σ (which correspond to configurations with the same energy), the system is said to have Z_2 symmetry.

If the temperature is increased again, the system will be in the nematic phase (Fig. 14). The continuous $O(2)$ symmetry will be **almost** broken, in the same sense as before with the QLRO 1 and PM phases, but this time, the nematic order parameter continues being nonzero (Z_2 symmetry remains). In other words, this nematic phase isn't a QLRO phase anymore and the correlation length doesn't diverge. Because of this, the helicity modulus should be able to predict the transition temperature from the QLRO 2 phase to the nematic phase.

Increase the temperature once more and the system will reach the PM phase (Fig. 12). In the transition from the nematic phase ($m_\sigma \neq 0$) to the PM phase ($m_\sigma = 0$), the discrete Z_2 symmetry is broken. Because of this, it is expected from universality (30) that this phase transition is in the same universality class as the 2D Ising model. The order parameter in this case is m_σ , so, using the same logic as before, an appropriate Binder cumulant U_σ can be defined (Eq. 54), along with a nematic order parameter susceptibility (Eq. 55), in analogy to the magnetic susceptibility:

$$U_\sigma := 1 - \frac{\langle m_\sigma^4 \rangle}{3\langle m_\sigma^2 \rangle^2}, \quad (54)$$

$$\chi_\sigma := \beta \left(\langle m_\sigma^2 \rangle - \langle m_\sigma \rangle^2 \right). \quad (55)$$

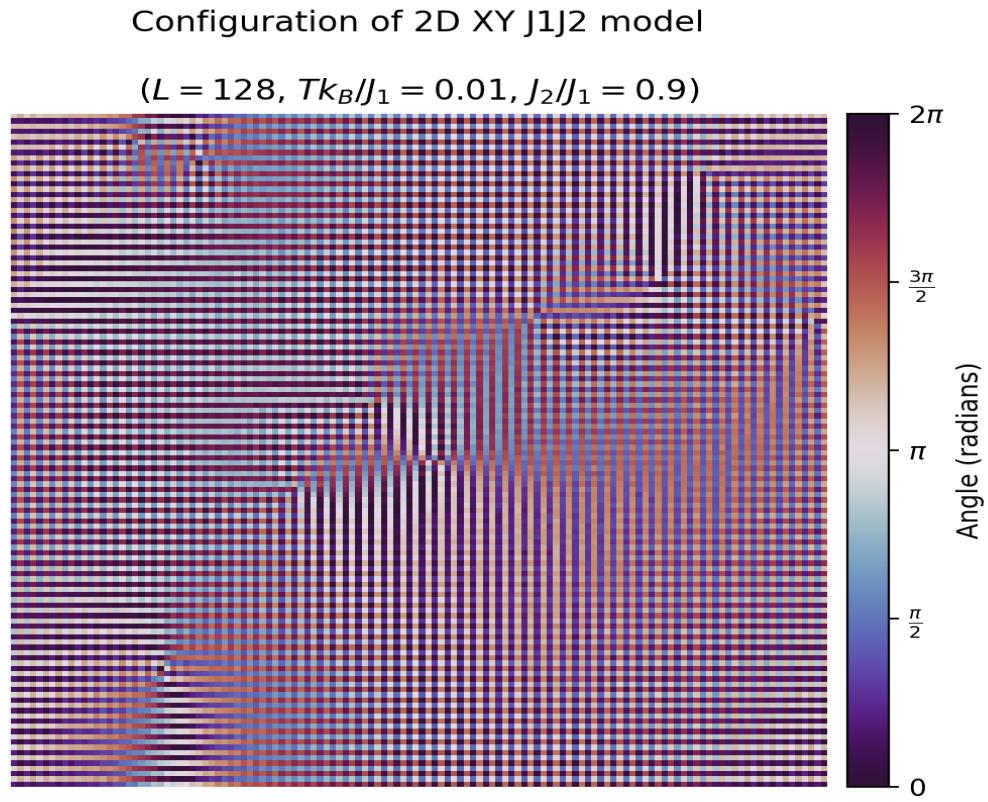


Figure 13 – Snapshot of the spin configuration for the 2D XY $J_1 - J_2$ model in the QLRO 2 phase.

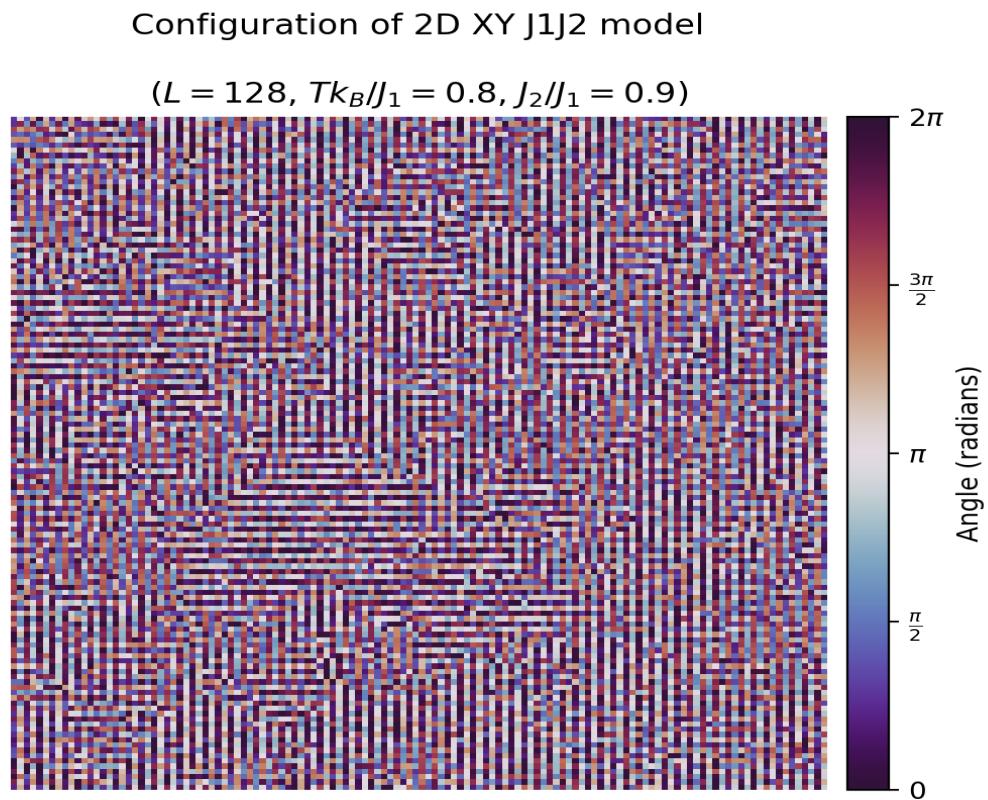


Figure 14 – Snapshot of the spin configuration for the 2D XY $J_1 - J_2$ model in the nematic phase.

2.3 Finite-size scaling

Empirically, at a continuous phase transition, the correlation length ξ of the order parameter diverges as the critical temperature is approached according to a power law characterized by the critical exponent ν : $\xi \sim |t|^{-\nu}$. Let Q be an observable that diverges at the critical temperature according to a power law with critical exponent k ,

$$Q \sim |t|^{-k} . \quad (56)$$

Eq. 56 assumes systems in the thermodynamic limit, i.e., infinitely large systems. However, since Monte Carlo simulations can only deal with finite systems, to precisely estimate critical exponents, beyond just naively simulating very large system sizes as a surrogate for the thermodynamic limit, one has to extrapolate the results for finite systems to infinite ones. To do this, finite-size scaling is usually employed, as it yields accurate results and is simple to implement.

In finite systems, the correlation length cannot truly diverge because it will be limited by the length of the system, say L . So near the phase transition, the relevant length scale will be L and not ξ (which is the correlation length for systems in the thermodynamic limit). Hence, from Eq. 56:

$$Q \sim |t|^{-k} = (\xi^{-1/\nu})^{-k} = \xi^{k/\nu} \sim L^{k/\nu} . \quad (57)$$

Eq. 57 quantifies the behavior of the observable $Q = Q(t, L)$ near the phase transition as a function of the system size. The behavior of this observable in other regions isn't relevant for the calculation of the critical exponents, so it is assumed to be given by a generic non-singular real function ¹³ called the “scaling function” $g_Q(x)$ (or $f_Q(x)$), whose functional form depends on the observable in question.

The finite-size scaling hypothesis states that $g_Q(x)$ (and $f_Q(x)$) will be a function of exactly one variable and that this variable is the ratio of the correlation length and the system size. So the observable Q will be given by Eq. 58 ¹⁴ ($\xi = t^{-\nu}$ was used, and the scaling function was given a different, more convenient, appearance):

$$Q(t, L) = L^{k/\nu} g_Q \left(\frac{\xi}{L} \right) \equiv L^{k/\nu} f_Q \left(t L^{1/\nu} \right) . \quad (58)$$

In practice, one can run simulations for multiple system sizes across a temperature range that contains the critical temperature and register the observable Q at each temperature for all system sizes. Then, on a single graph, one can plot $QL^{-k/\nu}$ as a function of $tL^{1/\nu}$, which is precisely the scaling function, and leave T_c , k and ν as free variables to be

¹³ A non-singular real function is simply a function that is defined, finite and differentiable in all points of its domain.

¹⁴ At the critical temperature, $t = 0$, so the scaling function is just a number and the behavior of Q given by Eq. 57 is recovered.

modified as to minimize the distance between all the plots, thus “collapsing the data”. This data collapse should occur when the correct values of exponents and critical temperature are used, because then Eq. 58 holds true and the data lies along the scaling function¹⁵. To carry out this minimization, it is customary to define the quantity P_b , given by Eq. 59:

$$P_b := \frac{1}{N} \sum_{p=1}^N \sum_{j \neq p} \sum_i (L_j^{-k/\nu} Q_{i,j} - E_p(L_j^{1/\nu} t_{i,j}))^2 , \quad (59)$$

where the indices i , j , and p refer to the i th index of the temperature in the range simulated, the j th set of L used and the p th set of L used. Also, the function E_p is just an interpolation of the data in set p . In other words, given an initial guess of (k, ν, T_c) , a set of observables (Q, t) are chosen from a system simulated with size L_p , these observables are interpolated to give E_p and the distance between this function and every data point in every other data set is calculated (on the graph of $Q(t, L) L^{-k/\nu}$ vs $t L^{1/\nu}$), and this entire process is repeated for each system size and then averaged to yield P_b . This gives the value of the function to be minimized with the arguments (k, ν, T_c) .

It should be noted that in some observables, a power law divergence near the phase transition doesn't occur, so the scaling laws take on a different form. For example, it is known that the specific heat in the 2D Ising model diverges logarithmically, so the proposed scaling law for it will be given by Eq. 60:

$$C_v \sim \log(|t|) = \log(\xi^{-1/\nu}) \sim \frac{-1}{\nu} \log(L) \implies \frac{C_v \nu}{\log(L)} = f_{C_v}(t L^{1/\nu}) . \quad (60)$$

To estimate its critical exponent α , Eq. 60 will be manually changed to:

$$\frac{C_v \nu}{\log^{1-\alpha}(L)} = f_{C_v}(t L^{1/\nu}) . \quad (61)$$

This way, the data collapse in Eq. 59 can be performed (with $L_j^{-k/\nu}$ being replaced by $\nu / \log^{1-\alpha}(L_j)$) and an estimate for α can be made (ideally, $\alpha = 0$).

Another example of a different scaling form is in the helicity modulus. Eq. 51 is only valid for infinite systems. To take into account finite systems, a logarithmic size correction must be employed (19):

$$\Upsilon(T, L) = \frac{2T_{\text{BKT}}}{\pi} \left(1 + \frac{1}{2 \ln L + c} \right) , \quad (62)$$

¹⁵ To be even more accurate, the finite-size scaling hypothesis can be weakened to have the scaling function depend on the system size ($g_Q(x) = g_{Q,L}(x)$), but the sequence $(g_{Q,L}(x))$ as $L \rightarrow \infty$ should be uniformly convergent, as to recover the behavior in the thermodynamic limit. If this is assumed, then Eq. 58 becomes:

$$Q(t, L) L^{-k/\nu} = f_Q(t L^{1/\nu})(1 + a L^{-w}) \equiv f_{Q,L}(x)$$

where the 1 is to recover Eq. 58 as $L \rightarrow \infty$, and $a L^{-w}$ is just a generic power law dependence (with two more free variables to adjust: a and w), because if $t = 0$, then the behavior of Q would continue being a power law, as is required.

where c is a system dependent free parameter. Now, a scaling hypothesis can be made by substituting $L^{k/\nu}$ in Eq. 58 with this logarithmic correction in Eq. 62. Using the fact that in the QLRO phase, the correlation length diverges according to $\xi \sim \exp(a/\sqrt{T - T_{\text{BKT}}})$:

$$\Upsilon(T, L) = \left(1 + \frac{1}{2 \ln L + c}\right) g_\Upsilon \left(\exp\left(a/\sqrt{T - T_{\text{BKT}}}\right)/L\right). \quad (63)$$

Alternatively, taking the logarithm of the argument of g_Υ ¹⁶:

$$\Upsilon(T, L) \left(1 + \frac{1}{2 \ln L + c}\right)^{-1} = f_\Upsilon \left(\ln L - a/\sqrt{T - T_{\text{BKT}}}\right). \quad (64)$$

Once again, Eq. 59 should be used to estimate (a, c, T_{BKT}) by substituting $L_j^{-k/\nu}$ with $\left(1 + \frac{1}{2 \ln L_j + c}\right)^{-1}$ and $L_j^{1/\nu} t_{i,j}$ with $\ln L_j - a/\sqrt{T_{i,j} - T_{\text{BKT}}}$.

