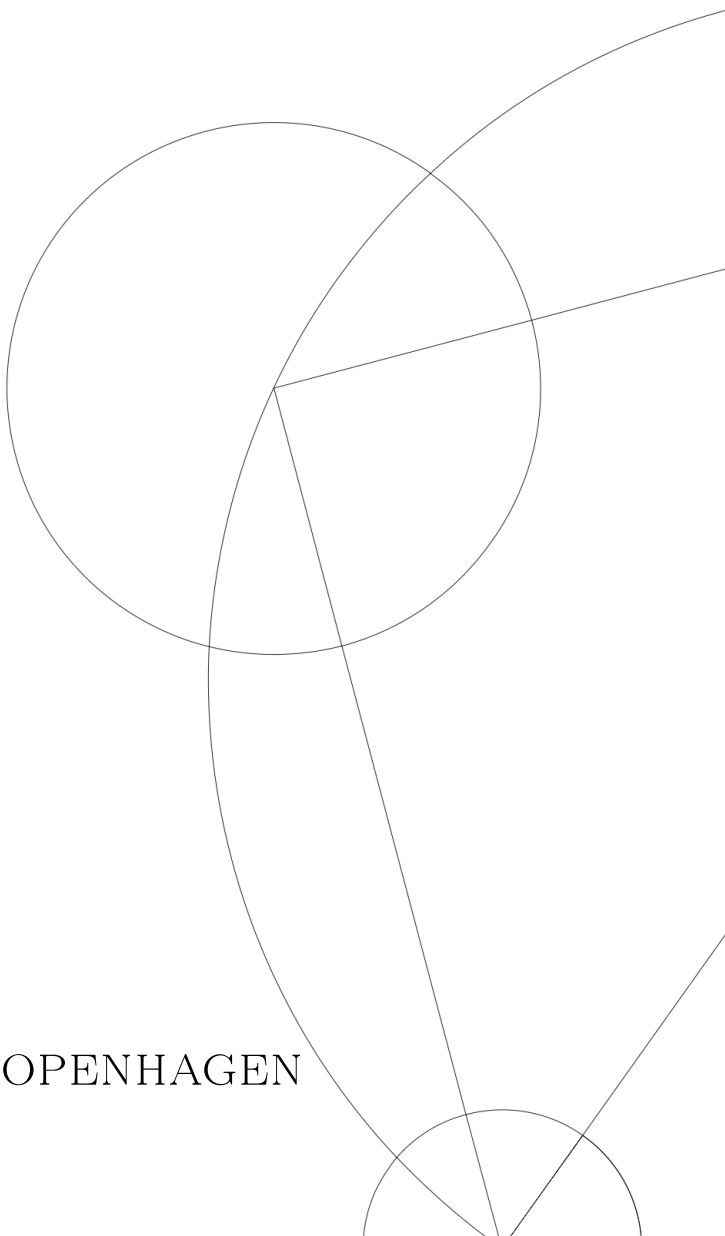
Temporal behaviour of the 2D Ising model

Quantifying hidden order of non-equilibrium transitions in the Ising model

Written by

Morten Holm (QGF305)

31. Jan. 2022



University of Copenhagen

Contents

1 Introduction

Less than a century ago in 1925, Ernst Ising whom the Ising model is named after, as part of his PhD, used a 1-dimensional string of magnetic moments that could be either spin up or spin down and calculated its statistical properties in thermodynamic equilibrium [1]. This is a system that displays complex behaviour, i.e. non linearity, due to its coupled interactions and random flipping of spin states. One example is a ferromagnetic system [7, Sec. 6.4.2] wherein iron can become magnetised in a magnetic field resulting in aligned spins, if heated, it loses magnetisation due to the opposite spin states cancelling out. Ernst concluded that the behaviour of the Ising model showed no phase transition for ferromagnetic ordered states in 1-dimension at any finite temperature and made the conclusion that no phase transition existed for any dimension. Reimagining Ernsts' Ising model as a single 1-dimensional domain [1] of down spins of a given length within a sea of spin up. We can imagine that expanding beyond the single domain of complex behaviour surely displays more interesting interactions. This was proven when Lars Onsager presented an exact analytical solution for the thermodynamic properties of the 2-dimensional Ising model in 1944. He demonstrated that a phase transition does exist for configurations in higher dimensions. His discovery has been important to science as phase transitions occur everywhere in nature and has been demonstrated, but not proven, to have universal properties. For Ernsts' ferromagnetic Ising model, these universal properties are the dimension of the system, the range of the interactions and the spin dimension. This universality exists as spatial indicators for the behaviour as a system approaches a phase transition and are defined using critical exponents. Now using a recently developed approach of compression based entropy, we can further show that additional such behaviour is indicated to exist not only for the spatial but also the temporal behaviour of the 2-dimensional Ising model. This could be important for many fields since the model was first used to describe ferromagnetism but has since then been used to describe many different systems from condensed matter, to Immunology and descriptions of long-range correlations in DNA sequences [4] where a temporal link could be important for the outcome.

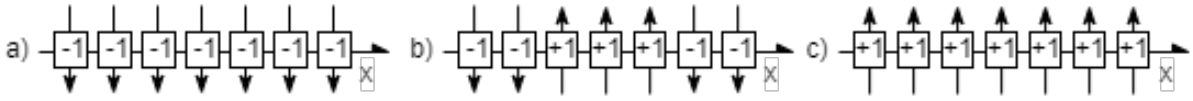


Figure 1: 1-dimensional system displaying spin flipping behaviour in time. a) 1-D chain consisting of coupled spins of spin down. b) Spin flipping several times where we see that three states have flipped their state from spin down to spin up. c) All spins have switched spin state to spin up.

2 Theory

2.1 Ising Model

We may consider N particles configured in a system of length, L, then the total amount of particles in a given dimension, d, is that of $N = L^d$ particles. Each particle demonstrates the behaviour of dipole moments also known as the magnetic spin, s, for which it can be either spin up or spin down, $s = \pm 1$, as seen in figure 1. This is also known as the spin-$\frac{1}{2}$ Ising model for which many other configurations exist [15, ch. 3], [3, pg. 139]. The spins will interact in pairs with a given energy configuration for given spin alignments. In using the ferromagnetic case where spins are aligned, we have a thermodynamic limit as the size of our system, N and with no external field as the average energy of the ground state per site of $E_0/N = -2J$. With an increase in dimension, the number of neighbours that a single particle experiences will also increase as an order of 2^d neighbours.

[1] In a given magnetic material, magnetic domains are regions of uniform direction. As such, a domain is a state of a ferromagnet in which the magnetisation does not vary across the magnet.