Finally, a scaling relationship for Binder cumulants can also be obtained. Let A be a constant and $Q_2(t, L)$ and $Q_4(t, L)$ the second and fourth moments of an observable that obeys the scaling relation given by Eq. 58¹⁷. The general form of a Binder cumulant could then be given by:

$$U(t, L) = 1 - \frac{1}{A} \frac{Q_4(t, L)}{Q_2(t, L)} \sim 1 - \frac{1}{A} \frac{\left(L^{k/\nu} f_{Q_4}(t L^{1/\nu})\right)^4}{\left[\left(L^{k/\nu} f_{Q_2}(t L^{1/\nu})\right)^2\right]^2} \equiv g_U(t L^{1/\nu}), \quad (65)$$

where the last equality was just a convenient choice of a scaling function. Since the $L^{k/\nu}$ cancel each other, the Binder cumulant is size independent up to subleading finite-size corrections.

3 Practical considerations

In this work, the Fortran programming language was used for the simulations, along with the “rkiss05” and “ranlux++” (23, 24) pseudorandom number generators. A computer cluster was used with the average node having an Intel core i7 – 3770K (x86_64, 4 cores / 8 threads, 1.6–3.9 GHz, 8 MiB L3 cache). The compiler used was gfortran 9.4.0 with compilation options -O3 -march=native .

To efficiently implement periodic boundary conditions, the spins were stored in a look up table. To this end, two data structures are essential: a map of neighbors and a vector of the states of each spin. The map of neighbors is just a matrix where column j has x lines which correspond to the addresses of the x neighbors site j has, imposing periodic boundary conditions¹⁸. This matrix is calculated once for each system. Then the vector of states will have an entry representing the address of each spin, and the value at that location representing the state of said spin.

¹⁶ Because if $g(x)$ is a function of x , then $g(\ln x)$ is still a function of x .

¹⁷ These could be moments of the magnetization, angular magnetization or nematic order parameter.

¹⁸ It's also convenient to, instead of using 2 numbers to represent the address of a spin, let each site have an address given by just 1 integer. Then, a site in position (x, y) will have a unique address $A = L(x - 1) + y$, where $x, y \in [1, L]$

Another important optimization to make is to precalculate trigonometric and exponential factors that appear in each algorithm once at the start of a simulation at a given temperature (25). This is only possible in the q -state clock model because the minimum difference in angles between any two spins is $2\pi/q$, so vectors containing precalculated values of $\cos(n2\pi/q)$ and $\sin(n2\pi/q)$ are established (for $n = 0, 1, \dots, (q - 1)$), then an array of $\exp(-\beta(\cos(n_12\pi/q) - \cos(n_22\pi/q)))$ is made too (for $n_1, n_2 = 0, 1, \dots, (q - 1)$), and the simulation for that temperature is performed. It should be noted that this isn't possible when simulating the $XY J_1 - J_2$ model, because the minimum angle difference between any two spins can be as small as the precision of the machine used.

Because of the efficiency of these look up tables, instead of simulating the original $XY J_1 - J_2$ model directly, the q -state $J_1 - J_2$ model with large q will be simulated. This is further justified in the “results” section, but essentially, the error in the results that is accrued by making this approximation is smaller than the finite-size effects, so, since the continuous model is considerably slower to simulate, only small systems would be accessible, as opposed to the discrete version, which has permitted simulations of considerably large systems and has yielded results that agree with the literature on the continuous model.

If the Wolff or Swendsen-Wang algorithm is used, it's instrumental to also have a mask for the sites, to be able to index and efficiently find out if a given site is in the cluster or not. This can be done with a simple boolean vector, where the indices represent the addresses of the spins and the values therein are either true or false, signaling if that spin is either in the cluster or not. This avoids having to iterate through every spin in linear time complexity to determine if it's already in the cluster.

If the heat bath algorithm is used for a discrete spin model, after calculating \mathcal{Z}_i^{-1} in Eq. 22, the next step is sampling the cumulative distribution function (CDF) given by $p_n = \exp(-\beta E_n) \mathcal{Z}_i^{-1}$. There are many ways to do this, but one efficient way is to put $(p_0, p_0 + p_1, \dots, 1.0)$ in a vector with q total entries, generate a random number r , then use a simple binary search procedure to find the first index i in this vector such that r is less than the value at index i ¹⁹.

Finally, for the minimization procedure used for Eq. 59, the Nelder-Mead, or “downhill simplex”, method (22, 28) was used, along with the Lagrange polynomial interpolation scheme (22). These methods were used largely because of their simplicity to implement. The vast majority of time spent by the computer was in simulating systems. Once the results were obtained, however, analyzing the data is a comparatively easier task on the computer, albeit time consuming for the programmer.

The Nelder-Mead algorithm aims to solve the following problem: given a multi-

¹⁹ The time complexity for a naive sampling of a CDF is $\mathcal{O}(q)$, but with a binary search, this becomes $\mathcal{O}(\log q)$, which is substantially better for large q , since this procedure will normally be performed many times.

dimensional function $f : A \subset \mathbb{R}^N \rightarrow \mathbb{R}$, what is the minimum value of $f(x_1, x_2, \dots, x_N)$ in A ? First, the algorithm starts off with an initial guess $P_0 \in A$. Then, N other points will be saved by adding $\Delta \vec{e}_i$ to P_0 , where Δ is a constant and \vec{e}_i is one of the N linearly independent unit vectors. These $N + 1$ points form a nondegenerate “simplex”, *i.e.*, a geometrical figure of $N + 1$ vertices in N dimensional space, such that no three vertices are collinear (in 2D, a nondegenerate simplex is a triangle, in 3D it is a tetrahedron etc.). Now $f(P_0), f(P_0 + \Delta \vec{e}_1), \dots, f(P_0 + \Delta \vec{e}_N)$ will be calculated and sorted. The centroid is then calculated for this simplex (this is just the mean of all $N + 1$ points) and the point P_l with largest image gets reflected about this point towards the point P_s with the smallest image. If the resulting point P_r has an image smaller than P_l , then it replaces P_l as a new vertex of the simplex, else, P_r is scaled by a certain factor such that it becomes an intermediate point. If P_r still has an image larger than P_l , then another scaling factor can be tried, or this whole process can be repeated on the other points with largest images until a satisfactory P_r is found²⁰. When this happens, if the difference in images from P_r and the point it replaced is fractionally smaller than some predetermined tolerance, then the algorithm terminates, otherwise, it just repeats this whole process on the new nondegenerate simplex formed. If the algorithm terminates, it is restarted a couple of times at the point it claims is a minimum to ensure convergence.

Given a set of N points $\{(x_i, y_i)\}$, where $x_i \neq x_j$ for all i and j , the Lagrange polynomial given by Eq. 66 is the unique polynomial of at most order $N - 1$ that exactly interpolates these points:

$$P(x) := \sum_{i=1}^N \prod_{j \neq i} \frac{(x - x_j)}{(x_i - x_j)} y_i . \quad (66)$$

It can be seen that the Lagrange polynomial actually interpolates these points by calculating $P(x_k)$ in Eq. 66 for $k = 1, 2, \dots, N$: $P(x_k) = y_k$. Every term in the sum will be zero except the one with y_k , and the product will have all the numerators and denominators cancel exactly. Furthermore, this polynomial is unique because, suppose another polynomial $Q(x)$ of degree $\leq N - 1$ also interpolates these N points. Then the difference $P(x) - Q(x)$ will have N zeros at each interpolated point. The only polynomial of degree $\leq N - 1$ with more than $N - 1$ zeros is the constant zero polynomial. Therefore $P(x) - Q(x) = 0 \implies P(x) = Q(x)$ ■.

²⁰ If no satisfactory P_r is found, the whole simplex should be contracted about P_s and the algorithm should continue.

4 RESULTS

4.1 q-state clock model

4.1.1 $q = 2$ (Ising)

The energy per site, magnetization per site squared²¹, specific heat and magnetic susceptibility for the Ising model with lattice sizes $L = 32, 64, 128, 256, 512$, simulated with the Wolff algorithm are shown in Fig. 15. A total of 10^6 MC steps were performed for equilibration and measurements. The error bars for all of the following plots were obtained by performing the same task over multiple computers (~ 10) and calculating the standard deviation of the set of results.

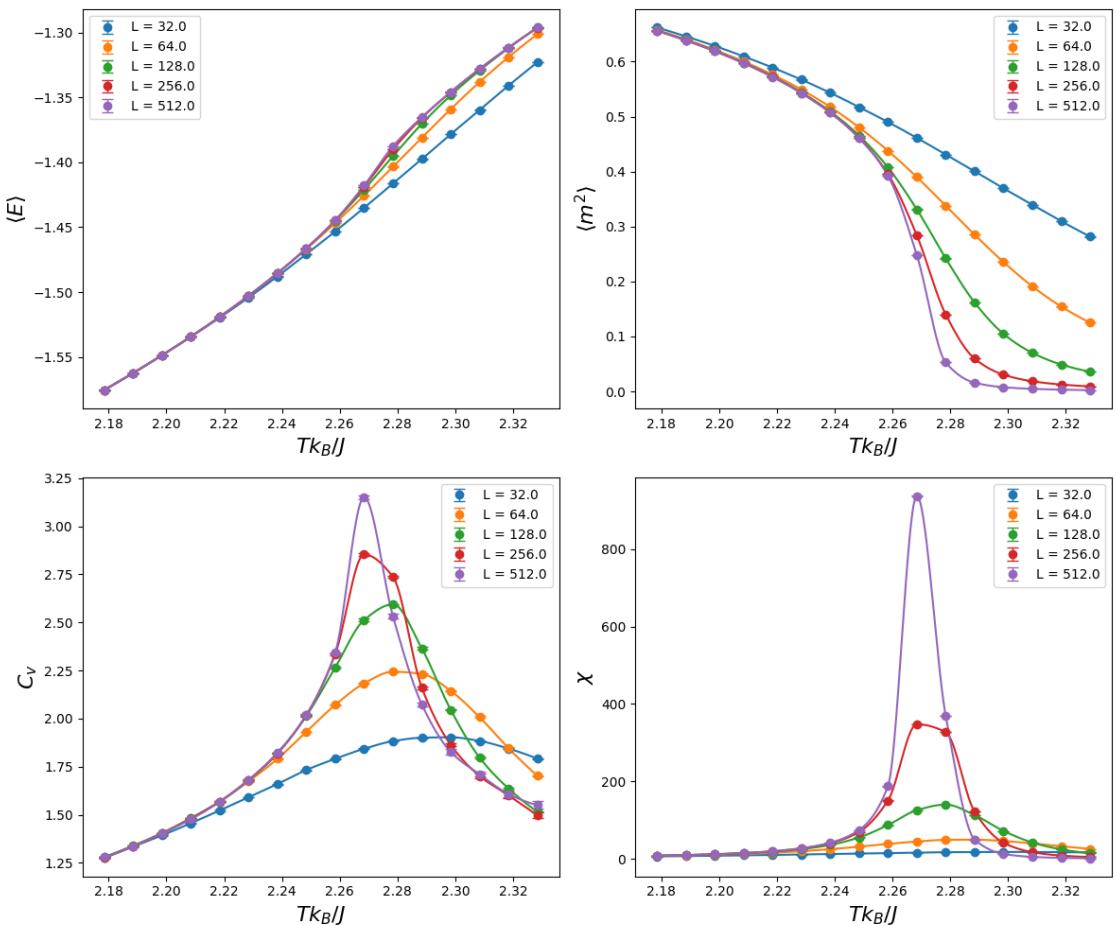


Figure 15 – Energy per site, magnetization per site squared, specific heat and magnetic susceptibility for the Ising model simulated with the Wolff algorithm.

The correlation lengths and correlation times simulated with the Wolff and Metropolis algorithms are shown in Fig. 16. It should be noted that the correlation time for the Wolff algorithm is defined as the correlation time calculated (as in Eq. 39) multiplied by

²¹ The magnetization squared was used because, below T_c , the Wolff algorithm essentially ignores the spontaneous breaking of symmetry and flips most spins after each MC step, such that the average magnetization always comes out to zero. Incidentally, in calculating m^2 , the Binder cumulant analysis becomes practically unchanged: the only difference being that now it should be defined as $U_4 = 1 - \langle m^8 \rangle / (3\langle m^4 \rangle^2)$, so now, for $T \rightarrow \infty$, $U_4 \rightarrow -2.89$

the average cluster size over the course of the simulation (because an MC step for the Wolff algorithm was defined as N total cluster flips, while the Metropolis algorithm has an MC step defined as N possible site flips). From the higher peaks of the correlation times in the Metropolis algorithm, it can be observed that the Metropolis algorithm has a harder time decorrelating successive measurements than the Wolff algorithm in the 2D Ising model (near T_c). This is further corroborated in Fig. 17, which performs a data collapse on the correlation times for the Wolff and Metropolis algorithms to estimate the dynamical critical exponents. This is done by identifying $Q \equiv \tau$ and $k \equiv z\nu$ in Eq. 58.