2.1.1 2D Ising model

In this paper, we will be examining the 2-D spin-$\frac{1}{2}$ Ising model which consists of a lattice configuration in 2 spatial dimensions, thus with $N = L^2$ particles. This spin interaction will happen according to the following generalised Hamiltonian

$$H = H_{spin} + H_{ext} = -\sum_{<i,j>} J s_i s_j - \sum_j h s_j. \tag{1}$$

The Hamiltonian consists of two terms, the first term is the the spin-spin interaction, H_{spin}. For which J is the exchange energy[2] between neighbouring particles and $< i, j >$ is a shorthand for the double sum over nearest neighbours and counts each spin pair only once.
The second term, H_{ext}, denotes the external force on the system in the form of a magnetic field with strength, h, and s_j being the individual spin site.
With no external magnetic field applied we have the total energy as

$$E = -J \sum_{<i,j>} s_i s_j. \tag{2}$$

For the exchange energy, we can outline its three cases resulting in the following energy configurations

$$J > 0 \Rightarrow \text{Ferromagnetic}$$
$$J = 0 \Rightarrow \text{No spin interaction}$$
$$J < 0 \Rightarrow \text{Antiferromagnetic}$$

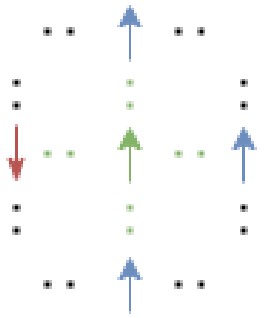


Figure 2: Demonstration of the bonding to nearest neigbours in a 2D lattice and the spin states. Green being the center spin coupling to four neighbours. Blue of spin up and Red of spin down. For spin flip dynamics, the resulting energy configuration of this particular setup for flipping the green spin to spin down would be $\delta E = 4J$

For $J > 0$ the interaction will be ferromagnetic, in that the spins will align. If for each pair the exchange energy $J = 0$, no spin interaction happen. The last case is when $J < 0$ which is the antithesis, in that it will be antiferromagnetic and will anti-align spins. The distribution of spins in the lattice will depend on temperature according to the Boltzmann distribution, thus the probability of a certain configuration will be the frequency distribution of

$$P(s) = \frac{e^{-E(s)/kT}}{Z} = \frac{e^{-\beta E(s)}}{Z}. \tag{3}$$

For the Boltzman constant, k and T the temperature of the system. Using the inverse temperature conversion of $\beta = (kT)^{-1}$ and Z being the partition function, also known as the normalisation constant of $Z = \sum_s e^{-\beta E(s)}$. Using the formulation of eq 3, the probability of a spin flip from state 1 to state 2 can be written as $P_{12} \propto exp(-\beta E_{12})$, where $E_{12} = E_2 - E_1$. For our spin lattice at a given value of T, a spin may or may not flip. We can differentiate between micro and macro states, where the micro state is a given spin configuration and a macro state is a given value of T which corresponds to many micro states. Suppose we have a given value of T, it follows that for $E_2 < E_1 \Rightarrow E_{12} < 0$ and it is therefore more likely flipping to a lower energy state as $P_{12} > P_{21}$. Similarly if $|E_{12}| \ll \beta^{-1}$ it is equally likely to flip in either direction as $P_{12} \approx P_{21}$. This will become important in a later section regarding the Metropolis algorithm when we discuss spin flip dynamics, as calculating the energy of a marcoscopic state $E = \sum_\alpha E_\alpha P_\alpha$ is computationally expensive.

[2] J is also often called the exchange integral or exchange constant.

2.2 Properties of the Ising model

The Ising model can be thought of as having several properties. Using the quantities of the partition function, energy and temperature from section 2.1.1 we can find these properties.

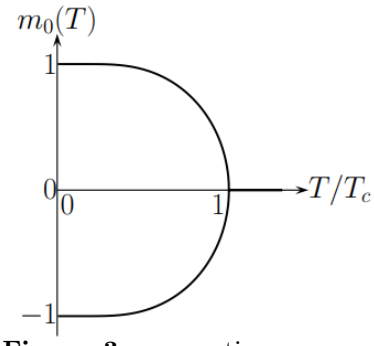


Figure 3: magnetisation for the 2D ising model in zero external field

2.2.1 Order parameter

The average magnetisation is the direct sum over all the configurations using the magnetisation itself and the frequency of the configurations

$$M_{avg} = \sum mexp(-\beta E). \tag{4}$$

This is also known as the order parameter, how ordered the magnetic spins are is simply the average magnetisation per spin as described by equation 4. [3, pg.127] Explicitly we have the average magnetisation per spin as

$$m(T,H) = \sum_{s_i} p_{s_i} m_{s_i} = \frac{1}{Z} \sum_{s_i} e^{(-\beta E_{s_i})} m_{s_i}. \tag{5}$$

When the temperature is low, we have a completely aligned state, as seen on figure 3. Our states are either in a spin up or down configuration, as the temperature increase the point of convergence of the pitchfork bifurcation is also the point of phase transition from an ordered state to a disordered state.

2.2.2 Magnetic susceptibility

With a description of the magnetisation and the partition function, we can find other properties of the model such as the magnetic susceptibility[7, Sec. 6.4], which tells us of the potential magnitude of the systems' capability of magnetisation by an external field

$$\chi(T,0) = \left(\frac{\partial m}{\partial H}\right)_T. \tag{6}$$

2.2.3 Heat capacity

The specific heat capacity, which is closely related to energy, is a measure of the change of the temperature due to the addition of energy at a constant magnetic field. The heat capacity at a constant magnetic field can be expressed as

$$C_v = \frac{\partial < E >}{\partial T} \approx \frac{1}{|T - T_c|^{\alpha}}. \tag{7}$$

The behaviour of the system can be described by the power law, α and we observe a discontinuity in the first derivative of energy w.r.t. time at the point T_c, later in the section of results we see this discontinuity in figure 7.

2.2.4 Correlation

How a spin at a given site influence another spin site can be determined using the correlation function. For any two points in our lattice we can examine the correlation between the spin sites using the expectation value $< s_i s_j >$[15, Sec. 2.4]. The correlation length, $\xi(T)$, also called the characteristic length is also a power law, $\xi(T) \sim |T - T_c|^{-\nu}$ and the correlation length gives the length scale for the decay of correlations between the spins. The independence of the spin observables, S, can be investigated by calculating the autocorrelation of the data for different lag times in succession. This autocorrelation is then given by

$$r_{ij} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} S_{mn} \bar{S}_{m+i,n+j}. \tag{8}$$

We can see this as averaging over an ensemble of realizations of the random variable and given that it is a random process, x(t) that evolves in time, t, we can write the autocorrelation function as

$$C(t) =< x(t)x(t+\tau) > \tag{9}$$

Usually, a correlation function makes the distinction between cross-correlation and auto-correlation. That is whether we have the two samples or one sample to correlate. As we only have spins from a single source, we can use the Wiener-Khinchin Theorem, which is a special case of the cross-correlation theorem for when $f(t) = g(t)$, i.e. the same sample[14]. This theorem utilises a Fourier transform and the formulation of the is then

$$C(t) = \mathcal{F}[|A|^2](t). \tag{10}$$

The Correlation function is then used to find the characteristic length,ξ and time, t. The correlation length we can find by applying the autocorrelation function for the lattice size for the time evolution of our system. This in turn yields a matrix of lag times in respectively x and y that then make up radial lag.

2.3 Phase Transitions: Criticality and Universality

The before mentioned power laws play a key role in the behaviour of a system such as this. A phase change in our system is a transition from one state to another. Specifically, we are moving from a completely ordered state where the system consists of particles of either entirely spin up or spin down and into a system of a completely disordered nature. In physical phase space, we have a critical point as defined by the divergence of the specific heat and susceptibility [15][3]. We are therefore interested in finding the critical temperature known as the Curie temperature, T_c that signifies where the singularity occurs. It can be a good idea to refactor the thermodynamic quantities as a function of the reduced temperature instead, which will centre the transition around zero. As the reduced temperature approaches zero at the critical point we have $\beta \to \beta_c$ and will therefore better explain the behaviour as we approach the critical point

$$t \equiv \frac{T - T_c}{T_c} \quad \texttt{or} \quad t \equiv \frac{\beta_c - \beta}{\beta_c}. \tag{11}$$

For the properties explained in the previous section, we can find so-called critical exponents that signify some universal behaviour. For the specific case of our 2D spin-$\frac{1}{2}$ Ising model the following quantities can be seen as some *quantity* $= A|t|^p$ for p being a power and A being some constant[15, Sec. 2.6]. Using eq 5, the critical exponent can be found and in Onsager's original formulation he found[4]

$$M = (1 - sinh(2/x)^{-4})^{1/8} \tag{12}$$

However, omitting the constant we have the critical exponents of a given property for our model as

$$M \sim |t|^{\beta}, \quad \beta = \frac{1}{8} \qquad\qquad C \sim |t|^{-\alpha}, \quad \alpha = 0$$

$$\chi \sim |t|^{-\gamma}, \quad \gamma = \frac{7}{4} \qquad\qquad \xi \sim |t|^{-\nu}, \quad \nu = 1$$

It is important to note that more critical exponents exist, but here we will only briefly mention the correlation length as it is particularly interesting when describing the system. The correlation length tells us the correlation for clusters of a particular spin direction and the physical meaning that the system organises itself into uniform blocks with a size approximately that of the correlation length, ξ. As $\xi \to \infty$ we have $T \to T_c$ and the correlation length diverges. A spatial structure smaller than ξ cannot be resolved, i.e. a big block of spin up or down.