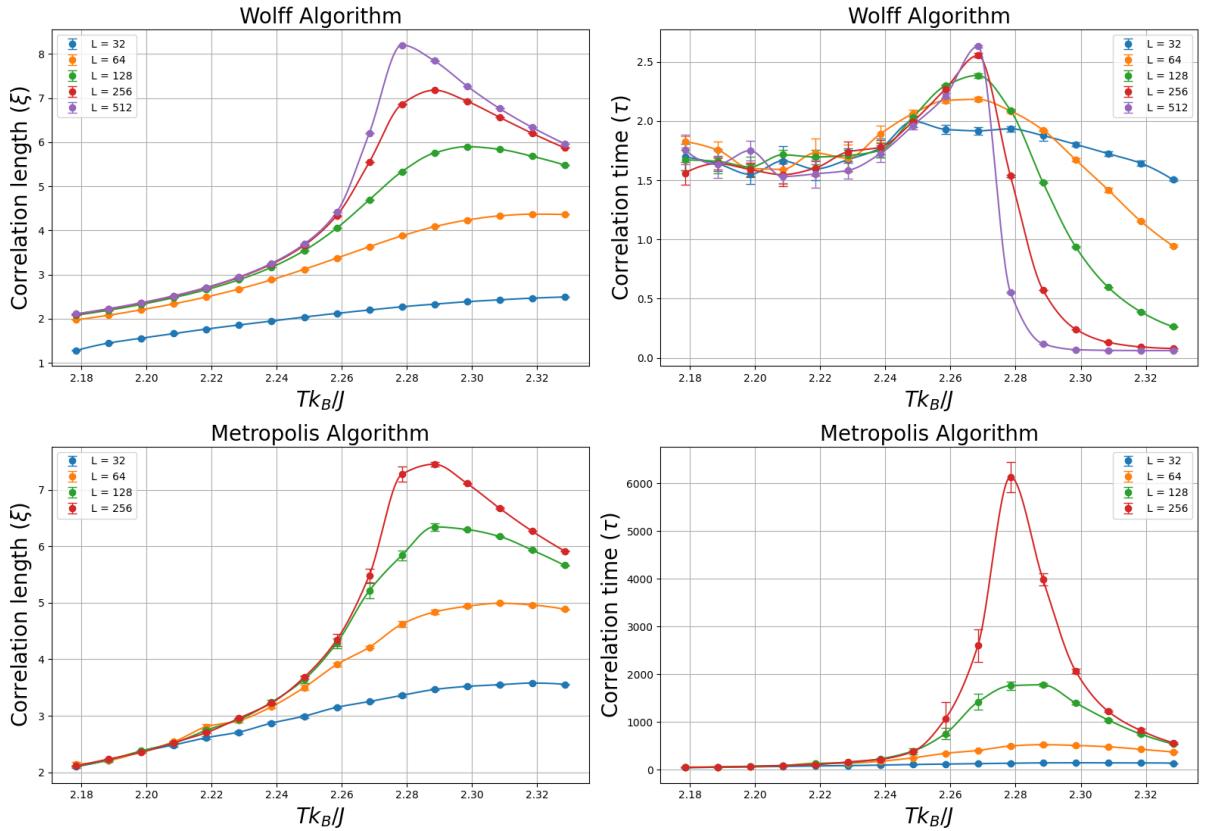


Figure 16 – Correlation lengths and correlation times calculated with the Metropolis and Wolff algorithms.

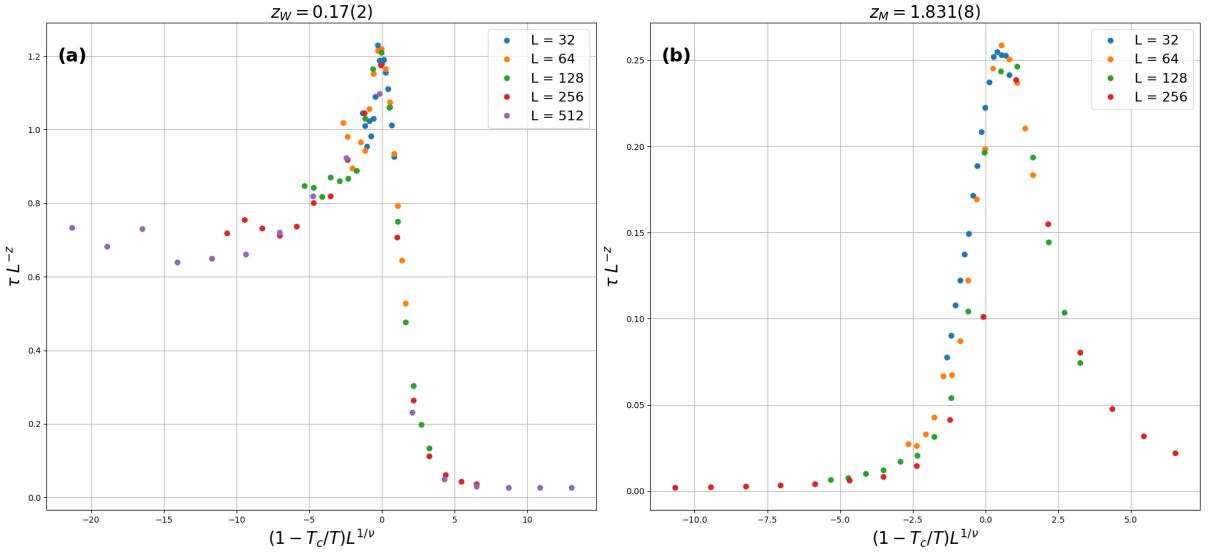


Figure 17 – Data collapses of correlation times: τL^{-z} versus $(1 - T_c/T)L^{1/\nu}$ ($T_c = 2.269185$ and $\nu = 1.0099$) performed for the (a) Wolff algorithm and the (b) Metropolis algorithm.

Data collapses for the specific heat (Eq. 61), magnetization squared ($Q \equiv m^2$, $k \equiv -2\beta$) and magnetic susceptibility ($Q \equiv \chi$, $k \equiv \gamma$) were also performed and are shown in Fig. 18. Furthermore, a plot of an estimate of the anomalous critical exponent η is shown as well: the temperatures corresponding to the largest correlation lengths in Fig. 16 (for the Wolff algorithm) were found and their respective correlation functions were obtained. Then, the natural logarithm of these correlation functions versus the natural logarithm of the distances were plotted and their slopes obtained for different system sizes. This slope is an approximation for η . It should get more accurate for larger systems and more fine discretizations of temperatures.

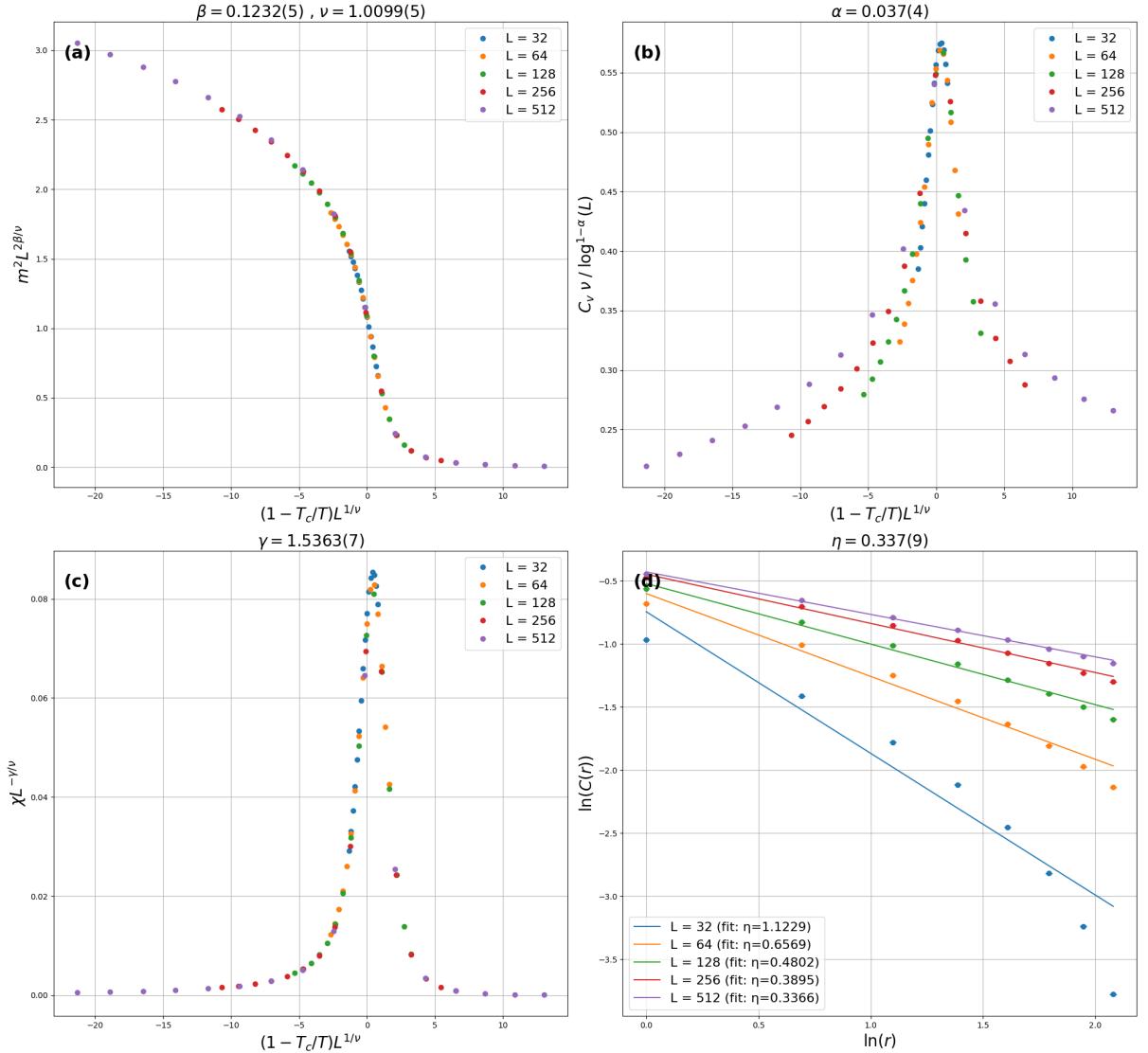


Figure 18 – Data collapses performed on data from Wolff algorithm. All plots have $T_c = 2.269185$ and $\nu = 1.0099$.
(a) Magnetization per site squared: $m^2 L^{2\beta/\nu}$ versus $(1 - T_c/T)L^{1/\nu}$. (b) Specific heat: $C_v \nu / \log^{1-\alpha}(L)$ versus $(1 - T_c/T)L^{1/\nu}$. (c) Magnetic susceptibility: $\chi L^{-\gamma/\nu}$ versus $(1 - T_c/T)L^{1/\nu}$. (d) Anomalous critical exponent: \log of spatial correlation function versus \log of distance at the temperatures with largest calculated correlation lengths.

Lastly, the Binder cumulant U_4 of the magnetization squared is shown in figure 19, alongside a collapse of data by plotting U_4 versus $(1 - T_c/T)L^{1/\nu}$ (Eq. 65) and a plot of the intersection temperatures for all the system sizes used versus the reciprocal of the minimum lattice size used in this intersection. This is done to further minimize finite-size effects of the lattices on the estimation of the critical exponents. As the lattice sizes get larger (Y axis is approached), the true critical temperature should be arrived at. A summary of all the critical exponents and critical temperatures estimated for the 2D Ising model with mean-field theory, numerical calculations and exact results (26) are shown in table 1. The values of the critical exponents obtained numerically are relatively close to the exact

results, but the errors are very small. This is due to the way the minimization procedure was carried out: after the Nelder-Mead algorithm converges on a minimum, it's restarted at that same point to see if it will indeed converge to it again. The problem is that the minima are very large, so most of the nodes typically converge to them quickly with little variation, producing errors that are much smaller than expected. If larger lattice sizes were used, along with a finer discretization of temperatures near T_c , then the estimates for the critical exponents would be even closer to the exact results.

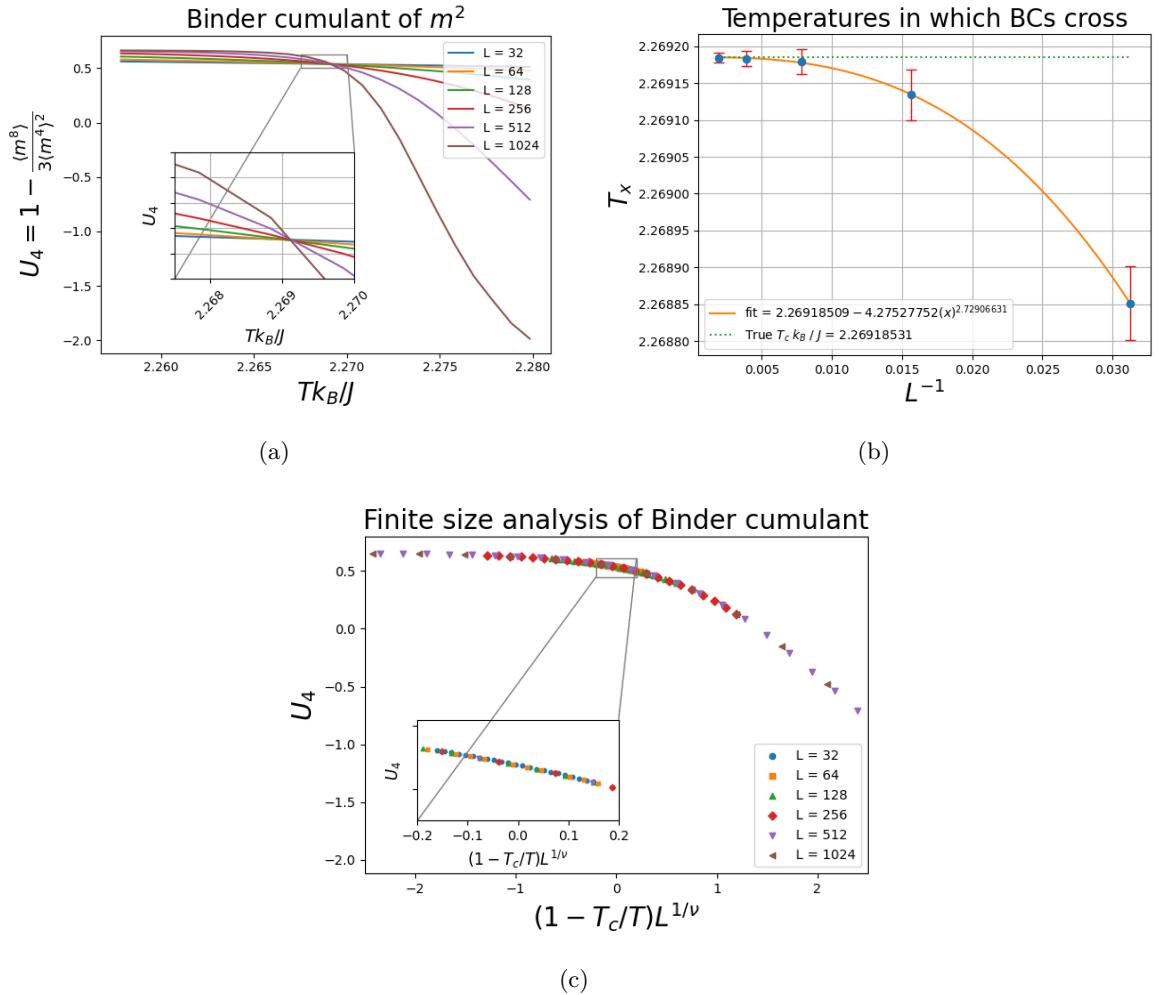


Figure 19 – (a) Binder cumulant $U_4 = 1 - \langle m^8 \rangle / 3\langle m^4 \rangle^2$ for various system sizes on the Ising model. (b) Temperatures in which Binder cumulants intersect for adjacent system sizes (T_x) as a function of the reciprocal of the smaller respective system size (L^{-1}). The true critical temperature is the dotted green line. (c) Binder cumulant for each system size plotted against $tL^{1/\nu}$, where $T_c = 2.269185$ and $\nu = 1.0099$ (values obtained numerically from finite size scaling).

	Exact	Mean-field	Numerical
$T_c k_B / J$	≈ 2.269185314	4	2.269185(1)
α	0 (log divergence)	0 (discontinuity)	0.037(4)
β	0.125	0.5	0.1232(5)
γ	1.75	1	1.5363(7)
ν	1	0.5	1.0099(5)
η	0.25	0	0.337(9)
$z_{\text{Metropolis}}$	—	—	1.831(8)
z_{Wolff}	—	—	0.17(2)

Table 1 – Comparison of the exact, mean-field and numerical values of the critical temperature and critical exponents for the 2D Ising model. Critical exponents shown are for the specific heat (α), magnetization (β), magnetic susceptibility (γ), correlation length (ν), anomalous critical exponent (η) and the dynamical critical exponents for the metropolis ($z_{\text{Metropolis}}$) and Wolff algorithm (z_{Wolff}), which are only defined for the numerical results.

4.1.2 $q \geq 3$

In Fig. 20, plots of the energy per site, magnetization per site, specific heat and magnetic susceptibility are shown for $q = 3, 4, 5, 6$ ($L = 16$). These results, along with the others for the q -state clock model, were obtained with the Wolff algorithm, using 10^6 MC steps for equilibration and measurements. For $q \geq 5$, the QLRO phase appears between the LRO and PM phases. This is seen by the two distinct peaks in the specific heat and magnetic susceptibility.

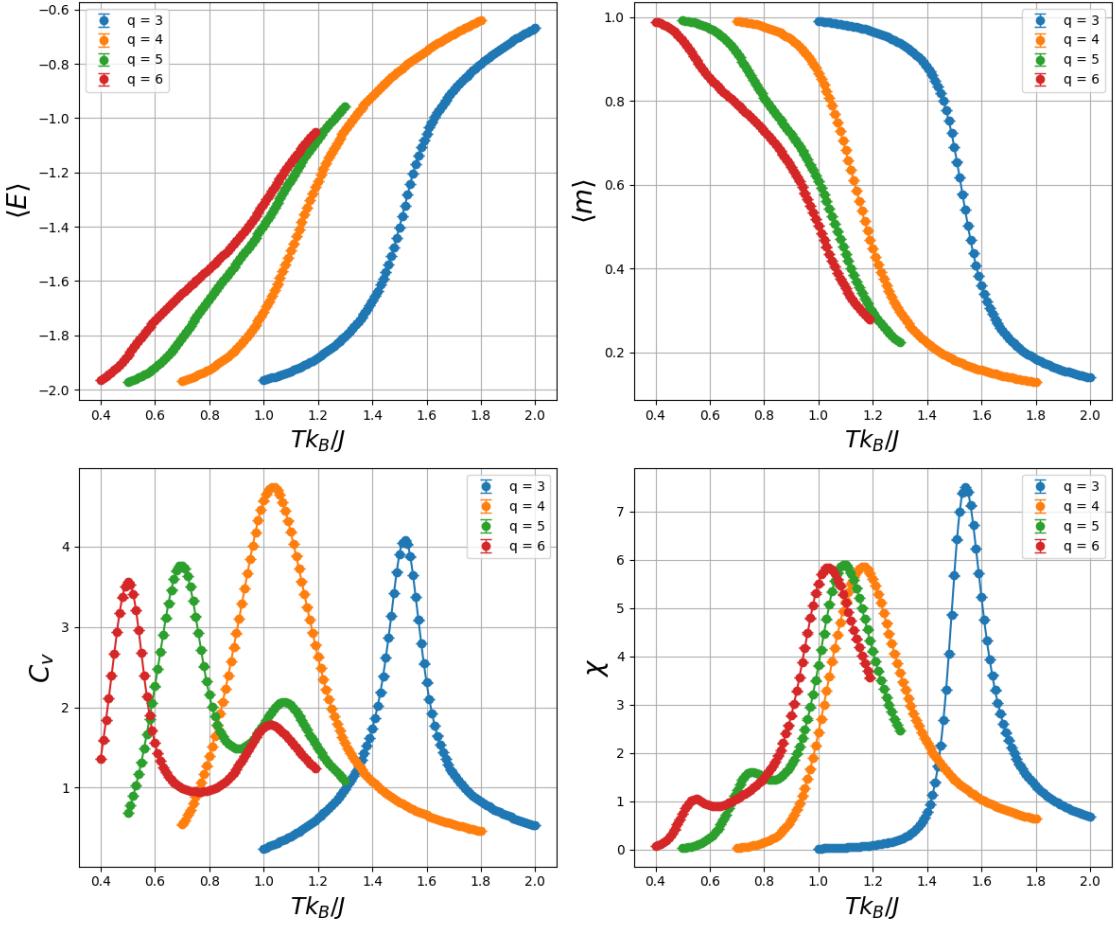
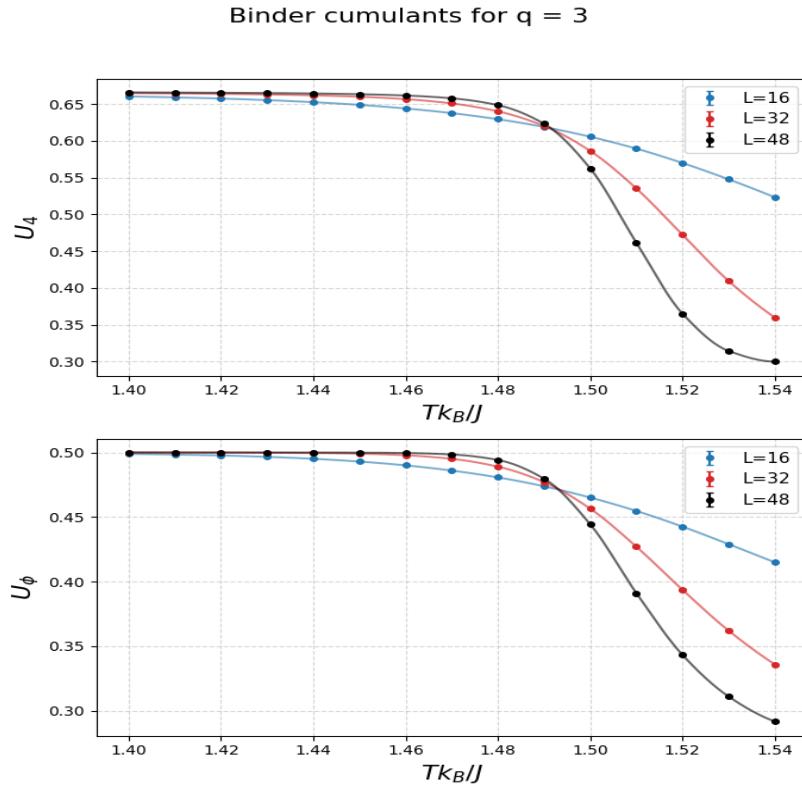
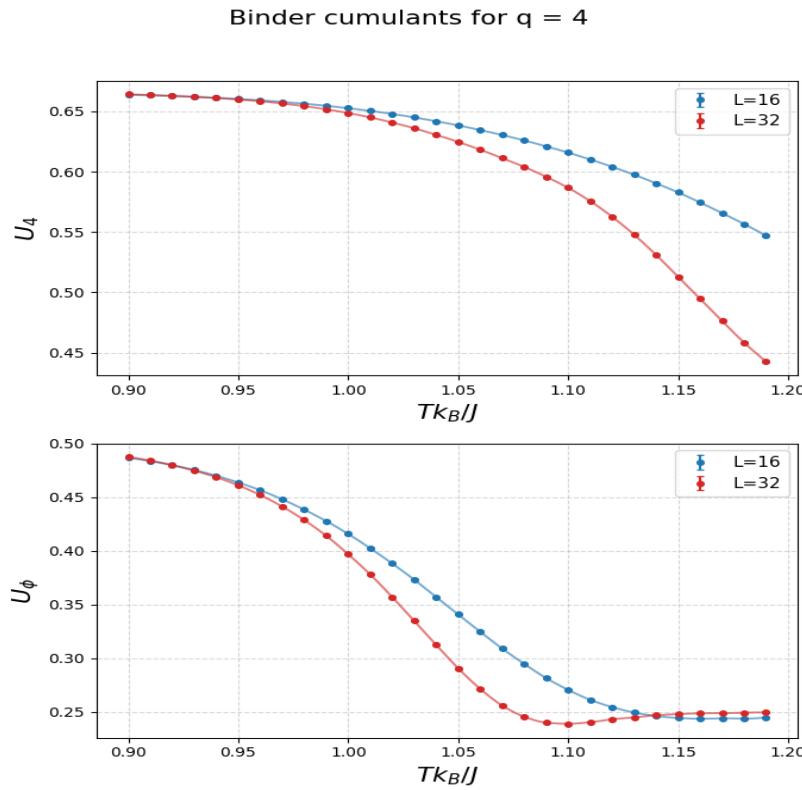


Figure 20 – Energy per site, magnetization per site, specific heat and magnetic susceptibility for $q = 3, 4, 5, 6$ ($L = 16$ for all plots). For $q \geq 5$, two distinct peaks form in the specific heat and magnetic susceptibility.

In Figs. 21 - 24, plots of U_4 , U_ϕ and Υ are shown for $q = 3, 4, 5, 100$ ²². The estimates obtained with U_ϕ (T_c from LRO to QLRO phases) are in agreement with results from the literature (27), however, results with U_4 (T_c from QLRO to PM phases) aren't good for $q \geq 4$ because an intersection cannot be determined. This is most likely because of the small lattice sizes used, since the authors in reference 27 were able to determine an intersection in U_4 by simulating systems up to $L = 256$. Larger lattice sizes were attempted to be simulated, but time constraints limited this analysis (the larger the system used, the more time had to be spent equilibrating it and decorrelating successive microstates generated). So, instead of using the results for U_4 , the critical temperatures from the QLRO to the PM phases were obtained by performing the data collapse in Eq. 64 on the spin stiffness data. This is shown for $q = 100$ in Fig. 25 explicitly²³. This data collapse is what allows for accurate assessments of T_{BKT} despite having the largest lattice size simulated be $L = 48$.

²² These observables for $q = 6, 8, 12$ are very similar in appearance to the $q = 5$ case, so they have been omitted.

²³ The other values of q have very similar collapses so they have also been omitted.

Figure 21 – U_4 and U_ϕ for $q = 3$.Figure 22 – U_4 and U_ϕ for $q = 4$. No intersection can be identified for U_4 , so T_c from U_ϕ has been used instead.