[3]This divergence can be seen from figure 12 c) and d)

[4]The derivation of which has been found in numerous publications and books such as in [15] or even on Wikipedia.

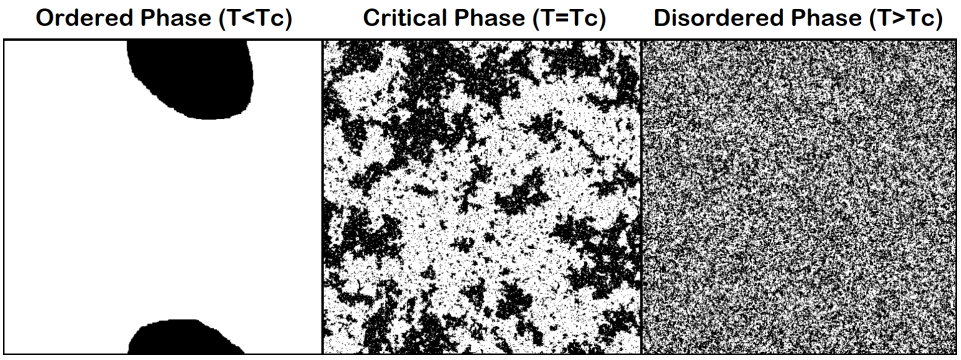


Figure 4: Image demonstrating the spin directions of spin up (white) and spin down (black) for three-phase states; from the left, we see $T < T_c$ meaning almost entirely spin up and in this specific case going to T=0 would mean only having spin up resulting in the entire space becoming white. In the middle we see how distinct clusters have formed, in this space we have the correlation length approaching $\xi \to \infty$ corresponding to $T = T_c$. At that point, a spin down cluster would occupy the length of the system. In the last image, it is shown how $T > T_c$ affects the disorder where all clusters have broken up into almost individual spins akin to white noise.

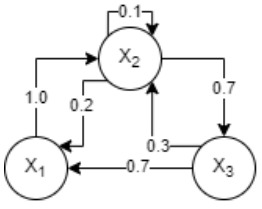
These values can also change based on the so-called Universality class. This is because critical exponents allow us to divide systems into universality classes, meaning that even though a system can have different Hamiltonians, it may exhibit the same behaviour as the critical point is approached. An example of a different universality class could be a change in dimension, where we would also see critical exponents would change.

2.4 Metropolis Algorithm

To estimate average values of the thermodynamic properties we employ a Metropolis Monte Carlo simulation. Using this algorithm, new configurations are generated from a previous state using a transition probability. This is done by generating large numbers of different and independent configurations of the system so we have enough statistical sampling. [11]

2.4.1 Markov Chain Monte Carlo

In calculating the expectation value of large systems, we exceed the available computational capabilities of most if not all computers. As an example in our 2-dimensional system, $d = 2$, we have a system length, L^d for $N = L^d$ points. Given our two states s, the number of possible configurations of the system to visit can become immense, 2^N. Instead, we can construct a probability distribution and use it to sample possible system configurations. This can be done using Markov Chain Monte Carlo (MCMC), where the basic idea is that we perform a pseudo-random walk through a probability distribution, favouring values with a higher probability. We use an initial starting point and randomly pick a nearby point and evaluate its probability. Given some acceptance criteria, we either move on to the new point with some probability or stay put. Given the right conditions and time, we can visit every energy configuration with a frequency proportional to its probability. Specifically the Markov chain part of MCMC can be explained using a simple set of numbers $\mathbf{x} = \{x_1, x_2, x_3, x_4\}$. Here, any event that follows, is only dependent on the event that came before it. Such that a probability $P(x_4|x_3, x_2, x_1) = P(x_4|x_3)$. The probability of a state being reached is therefore only dependent on the previous state of the chain. From figure 5 we see how this would work in practice. Given some initial probabilities of being at each of the locations $L_{t=0} = (0.1, 0.6, 0.3)$, we can write the next step as $L_{t=1} = L_{t=0} \cdot T = (0.33, 0.25, 0.42)$.



$$\begin{bmatrix} 0 & 1.0 & 0 \\ 0.2 & 0.1 & 0.7 \\ 0.7 & 0.3 & 0 \end{bmatrix}$$

Figure 5: Example of a transition graph and its transition matrix

Given enough iterations, this model will at some point converge at (0.289, 0.418, 0.293)[5], describing the distribution of probabilities to be at each point in equilibrium. An important property of MCMC, as described, is that it is capable of visiting every point in the system also known as ergodicity. A requirement for this is that the chain does not get trapped in a cycle[6] such that for every state, there is a possibility to move to any other state[7].

2.4.2 Metropolis-Hastings

We assume a well designed transition probability (Markov process), where our previous conditions are met. The stationary distribution of our transition probability distribution matrix, T, can be chosen as the probabilities $P(x_j|x_i)$, then detailed balance is satisfied as $P(x_j|x_i)p(x_i) = P(x_i|x_j)p(x_j)$.
We can then set up a conditional probability, $U(x)$, that proposes a new state close to its current state, as well as an acceptance/rejection distribution, $A(x)$, that determines the probability of accepting the given new state

$$P(x_j|x_i) = U(x_j|x_i)A(x_j, x_i). \tag{13}$$

We therefore either accept the proposal, such that $x_j = x_i$, or otherwise reject it, but retain the old point with probability

$$A(x_j, x_i) = p_{\text{accept}} = \min\Big(1, \frac{p(x_j)U(x_i|x_j)}{p(x_i)U(x_j|x_i)}\Big). \tag{14}$$

Taking it all back to the statistics of our system we specifically in the above, want to select the acceptance probability to satisfy

$$\frac{A(x_j, x_i)}{A(x_i, x_j)} = \frac{P(x_i)}{P(x_j)} = \frac{\frac{1}{Z}e^{-\beta H_{x_i}}}{\frac{1}{Z}e^{-\beta H_{x_j}}} = e^{-\beta(H_{x_i} - H_{x_j})}. \tag{15}$$

for x being a configuration in the lattice, Z the partition function and $\beta = \frac{1}{T}$. Resulting in the criterion in eq. 14 that if $H_{x_i} > H_{x_j}$ it means that $A(x_j, x_i) > A(x_i, x_j)$.

$$A(s_i \to s_i^{'}) = \frac{P(s_i^{'})}{P(s_i)} = \begin{cases} exp(-\beta\Delta E) & \text{if } \Delta E < 0 \\ 1 & \text{otherwise} \end{cases}$$