Binder cumulants and spin stiffness for $q = 5$

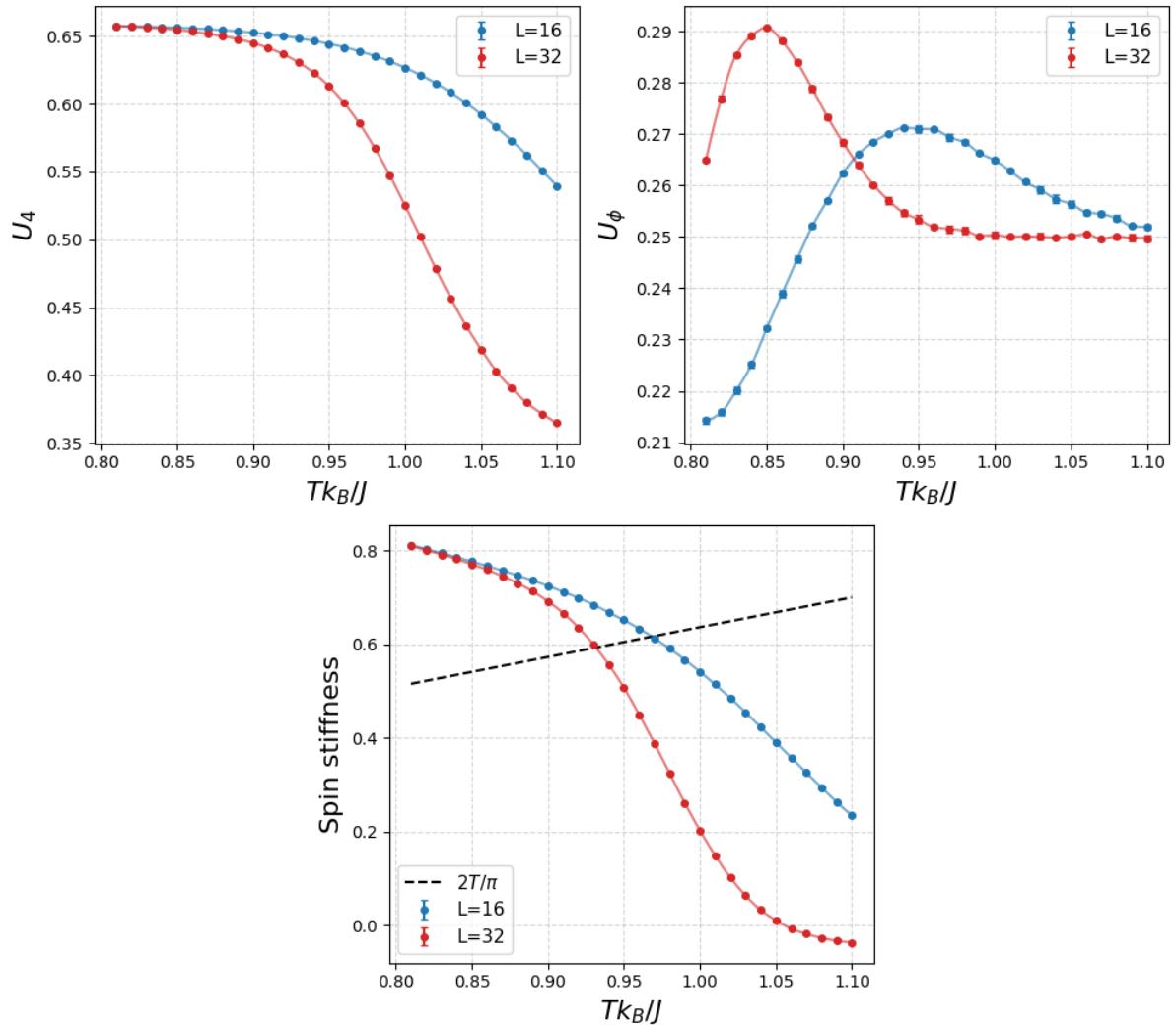


Figure 23 – U_4 , U_ϕ and Υ for $q = 5$. No intersection can be identified for U_4 , so T_c from the spin stiffness has been used instead.

Binder cumulants and spin stiffness for $q = 100$

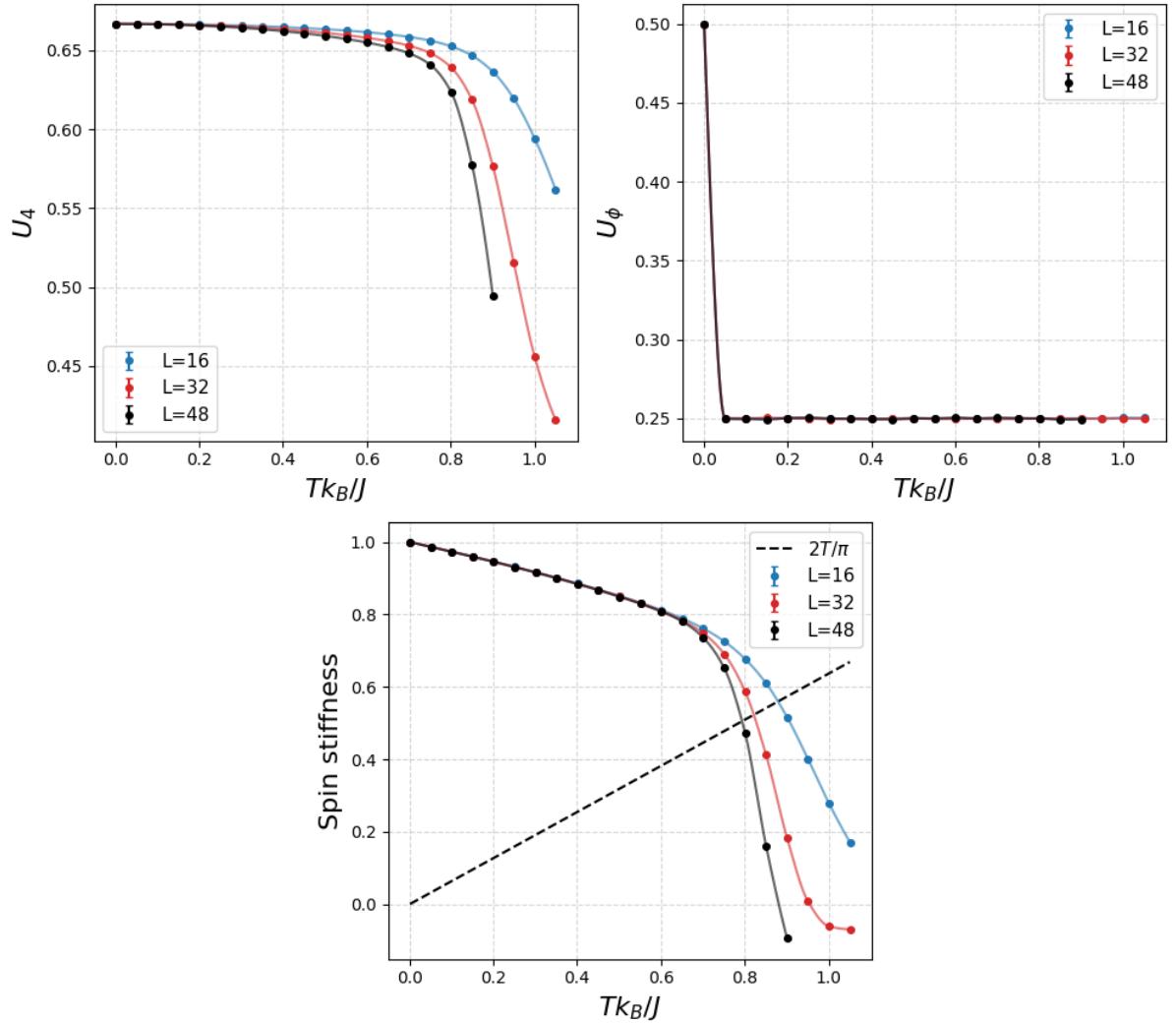


Figure 24 – U_4 , U_ϕ and Υ for $q = 100$. No intersection can be identified for U_4 , so T_c from the spin stiffness has been used instead. Also, no intersection can be identified for U_ϕ either.

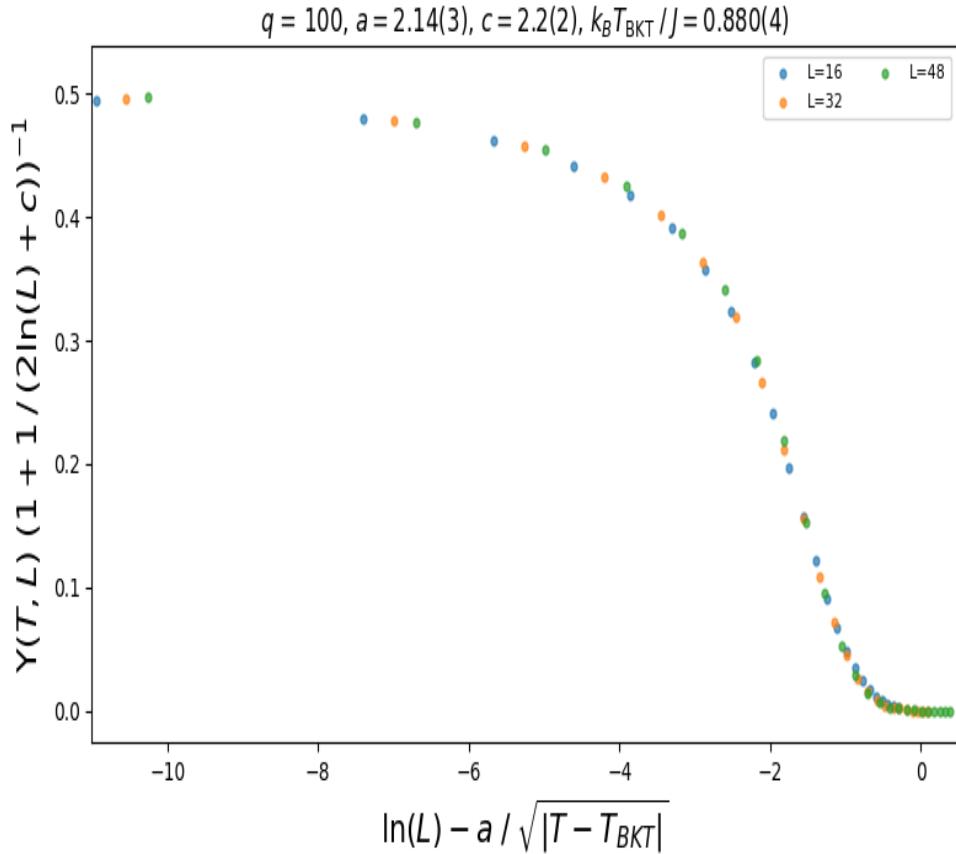


Figure 25 – Collapse of Υ for $q = 100$ according to Eq. 64.

In Fig. 26, a phase diagram is made from the critical temperatures calculated with U_4 , U_ϕ and Υ . These critical temperatures are also summarized in table 2. The critical temperatures obtained are in relatively good agreement with those obtained by the authors in reference 27. However, estimates of the critical temperature from the LRO to the QLRO phase for $q = 100$ were not able to be made, as can be seen from U_ϕ in Fig. 24: at $T = 10^{-2}$, $U_\phi \approx 0.5$ (signaling the LRO phase), but at $T = 2 \times 10^{-2}$, $U_\phi \approx 0.25$ (signaling the QLRO phase). This transition was too abrupt to produce meaningful estimates of the critical temperature in this range (hence why in table 2 there is a question mark for $q = 100$).

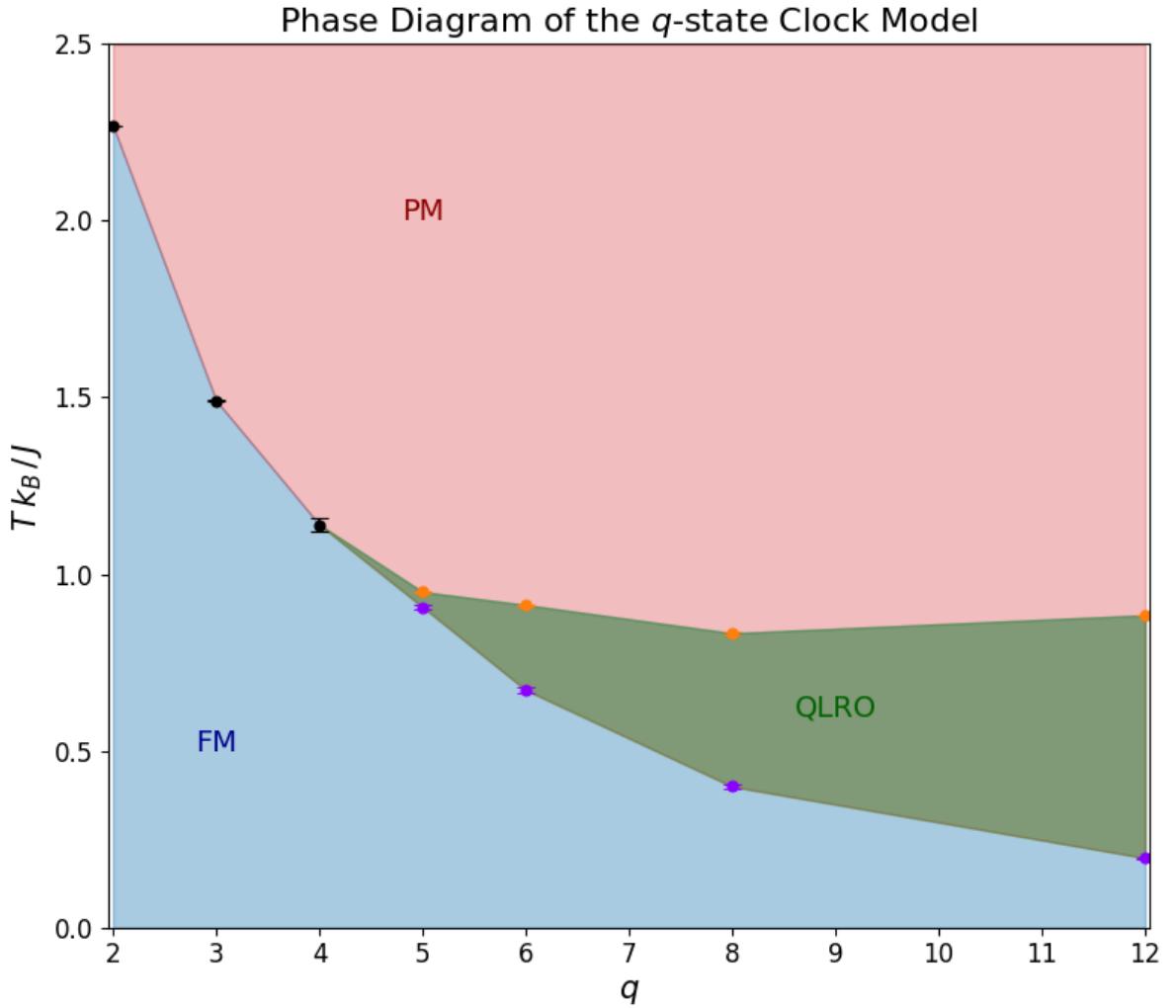


Figure 26 – Phase diagram of q -state clock model. Simulated systems were $q = 2, 3, 4, 5, 6, 8, 12$. For $q \geq 5$, the QLRO phase appears.