2.5 Kawasaki dynamics

There exist several types of spin-flip dynamic algorithms, commonly differentiated between single spin-flip and cluster flip algorithms[8]. Whereas the Metropolis algorithm is the most commonly used, due to its property of evolving the system magnetically. We can also use Kawasaki dynamics for its property of keeping the magnetisation of the system constant. Mathematically we can follow much of the same process as for the Metropolis algorithm, but the update scheme changes. Each simulation step, we choose a random spin site, s_i and may swap its proposed next spin state, s_i' with one of its neighbours present and proposed states as $s_i' = s_j$ and $s_j' = s_i$ with probability $A(s_i, s_j \to s_i^{'}, s_j^{'})$

$$A(s_i, s_j \to s_i^{'}, s_j^{'}) = \frac{P(S^{'})}{P(S)} = \begin{cases} exp(-\beta\Delta E) & \text{if } \Delta E < 0 \\ 1 & \text{otherwise} \end{cases}$$

[5]The numbers are rounded off and no matter the initial points, it converged to the given values.
[6]The random walk is aperiodic
[7]the random walk is irreducible
[8]We want to stay in the domain of single spin-flip algorithms as using cluster flip removed the "shoulder".

This then ensures that the magnetisation remains constant, that is the fraction of spin up and spin down remains the same. We can therefore see the process as moving the spins spatially where clustering still occur. Depending on the initial system and how we evolve the system[9] we expect that if we run a system long enough that the clusters will segregate into respectively spin up and spin down clusters

2.6 Quantifying information

In essence, the Ising model is a system that can be represented as a binary, i.e. 1 and 0, either we have spin up or spin down. Using Information theory we can, given a sequence of data and their probability arithmetic, find an optimal way to represent the data. In this paper, we use the modelling behaviour of LZ77, also known as Lempel-Ziv compression.

2.6.1 Lempel-Ziv

In using Lempel-Ziv compression we can ask the question of how many distinct phrases can a string with length n be parsed into. The length of a bit string, composed of 0 or 1, can take on $n_k = \sum_{j=1}^{k} j2^j$, where it contains 2^j phrases of length j for which we could show an example of a string with length. LZ77 iterates sequentially through the input string and stores any new match into a search buffer. Given the example string of [001010], we start from the left with an empty search buffer. We have and offset, o, a length, l and a character, c. So the LZ77 encoding will look like (o, l, c), for the first digit we have [[0]01010], as $(o = 0, l = 0, c = 0)$. Since we are not moving backwards, the offset is zero. As the buffer is currently empty we start at $l = 0$ and the character is a zero. likewise for the next character, we now move $l + 1$, as the pattern already exist we can write $(o = 1, l = 1, c = 0)$. For the next one, since no existing pattern exist in the buffer we still have $l = 0$, we therefore have $(o = 0, l = 0, c = 1)$ for [00[1]010]. We can then easily imagine that we'll find patterns such as [01], [001] or larger[13, chap. 1]. As the LZ77 algorithm require a 1D sequence and the dimension of our data is larger than 1D, the data will have to be reordered down into a 1D array. This is done using a scanning approach similar to the work found in the paper by Andrea C. et al [2], but here we utilise a Hilbert curve to generate the sequence The scanning is composed of two distinct parts, a Serialized Time Coding (STC), that scans a 2D Hilbert curve and an Interlaced Time Coding (ITC), which scans a 3D Hilbert curve. For the STC a 2D configuration matrix is scanned for each time step using a 2D hilbert curve that is then concatenated into an array of size $A(T, L, L)$. For the ITC, the entire space-time array is scanned with regard to a 3D Hilbert curve, into an array of size $A(T, m, m)$. These two different ways of treating the data enables ITC to conserve time correlations, while STC does not.[2]

2.6.2 Computable Information Density (CID)

If we wish to quantify the information of our system we can use Computable Information Density (CID), which is a measure of the ratio of the length of losslessly compressed data sequence to that of the uncompressed length of the original data sequence

$$\texttt{CID} = \frac{\mathcal{L}(x)}{N}. \tag{16}$$

Here $\mathcal{L}(x)$ denotes the total binary data length of a compressed sequence and N is the length of the original sequence of x.[2]

[9] This is discussed in the section on methods, where the choice of initial system is important.

2.6.3 Q measure

While CID is system dependent and can tell us about the properties of a first and second order phase transition [10], we cannot necessarily always find the inherent correlation of a system. The Q measure is instead independent of the analysed system by sampling the system at various length scales. The CID is here compared to its uncorrelated counterpart, where all the degrees of freedom have been subject to random shuffling[2]

$$Q = 1 - \frac{CID}{CID_{\text{shuffled}}}. \tag{17}$$

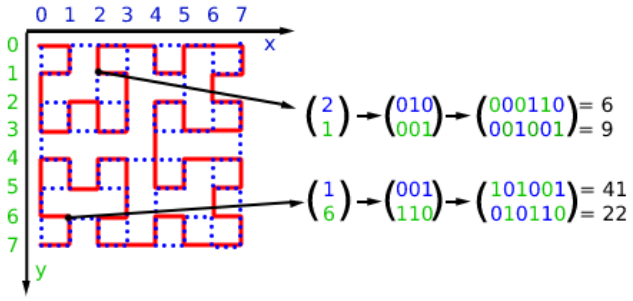


Figure 6: Two Hilbert curves overlaid on top of each other, a solid red line and a mirrored dotted blue. The distance between two points on the Hilbert curve depends on the order, as for this example we have a grid length of size $2^3 = 8$, thus an order of 3. For the given grid we have the distance as $41-6 = 35$ and $22-9 = 13$ and a total distance of $35 + 13 = 48$

3 Method and Result

In executing the model, the algorithm is composed of two parts. Firstly there is the Ising model itself for which the available variables when running the model are, L corresponding to the number of spin sites in one direction. The number of temporal measurements of $nsteps$, from here on is denoted as `n`, as well as the temperature range and temperature step size. In this part, all the regular properties of the Ising model are found and we reduce the output to the average quantity per number of spin sites in the following sections. Secondly, the Computable Information Density analysis runs separately after the Ising model has been simulated and find both the CID and Q measure. Here we have the order, p given by L. The number of shuffling $nshuff$, for which a default value of 8 was used initially but dropped to $nshuff = 2$ as the computational resource required proved too significant. The number of shuffling done did not appear to affect the results in an apparent or meaningful manner with regards to the results and thus kept at $nshuff = 2$. The last variables are then the number of `n` and what to use as the base decimation of `n`.
The results of this paper are from the initialisation of the Ising model from the lower temperature region, although it can also be reversed to start from a high-temperature region. Initial trials from the high-temperature region did not yield different results overall. An initial temperature of $T < 0.5$ consistently created lattice structures purely in a single spin direction, making sure to start from a completely ordered state, this is important when using Simulated annealing outlined in section 3.1. A thermalisation of 1024 was determined to be sufficient to reach a stable state and used for all cases. Re-scaling of the temperature axis has been done according to ordinary phase transition analysis of the reduced temperature $\tau = (T - T_c)/T_c$. The Temperature range ran from $T = 0.3, 3$ with step size of $dT = 0.05$.