	$k_B T_c / J(\text{LRO-PM})$	$k_B T_c / J(\text{LRO-QLRO})$	$k_B T_c / J(\text{QLRO-PM})$
$q = 2$	2.269185(1)	—	—
$q = 3$	1.491(1)	—	—
$q = 4$	1.14(2)	—	—
$q = 5$	—	0.907(7)	0.950(1)
$q = 6$	—	0.672(8)	0.913(2)
$q = 8$	—	0.399(5)	0.833(5)
$q = 12$	—	0.197(2)	0.884(1)
$q = 100$	—	?	0.880(4)

Table 2 – Numerical calculations of the 3 possible transition temperatures (T_c) in the q -state clock model for various values of q . T_c for $q = 2$ and $q = 3$ were obtained from U_4 ; T_c for $q = 4$ was obtained from U_ϕ ; T_c (LRO - QLRO) for $q \geq 5$ were obtained from U_ϕ and T_c (QLRO-PM) for $q \geq 5$ were obtained from Υ . Only T_c for $q = 100$ from the LRO to the QLRO phase could not be meaningfully calculated.

4.2 2D XY $\mathbf{J}_1 - \mathbf{J}_2$ model

Because of the clear tendency of the FM (or LRO) phase in the q -state clock model to disappear as $q \rightarrow \infty$ and because the transition temperature from the LRO phase

to the QLRO phase could not be estimated with $q = 100$, surely the LRO phase must occupy a very small temperature range for this value of q (at most it should be the range $(0, 2 \times 10^{-2})$). Furthermore, since the discretization of spins in the q -state clock model makes its simulation much faster²⁴ when compared to continuous spin models, such as the 2D XY $J_1 - J_2$ model, instead of simulating the 2D XY $J_1 - J_2$ model directly, the q -state $J_1 - J_2$ model with very large q value was simulated (the only change made was to make the spins discrete and be able to occupy many values as an approximation to a continuous spin). Namely, for all of the following Figs. a value of $q = 500$ was used and only the temperature range with $T \geq 2 \times 10^{-2}$ was used to guarantee that the LRO phase wouldn't appear.

The logic behind this discretization to the 2D XY $J_1 - J_2$ model was the following: continuous spin models demand costly operations at each MC step, so only relatively small system sizes would be simulated. Small lattice sizes means that finite-size effects would play a central role in systematically altering results for the worse. Even if additional finite-size corrections were to be used in the calculations of critical temperatures and critical exponents, more parameters would have to be fitted in the data collapses, making the errors significantly larger (albeit the results slightly better). On the other hand, if the q -state $J_1 - J_2$ model were used, much larger system sizes would be available for analysis even with large q . Of course this would only be an approximation to the 2D XY $J_1 - J_2$ model, but the error made in this approximation would be very small. This can be seen in theory by considering the energy difference in a bond between nearest neighbor spins²⁵ when they are aligned and when they have a difference in angle due to this approximation: $\delta E(q)/J_1 = 1 - \cos(2\pi/q)$. If $q = 5$ were to be used as this approximation, then $\delta E(q = 5)/J_1 \approx 0.69$, which is on the order of the thermal energy near the transition temperatures between the QLRO and PM phases in the q -state clock model (*c.f.*, table 2): $E_{\text{thermal}}/J = k_B T_{\text{thermal}}/J \sim 0.9$. On the other hand, if $q = 500$ is used, then $\delta E(q = 500)/J_1 \approx 8 \times 10^{-5} \ll 0.9$, meaning that the maximum energy difference per bond (due to J_1 and, thus, for J_2 as well) owing to the discretization of spins will be far smaller than the natural thermal scale near the QLRO-PM phase transition in the q -state clock model. In other words, the error in this discretization approximation should be negligible.

Even if this theoretical digression wasn't convincing, all of the main results that have been obtained agree with the literature on this subject, namely the phase diagram in Fig. 29 and the identification of the nematic-PM phase transition with the 2D Ising universality class. Therefore, evidently this approximation wasn't a bad one.

In Fig. 27, the energy per site, nematic order parameter, specific heat and nematic

²⁴ Because of the precalculation of exponentials.

²⁵ Only nearest neighbors are considered because the ratio J_2/J_1 will always assume a value between $[0, 1]$ in the following results. Thus, only the J_1 energy will be considered since it will always be larger than J_2 .

order parameter susceptibility for the 2D XY $J_1 - J_2$ model with $J_2/J_1 = 0.9$ are shown. These results were obtained from the Metropolis algorithm, using 5×10^5 MC steps for equilibration and 10^6 MC steps for measurements. These observables are analogous to those in Fig. 15, but for the 2D XY $J_1 - J_2$ model.

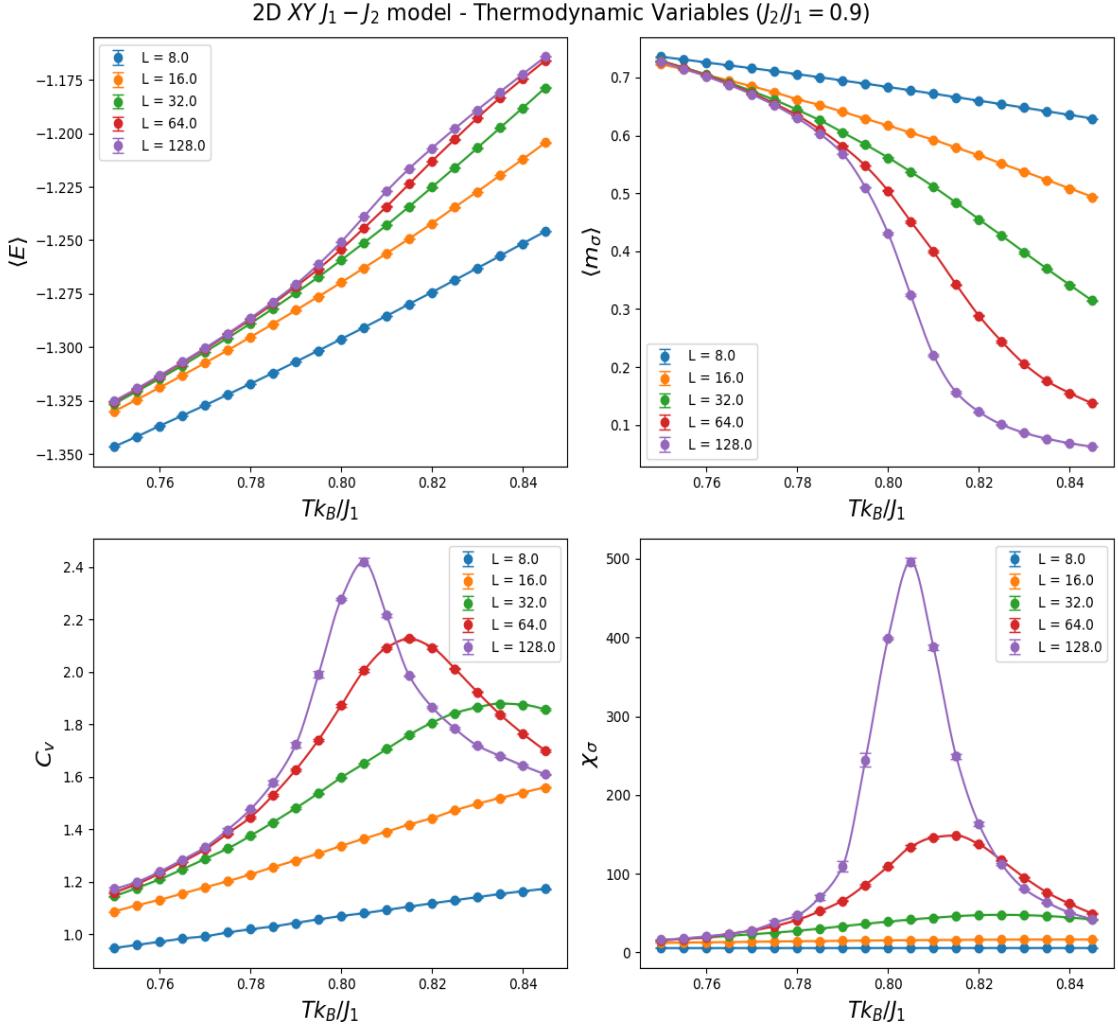


Figure 27 – Energy per site, nematic order parameter, specific heat and nematic order parameter susceptibility for the 2D XY $J_1 - J_2$ model with $J_2/J_1 = 0.9$.

In Fig. 28, the nematic order parameter Binder cumulant and spin stiffness for the 2D XY $J_1 - J_2$ model are shown in the two regimes where $J_2/J_1 = 0.8$ and $J_2/J_1 = 0.2$. This is to show that the nematic order parameter gains relevance only for $J_2/J_1 > 0.5$. From Fig. 28 it should also be noted that the large error bars in the spin stiffness for $J_2/J_1 = 0.8$ are due to the fact that there was a substantial amount of correlation between observables in successive MC steps. Since the Metropolis algorithm was used, this point can be corroborated by noting that in Fig. 32, the correlation times of the nematic order parameter using the Metropolis algorithm in this same regime ($J_2/J_1 = 0.8$) are on the order of 100 MC steps, meaning that only after 100 MC steps, an independent measurement of m_σ could be attained. Nonetheless, despite these error bars, since the intersection of

the spin stiffness with the line $2T/\pi$ occurs between points with small error bars, and the additional logarithmic corrections to size given by Eq. 64 will be used, the estimates for the critical temperatures shouldn't be spurious. This is most likely the case, because the phase diagram shown in Fig. 29 agrees with the results in the literature (31).

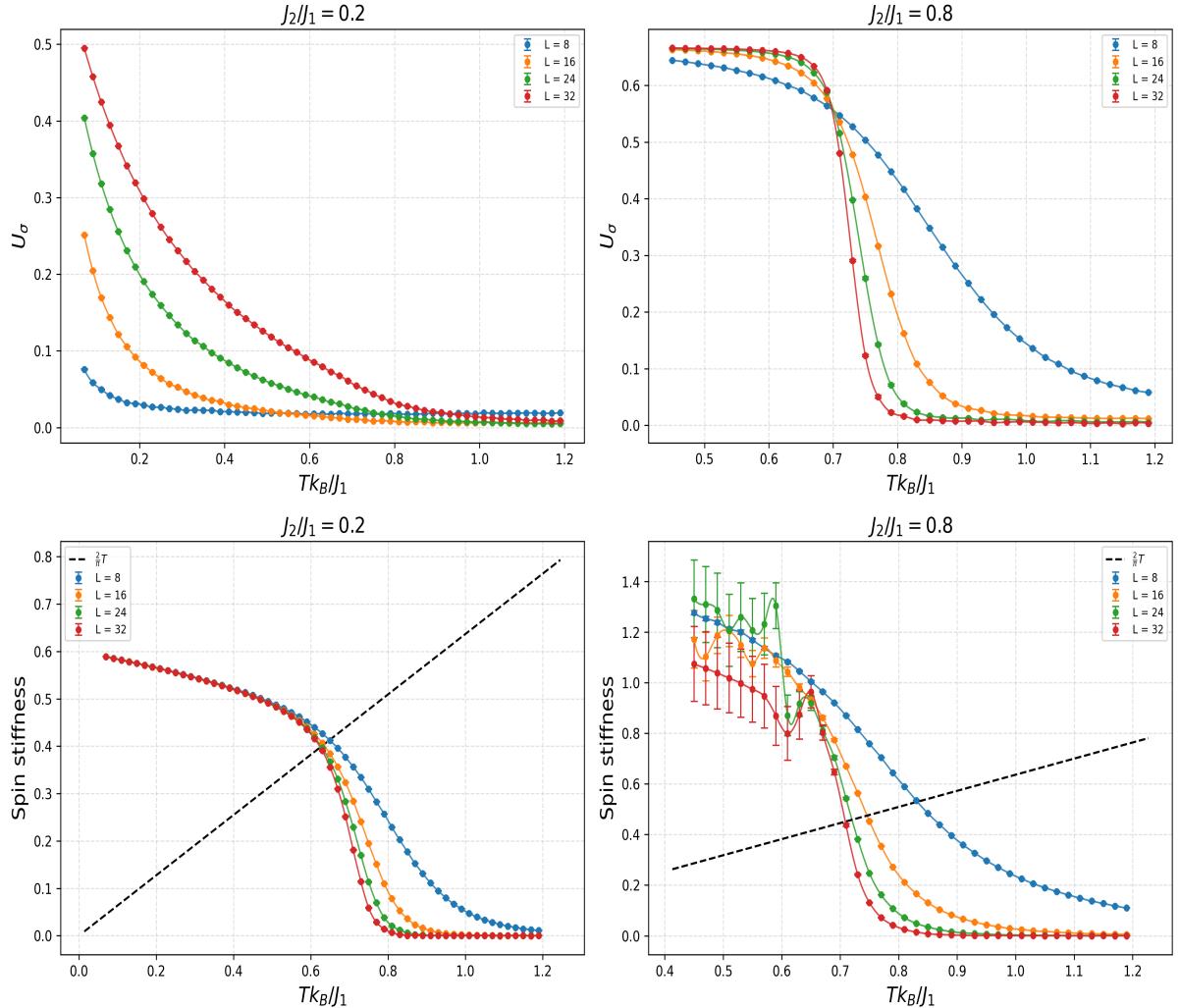


Figure 28 – Nematic order parameter Binder cumulant and spin stiffness for the 2D XY $J_1 - J_2$ model with $J_2/J_1 = 0.8$ and $J_2/J_1 = 0.2$.

With the spin stiffness, the transition temperature from the QLRO 1 phase to the PM phase can be detected for $J_2/J_1 < 0.5$ and the transition temperature from the QLRO 2 phase to the nematic phase for $J_2/J_1 > 0.5$ can be detected²⁶. The transition temperature from the nematic phase to the PM phase can be detected with the nematic order parameter Binder cumulant. These transition temperatures are summarized in table 3 and represented schematically as a phase diagram in Fig. 29.