3.1 Simulated annealing

The Ising model is a combinatorial optimisation problem and we want to find a configuration that can minimise the Hamiltonian energy function $H(s)$. From section 2.4.1 on MCMC, we know that the configuration space has a total number of possible configurations as 2^N. The system complexity in other words increases exponentially with the number of sites and it is thus immensely difficult to find a ground state to solve the minimisation problem. Luckily, we can use Simulated Annealing (SA). SA uses the MCMC method described to generate configuration samples based on the Boltzmann distribution $P(E) \sim exp(-\beta E)$ while evolving the temperature slowly. This ensures that we do not get trapped in a local minimum and instead find the global minimum. The spins are updated in

sequence and when we have performed one sweep, it means that we have done a complete update overall spins[12, pg.549-555]. Specifically for our Ising system and the kth sweep, the change in energy caused by flipping can be summarised using equation 2 as

$$\Delta E_i^k = -\sum_{j=1}^{i-1} J_{i,j} s_j^k (s_i^k - s_i^{k-1}) - \sum_{j=i+1}^{N} J_{i,j} s_j^{k-1} (s_i^k - s_i^{k-1}). \tag{18}$$

We, therefore, have a probability, using equation 14 of accepting the new state s_i^k with probability

$$p_{\text{accept}} = \min\left(1, exp(-\beta_k \Delta E_i^k)\right). \tag{19}$$

In running the code, we show in fig. 7, that the output performs as expected and the usual properties of the Ising model are found. The phase transition occurs at the correct point and we clearly see a discontinuity for the specific heat capacity and susceptibility.

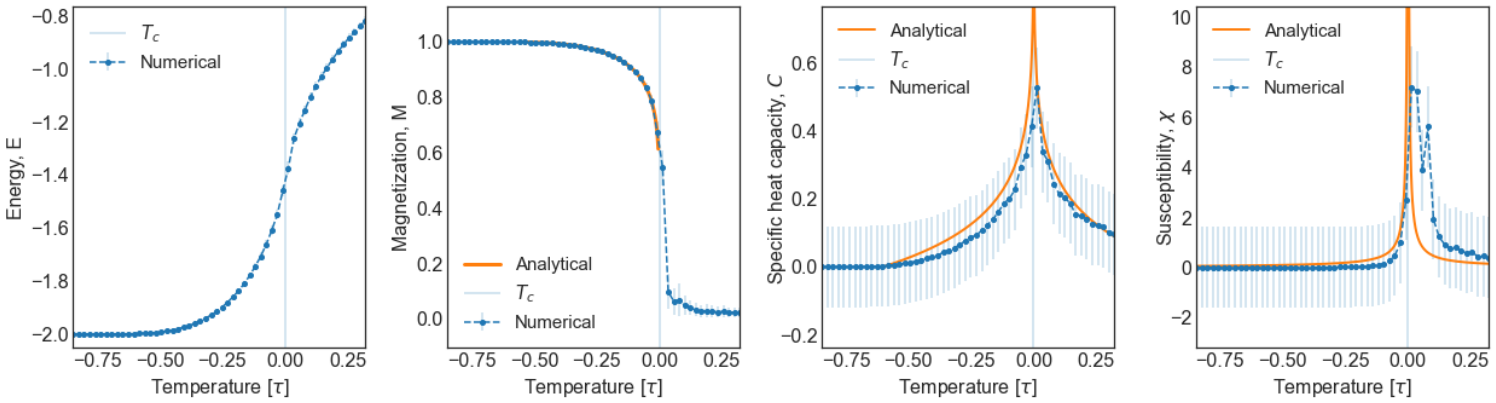


Figure 7: Properties of Metropolis Ising model second order phase transition for L128 and n4096. It is shown how a second order phase transition occur at the curie temperature point. The quantities are averages over spin sites. In the first plot we see how the average energy of the system increase. This increased energy results in the property of the second plot where the order of the system goes from an ordered state to a disordered state.

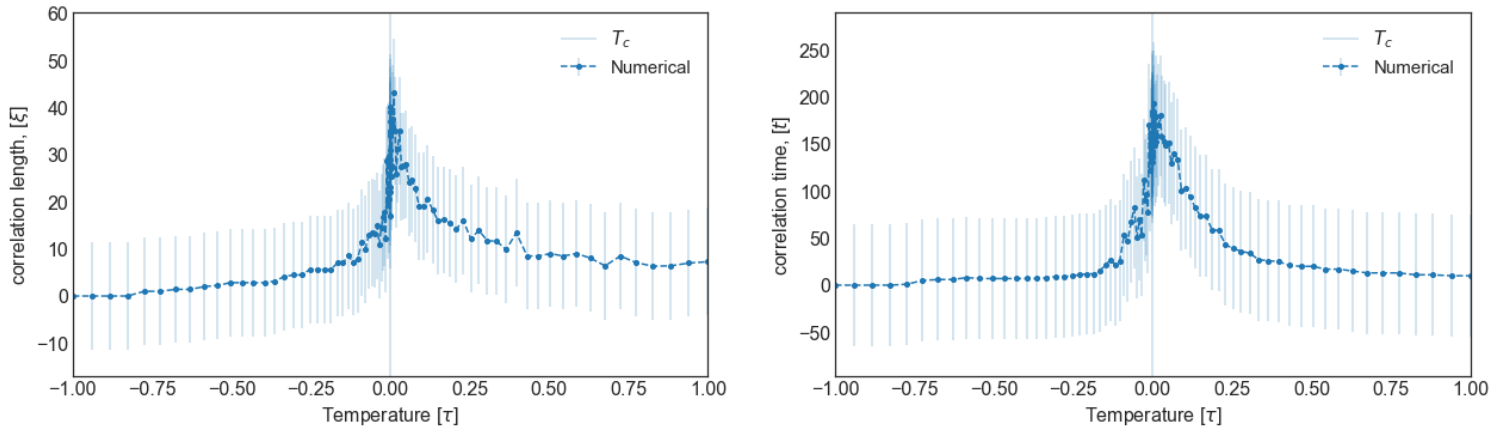


Figure 8: Both from the characteristic length and the characteristic time we see the phase transition occur as the discontinuity around zero. The characteristic time is how the temporal evolution of our microscopic variables given position, r, and time, t, influences the value of the same microscopic variable at a later time.

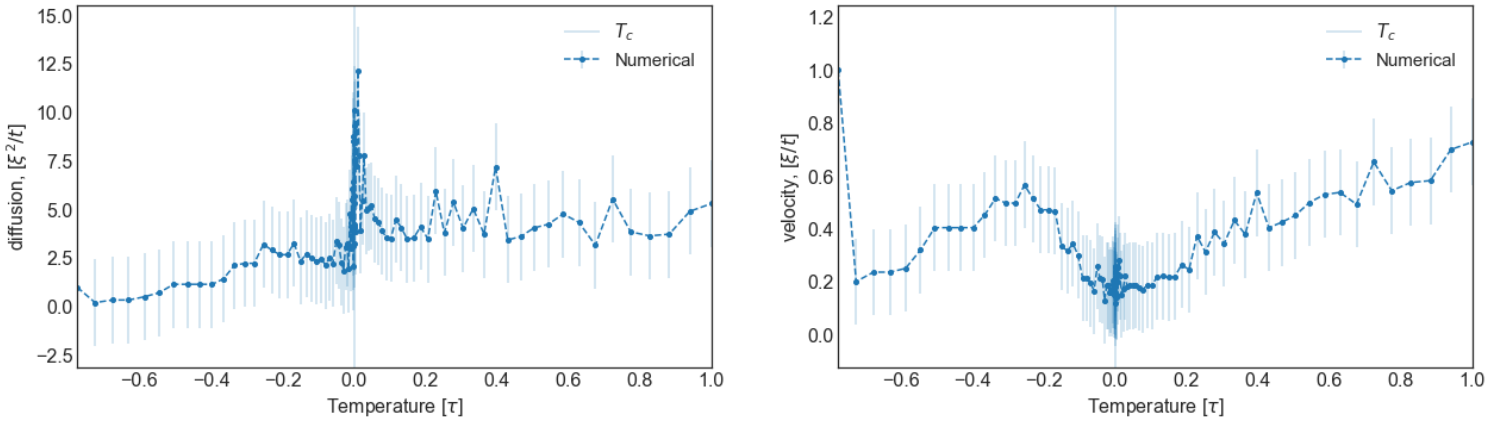


Figure 9: Shown is the velocity, given that we have a characteristic length and characteristic time and functions like ordinary velocity. How the correlation length changes over a given time displays a slow down around the phase transition on both sides before picking up speed again. Likewise we have a sort of characteristic area, signifying how a given area of characteristic length changes over time. This behaves as expected as when we approach the phase transition we also have a cluster inhabiting the width of the system and can therefore expect the area to be a maximum around the phase transition.

From figures 10 and 11, overall the data appears the same, but there is a slight movement to the left when we increase n. This slight movement is ultimately what we will investigate further.

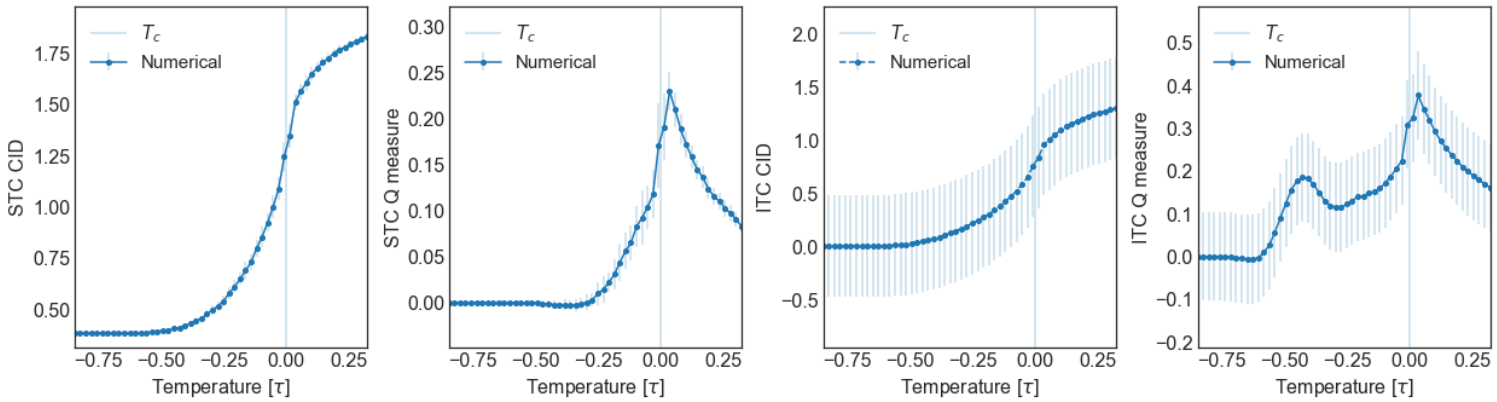


Figure 10: Quantified information of Metropolis-Hastings Ising model second order phase transition for L128 and n4096. The ITC displays a characteristic smooth relaxation

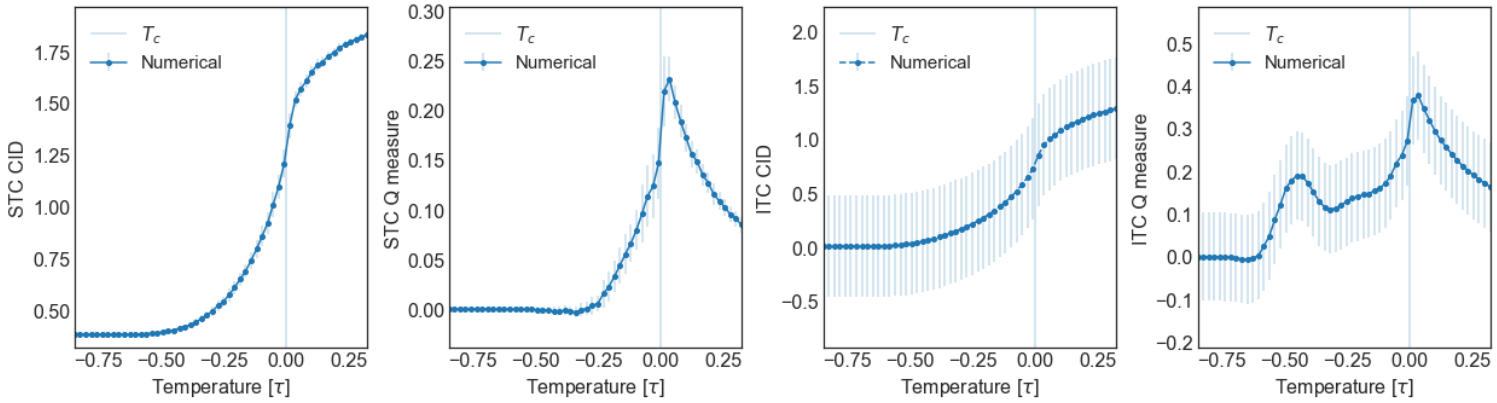


Figure 11: Quantified information of Metropolis-Hastings Ising model second order phase transition for L128 and n8192. An increase in n slowly starts moving the peak of the shoulder for ITC Q-measure to the left, a property seen from figure 16

3.2 Metropolis-Kawasaki hybridisation

The algorithm inherent to the Kawasaki dynamics is very sensitive to the initial conditions, as the fraction of spin directions remain constant. If we were to use simulated annealing for Kawasaki dynamics, we would find a state from which the system would never evolve. The first state would be ordered and it would have nowhere to go from there. Instead, we can take the last lattice configuration of the simulated annealing process that utilises the Metropolis-Hastings algorithm, as the input in our Kawasaki method. This ensures that each temperature has a point to make progress from and that it is initiated from a point of steady state[10]. The spin sites will only move spatially as described in a section 2.5, for our purposes, this will essentially isolate each temperature step from its original annealing approach and quench the high-temperature region after the phase transition.

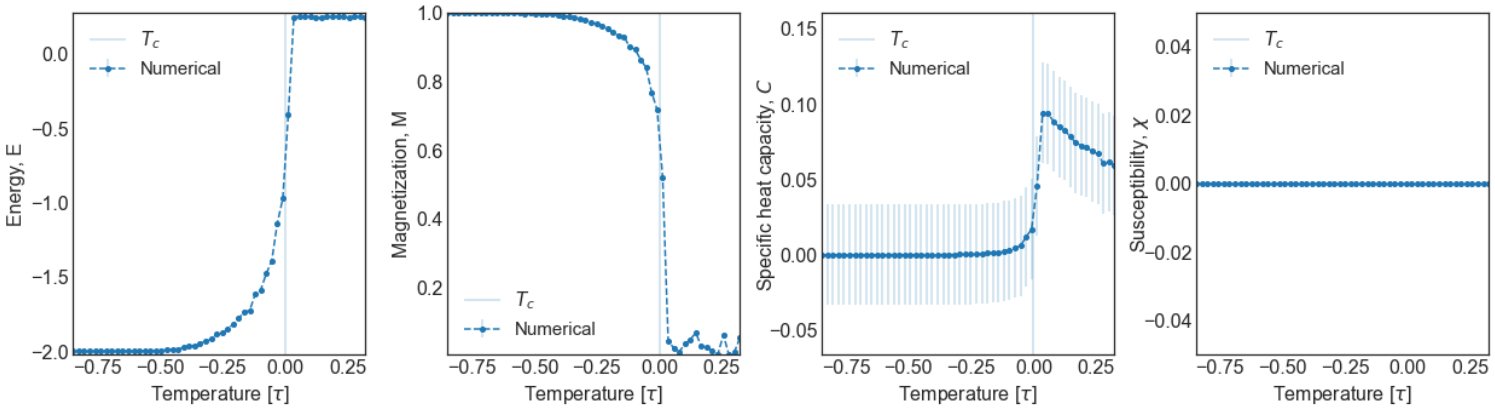


Figure 12: Properties of Metropolis-Kawasaki hybridisation method for Ising model second order phase transition for L128 and n4096. It is shown how a second order phase transition still occurs at the curie temperature point when compared to the pure metropolis-hastings method. The system has an expected constant susceptibility as a result of the kawasaki method.