²⁶ Once again, these critical temperatures can be estimated with the data collapse in Eq. 64. These were omitted due to their similarity in appearance with Fig. 25.

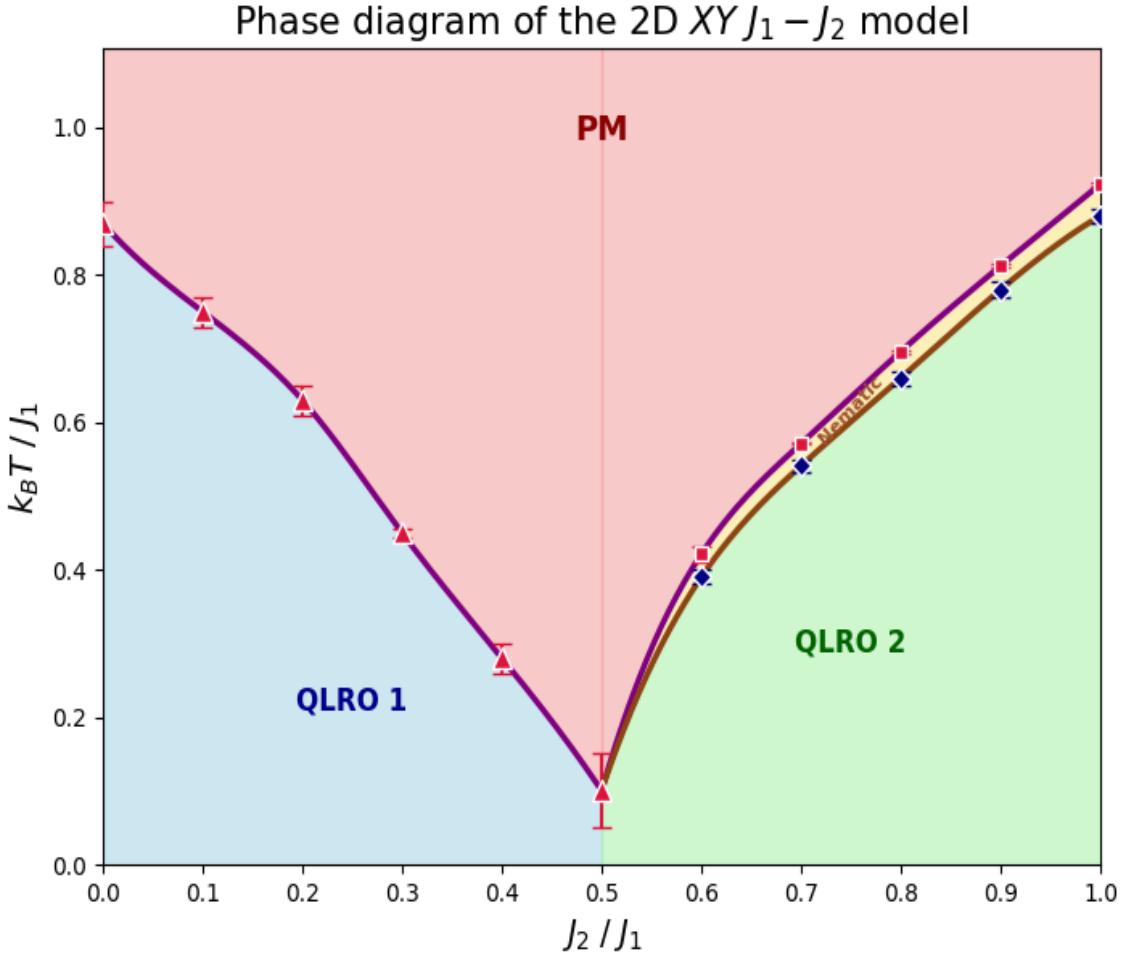


Figure 29 – Phase diagram of the 2D XY $J_1 - J_2$ model. The phase between the QLRO 2 and PM phases for $J_2/J_1 > 0.5$ is the nematic phase.

	$k_B T_c / J_1$ (QLRO1-PM)	$k_B T_c / J_1$ (QLRO2-nematic)	$k_B T_c / J_1$ (nematic-PM)
$J_2/J_1 = 0.0$	0.87(3)	—	—
$J_2/J_1 = 0.1$	0.75(2)	—	—
$J_2/J_1 = 0.2$	0.63(2)	—	—
$J_2/J_1 = 0.3$	0.450(6)	—	—
$J_2/J_1 = 0.4$	0.28(2)	—	—
$J_2/J_1 = 0.5$	0.10(5)	—	—
$J_2/J_1 = 0.6$	—	0.39(1)	0.42(1)
$J_2/J_1 = 0.7$	—	0.541(9)	0.5711(7)
$J_2/J_1 = 0.8$	—	0.66(1)	0.696(2)
$J_2/J_1 = 0.9$	—	0.78(1)	0.812(2)
$J_2/J_1 = 1.0$	—	0.88(1)	0.923(3)

Table 3 – Transition temperatures calculated from the spin stiffness and nematic order parameter Binder cumulant for the 2D XY $J_1 - J_2$ model. Lattice sizes of $L = 8, 16, 32, 64, 128$ were used.

As was hitherto noted, the transition from the nematic phase to the PM phase breaks the system's Z_2 symmetry. Thus, it's believed that this phase transition is in the

same universality class as the 2D Ising model. To numerically verify this, data collapses on C_v , m_σ , χ_σ and U_σ were performed to estimate the critical exponents: α , β , γ and ν , respectively. These data collapses are shown in Fig. 30 for $J_2/J_1 = 0.7$ (the other values of $J_2/J_1 > 0.5$ have collapses similar in appearance so they were omitted). First, T_c and ν were calculated with U_σ , then these values were used in the other collapses. The critical exponents are estimated for $J_2/J_1 = 0.6, 0.7, 0.8, 0.9$ and 1.0 , and are summarized in table 4. The results suggest that this phase transition has critical exponents that are experimentally indistinguishable from the 2D Ising model's critical exponents, making it plausible that they are indeed in the same universality class.

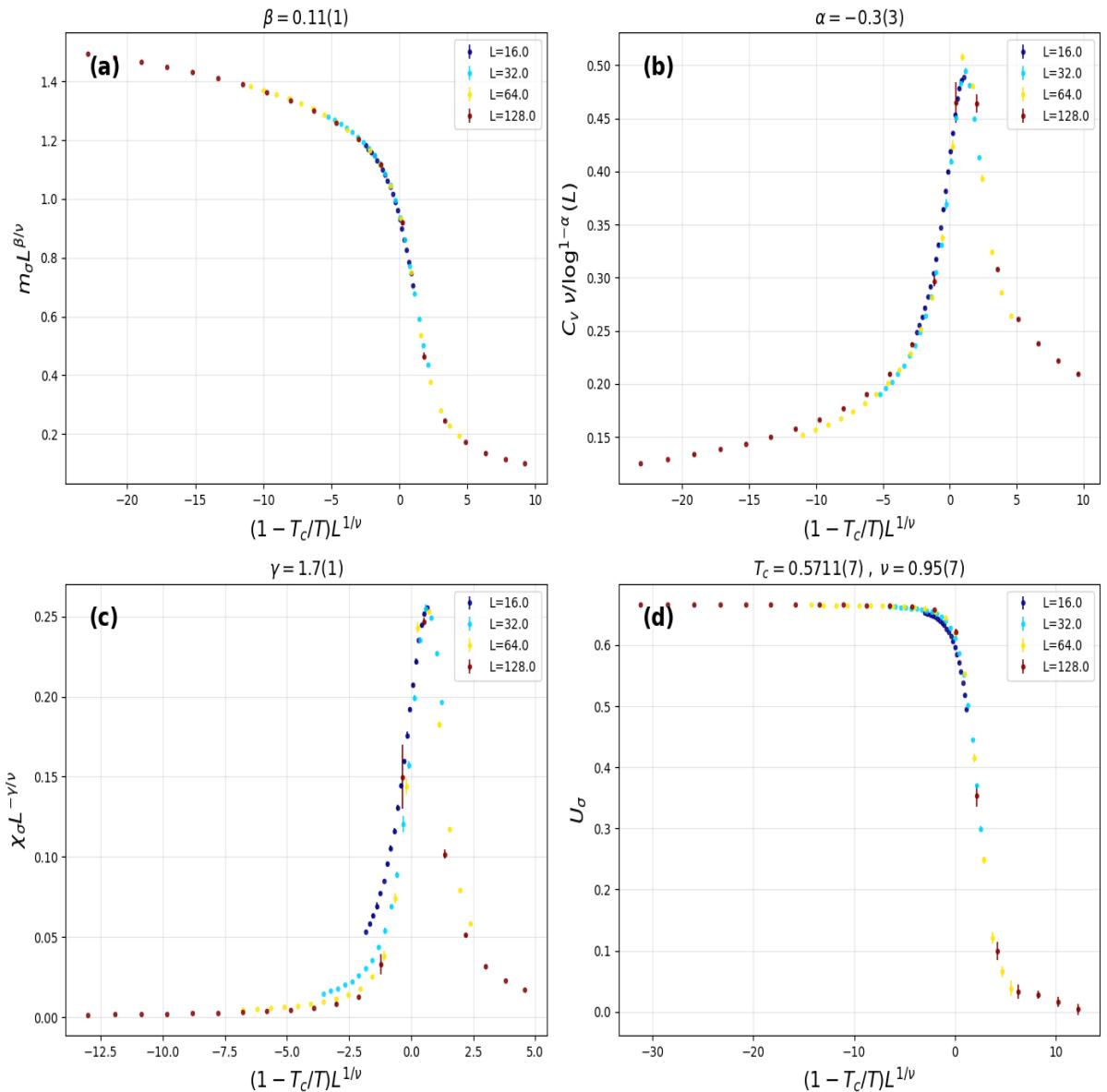


Figure 30 – Data collapses performed on data from Metropolis algorithm. All plots are from the data set with $J_2/J_1 = 0.7$. (a) Nematic order parameter: $m_\sigma L^{\beta/\nu}$ versus $(1 - T_c/T)L^{1/\nu}$. (b) Specific heat: $C_v \nu / \log^{1-\alpha}(L)$ versus $(1 - T_c/T)L^{1/\nu}$. (c) Nematic order parameter susceptibility: $\chi_\sigma L^{-\gamma/\nu}$ versus $(1 - T_c/T)L^{1/\nu}$. (d) Nematic order parameter Binder cumulant: U_σ versus $(1 - T_c/T)L^{1/\nu}$.

	α	β	γ	ν
2D Ising (Exact)	0	0.125	1.75	1
$J_2/J_1 = 0.6$	0.02(1)	0.098(8)	1.86(5)	0.94(3)
$J_2/J_1 = 0.7$	-0.3(3)	0.11(1)	1.7(1)	0.95(7)
$J_2/J_1 = 0.8$	0.03(2)	0.109(4)	1.5(2)	0.96(2)
$J_2/J_1 = 0.9$	0.1(1)	0.120(3)	1.74(9)	1.02(1)
$J_2/J_1 = 1.0$	-0.001(2)	0.124(3)	1.73(7)	1.03(2)

Table 4 – Critical exponents obtained from the transition of the nematic phase to the PM phase in the 2D XY $J_1 - J_2$ model. This is believed to be in the same universality class as the 2D Ising model.

Out of all the lines in table 4, $J_2/J_1 = 0.6$ contains the critical exponents that deviate the most from the 2D Ising universality class (specifically, the estimates of β and γ). This is most likely due to spurious convergences in the minimization subroutine. To get around this and numerically show that this phase transition ($J_2/J_1 = 0.6$) is in the same universality class as the 2D Ising model, Fig. 31 has been made, which imposes the exact critical exponents on the data. It can be seen that the data collapses are decent, so this doesn't completely eliminate the possibility that this phase transition can still be in the 2D Ising universality class (even though the estimates for β and γ aren't great).

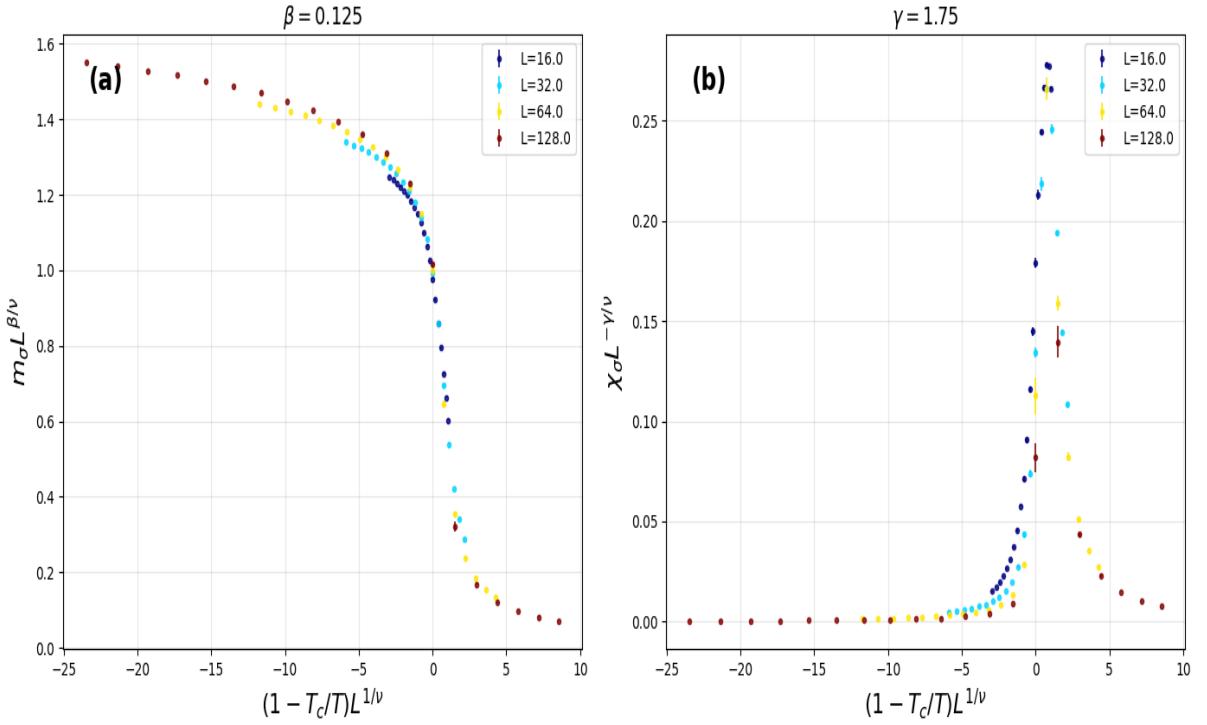


Figure 31 – Data collapses performed on $J_2/J_1 = 0.6$ imposing 2D Ising model's exact critical exponents ($\beta = 0.125$, $\gamma = 1.75$, $\nu = 1.0$ and $T_c = 0.42$). (a) Nematic order parameter: $m_\sigma L^{\beta/\nu}$ versus $(1 - T_c/T)L^{1/\nu}$. (b) Nematic order parameter susceptibility: $\chi_\sigma L^{-\gamma/\nu}$ versus $(1 - T_c/T)L^{1/\nu}$.