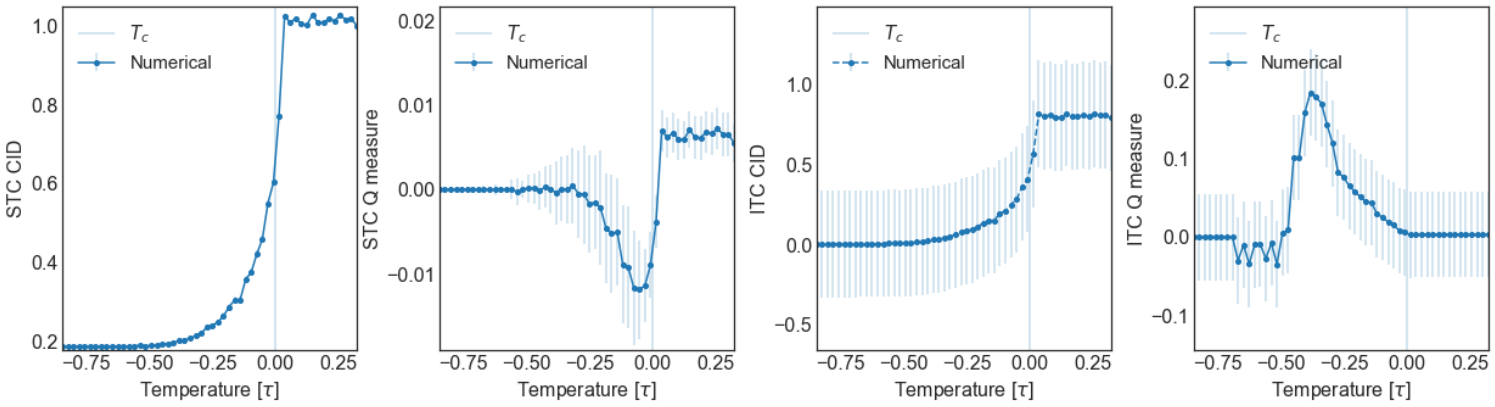


Figure 13: Quantified information of Metropolis-Kawasaki hybridisation for Ising model second order phase transition for L128 and n4096. The most notable feature is how the ordinary transition peak has been completely quenched and only the shoulder remain. The shoulder displays the same characteristics as with ordinary CID calculations for pure Metropolis-Hastings method. After the phase transition we see how the energy quickly reaches a plateau for its maximum energy.

[10]The code also run a new thermalisation process using the Kawasaki dynamics.

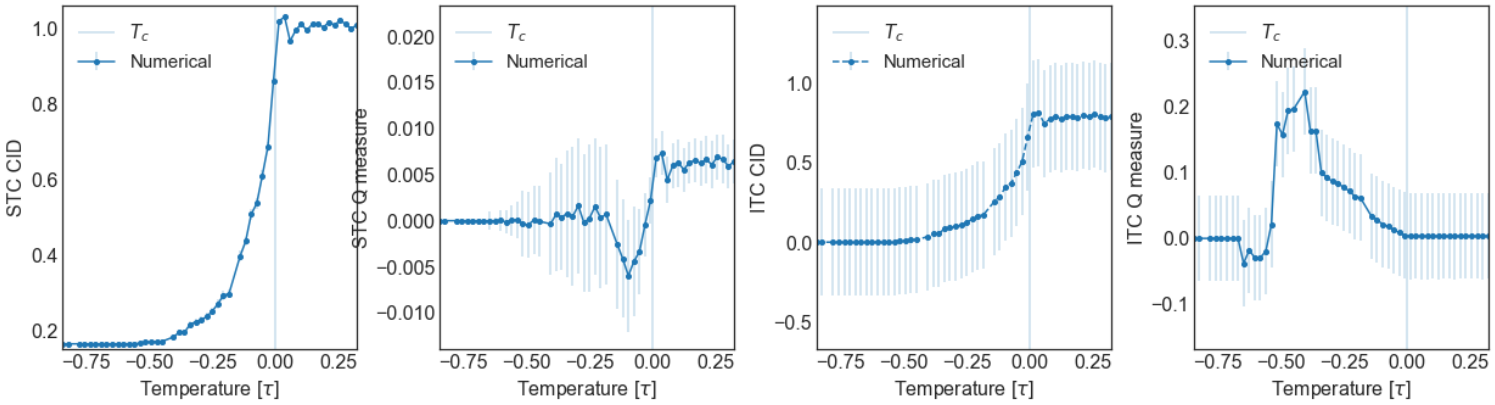


Figure 14: Quantified information of Metropolis-Kawasaki Ising model second order phase transition for L128 and n8192. Doubling nsteps have effectively halved the STC Q-measure, a feature not shared by the ordinary Metropolis-Hastings method. This could be a direct result of running the Kawasaki method for a longer period of time where the Kawasaki method "takes over" and further equalise compared to the pure metropolis.

3.3 Temporal decimation

With the inspiration of renormalization groups for Ising spins, we have used the approach of the decimation of the temporal n steps. Each subsequent evolution of the system acts like a windowpane from which we can remove e.g. every other frame. The decimation was done in skipped frame increments of $n4096$. This means that for $n = 4096$, no frames are skipped. For $n = 8192$ every other frame is skipped, etc. When temporal decimation is done we effectively move the system found via ITC towards that of STC, this can be seen by comparing figure 15 and figure 10. Both the observed shoulder and phase transition peaks are reduced, with enough decimation of a sufficiently large n, we can expect to converge.

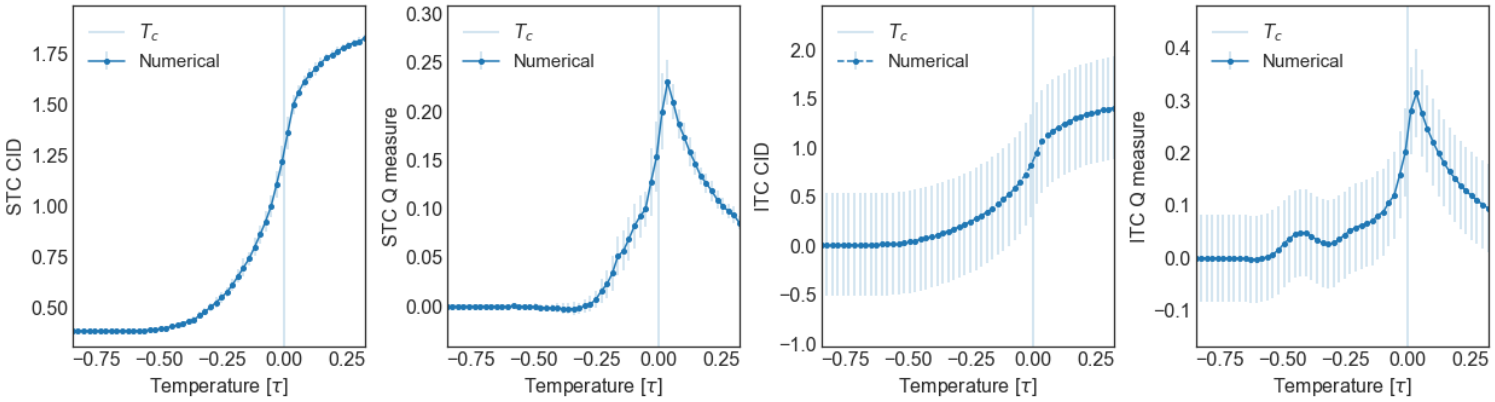


Figure 15: Decimation of CID of the Metropolis-Hastings Ising model second order phase transition for L128 and n16384. Decimation with base of $n4096$ effectively quench the ITC CID and Q-measure.