Finally, to ascertain which algorithm is the best to simulate this model in the frustrated regime $J_2/J_1 > 0.5$, in Fig. 32, the correlation times of the nematic order parameter calculated for the Metropolis, heat bath and Wolff algorithms are shown (for

$J_2/J_1 = 0.8$ and $L = 16$). It can be concluded that the Wolff algorithm is by far the worst choice, because not only is the correlation time roughly an order of magnitude greater for all temperatures simulated, but the mean cluster size over the course of many MC steps in a single temperature are on the order of the lattice size. In other words, the Wolff algorithm spends a lot of time creating and rotating large clusters of spins, but the measurements of the nematic order parameter are still very correlated. Between the heat bath and Metropolis algorithm, the better choice isn't as obvious as Fig. 32 might suggest, because even though the heat bath algorithm manages to decorrelate measurements slightly better than the Metropolis algorithm, the heat bath algorithm takes significantly longer to terminate. Therefore, maybe a mix of MC steps from both algorithms would be superior.

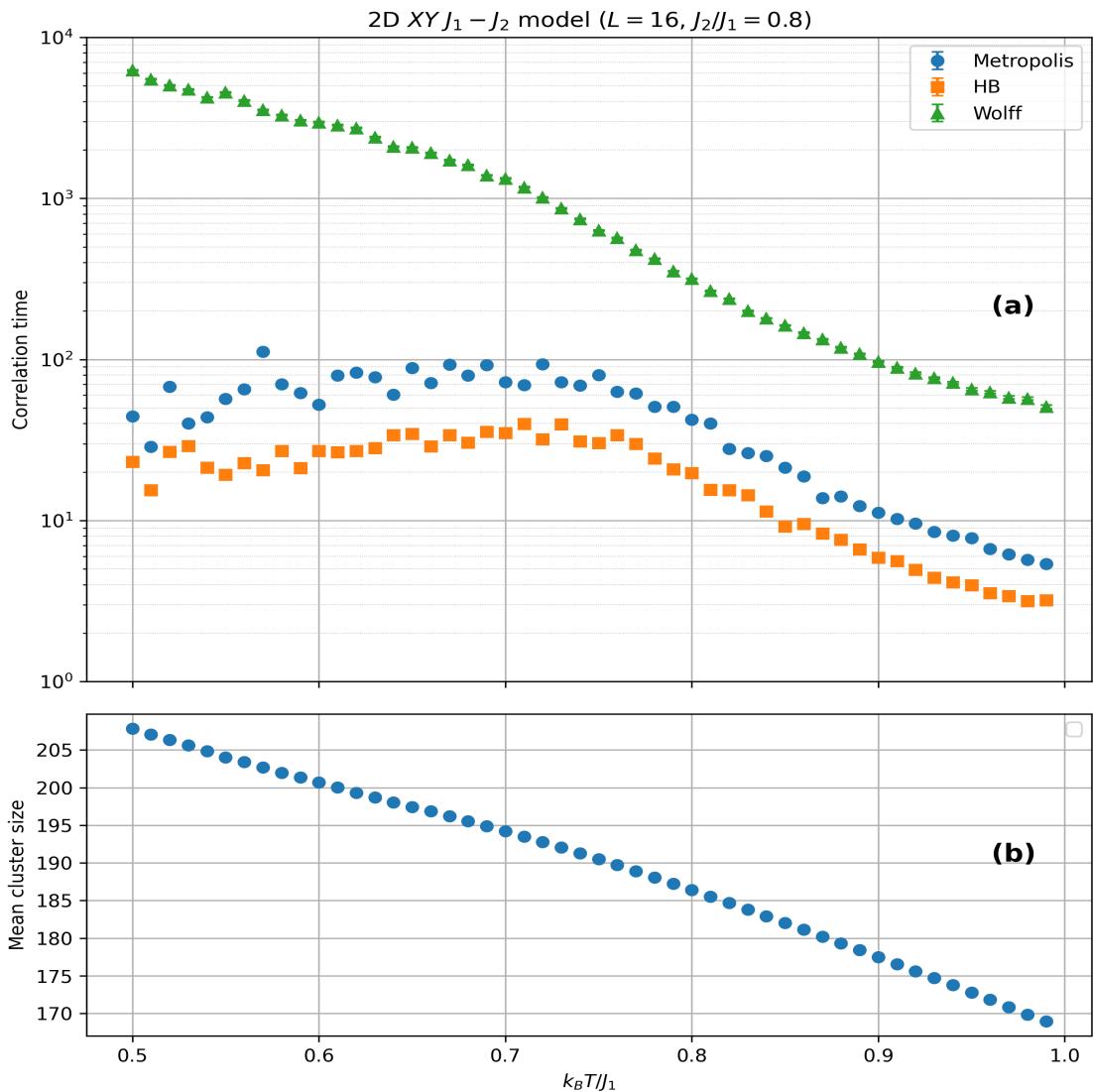


Figure 32 – (a) Correlation times of the nematic order parameter calculated for the Metropolis, heat bath and Wolff algorithms when simulating the 2D XY $J_1 - J_2$ model with $L = 16$ and $J_2/J_1 = 0.8$ (Y axis is log scale). (b) Average cluster size formed in the Wolff algorithm for this same system ($L^2 = 256$) over the course of 10^5 MC steps.

5 CONCLUSION

The objective of this work was to study mean-field theory and Monte Carlo techniques applied to phase transitions in the q -state clock model and 2D XY $J_1 - J_2$ model. The critical exponents for both models were obtained with Weiss' molecular field theory and Ginzburg-Landau theory. Then, using finite-size scaling, the critical temperature and critical exponents for the 2D Ising model were obtained through Monte Carlo simulations (with the Wolff algorithm), by verifying the intersection points of the Binder cumulant U_4 and “collapsing” the data generated from thermodynamic variables, respectively. They agree relatively well with the exact analytical results (*c.f.* table 1). The estimations of the dynamical critical exponents show that the Wolff algorithm is better than the Metropolis algorithm in resisting the critical slowing down effect near the critical temperature of the 2D Ising model.

By using Binder cumulants (U_4 and U_ϕ) and spin stiffness (Υ), simulations on the q -state clock model with $q = 3, 4, 5, 6, 8, 12, 100$ were performed and their critical temperatures were estimated to yield the phase diagram in figure 26. For $q \geq 5$, a new intermediate quasi-long-range order phase appears between the paramagnetic and ferromagnetic phases, exhibiting a BKT transition that ends up replacing the ferromagnetic (long range order) phase for $q \rightarrow \infty$.

Finally, using the spin stiffness and Binder cumulant of the nematic order parameter, a phase diagram of the 2D XY $J_1 - J_2$ model was obtained (Fig. 29). Through data collapses on the specific heat, nematic order parameter, nematic order parameter susceptibility and nematic order parameter Binder cumulant, the critical exponents for the transition from the nematic phase to the paramagnetic phase were obtained, for $J_2/J_1 > 0.5$, and are summarized in table 4. By comparison to table 1, this transition appears to be in the same universality class as the 2D Ising model. Lastly, the Wolff algorithm was shown to be more inefficient than the Metropolis and heat bath algorithms in simulating the 2D XY $J_1 - J_2$ model with $J_2/J_1 > 0.5$ because it generates successive microstates that are much more correlated with one another than the single-spin-flip algorithms.

6 REFERENCES

1. N. D. Mermin and H. Wagner, **Absence of Ferromagnetism or Antiferromagnetism in One- or Two-Dimensional Isotropic Heisenberg Models**. Phys. Rev. Lett., **17**, 1133 (1966).
2. L. D. Landau and E. M. Lifshitz, **Course of Theoretical Physics: Volume 5, Statistical Physics, part 1**. 3rd ed. Pergamon Press, 1980. p. 451-498.

3. R. K. Pathria and R. D. Beale, **Statistical Mechanics**. 3rd ed. Elsevier, 2011. p. 442-463.
4. S. R. A. Salinas, **Introduction to Statistical Physics**. Springer New York, 2001.
5. D. V. Schroeder, **An Introduction to Thermal Physics**. Addison-Wesley, 2000. p. 343-356.
6. M. E. J. Newman and G. T. Barkema, **Monte Carlo Methods in Statistical Physics**. Oxford University Press, Oxford, 1999.
7. D. Landau and K. Binder, **A Guide to Monte Carlo Simulations in Statistical Physics**. Cambridge University Press, Cambridge, 2000.
8. N. Metropolis, et al, **Equation of State Calculations by Fast Computing Machines**. Journal of Chemical Physics **21**, 1087 (1953).
9. Y. Miyatake, M. Yamamoto, J. J. Kim, M. Toyonaga and O. Nagai, **On the implementation of the “heat bath” algorithms for Monte Carlo simulations of classical Heisenberg spin systems**. Journal of Physics C: Solid State Physics **19**, 2539 (1986).
10. M. Abramowitz and I. A. Stegun (Eds.), **Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables**. Vol. 55, Dover, New York (1964). p. 376, Eq. 9.6.16 .
11. U. Wolff, **Collective Monte Carlo Updating for Spin Systems**. Phys. Rev. Lett. **62**, 361 (1989).
12. R. H. Swendsen and J. S. Wang, **Nonuniversal critical dynamics in Monte Carlo simulations**. Phys. Rev. Lett. **58**, 86 (1987).
13. M. M. J. Miranda, I. C. Almeida, E. C. Andrade and J. A. Hoyos, **Phase diagram of a frustrated Heisenberg model: From disorder to order and back again**. Phys. Rev. B, **104**, 054201 (2021).
14. M. Weigel, **Connected component identification and cluster update on GPU**. arXiv:1105.5804v2 (2011).
15. Y. Komura and Y. Okabe, **GPU-based Swendsen-Wang multi-cluster algorithm for the simulation of two-dimensional classical spin systems**. Journal of Computer Physics Communications, **183**, 1155 (2012).
16. J. M. Kosterlitz and D. J. Thouless, **Ordering, metastability and phase transitions in two-dimensional systems**. Journal of Physics C: Solid State Physics **6**, 1181 (1973).

17. D. R. Nelson and J. M. Kosterlitz, **Universal Jump in the Superfluid Density of Two-Dimensional Superfluids.** Phys. Rev. Lett., **39**, 1201 (1977).
18. M. G. Kendall and A. Stuart, **The Advanced Theory of Statistics, Vol. 1: Distribution Theory**, 3rd ed., Charles Griffin & Company, London (1969).
19. A. W. Sandvik, **Computational Studies of Quantum Spin Systems.** AIP Conference Proceedings **1297**, 135 (2010). p. 155-173.
20. S. K. Baek, P. Minnhagen, B. J. Kim, **True and quasi-long-range order in the generalized q -state clock model.** Phys. Rev. E, **80**, 060101 (2009).
21. S. M. Bhattacharjee and F. Seno, **A Measure of data-collapse for scaling.** arXiv:cond-mat/0102515v2, 2001.
22. W. H. Press, *et al.*, **Numerical Recipes: The Art of Scientific Computing.** 3rd ed. Cambridge University Press, 2007.
23. M. Luscher, **A portable high-quality random number generator for lattice field theory simulations.** Computer Physics Communications, **79** (1994), p.100.
24. J. Hahnfeld and L. Moneta, **A Portable Implementation of RANLUX++.** arXiv:2106.02504 (2021).
25. A. E. Macias-Medri *et al*, **Speedup of the Metropolis protocol via algorithmic optimization.** Journal of Computational Science, **66**, 101910 (2023).
26. L. Onsager, **Crystal statistics. 1. A two-dimensional model with an order-disorder transition.** Phys. Rev., **65**, p. 117 (1944).
27. S. Chatterjee, *et al.* **Ordering kinetics in the q -state clock model: scaling properties and growth laws.** Phys. Rev. E, **98**, 032109 (2018).
28. J. A. Nelder and R. Mead, **A simplex method for function minimization.** The Computer Journal, **7** (1965), p.308.
29. C. L. Henley, **Ordering Due to Disorder in a Frustrated Vector Antiferromagnet.** Phys. Rev. Lett. **62**, 2056 (1989).
30. C. L. Henley and S. Prakash, **Ordering Due to Disorder in a Frustrated XY Antiferromagnet.** J. Phys. Colloques, **49**, C8 (1988).
31. F. F. Song, *et al.*, **Emergent Intermediate Phase in the $J_1 - J_2$ XY model from Tensor Network Approaches.** arXiv:2412.18892 (2024).