Executing the model several times over the range of $n = [4096, 8192, 12288, 16364, 20480]$[11] for both the non-decimated data and decimated data with a base of $n = 4096$, enables us to compare the behaviour of both the shoulder and its convergence of ITC towards STC.

[11]For a L128 system at increasing n, the system is computationally very demanding, while the Data is almost done, it could not be done for this plot and earlier data was used containing $L64$. The two system sizes displayed the same behaviour although for $L128$ the resolution was of cause better.

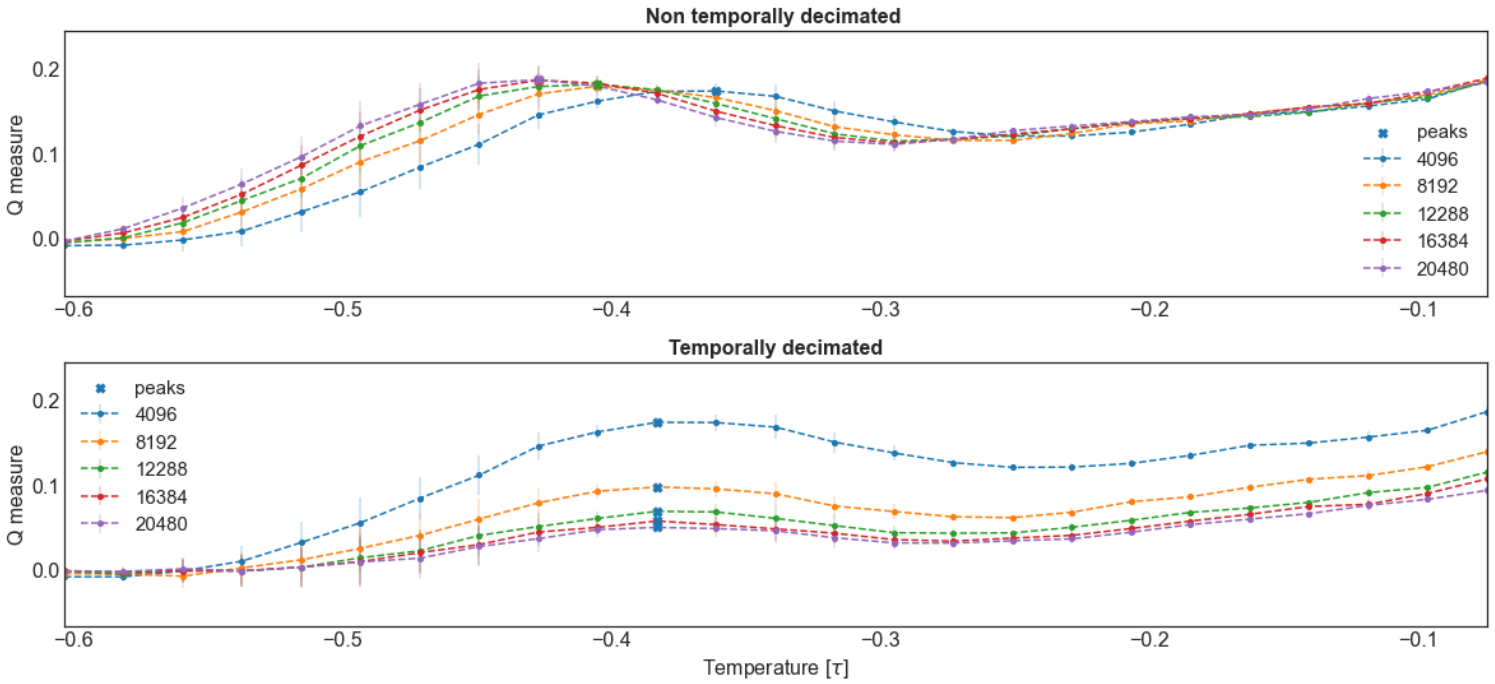


Figure 16: Quantified information of Metropolis-Hastings Ising model second order phase transition for L64
Top: ITC Q-measure plotted in range of n. From the figure we see how the shoulder moves to the left as n increase. For each increment in n, the pace for which the shoulder moves to the left decrease. Based on previous decimation runs $n4096$ was selected as a pseudo stable point from which to use as the basis for decimation.
Bottom: A base of $n4096$ was used for decimation, each increment in n yields diminishing returns, but we see that at $n16384$ and $n20480$ we reach a state where the values very much overlap and the uncertainty of the measurement makes it impossible to resolve without higher data resolution.

As seen on figure 16, the system displays a clear behaviour when we increase n, both in regards to its slow pace away from the transition peak on the x-axis and with decimation the diminishing shoulder.

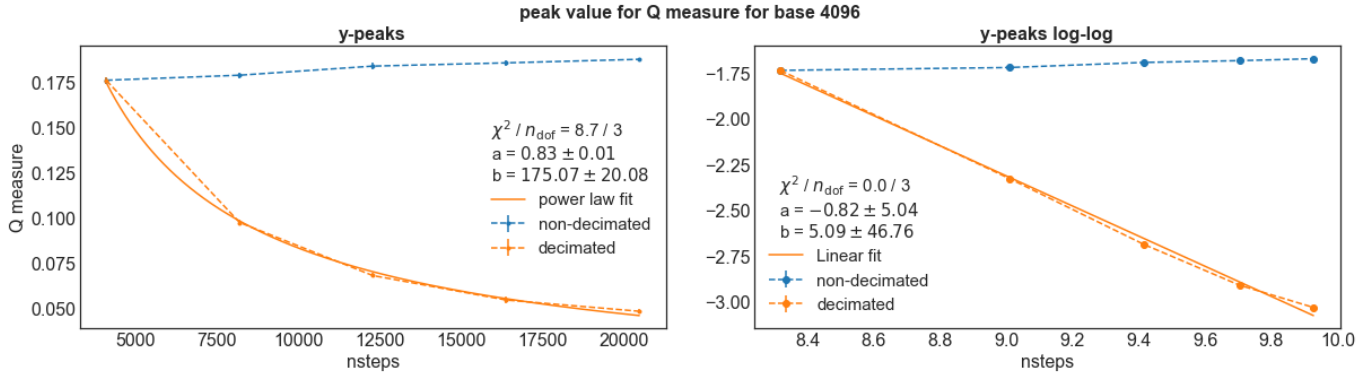


Figure 17: The peaks of decimated and non-decimated values of Q-measure for the shoulder in a range of n. Using iminuit to fit a power law of $b|x|^a$ via least squares we find a power of 0.83 4/5. Performing log-log operations further shows the clear linear correlation.

4 Discussion

It is well-established that properties such as magnetisation and specific heat of the 2D spin-$\frac{1}{2}$ Ising model uphold the expected second-order phase transition around the critical temperature T_c.

It is a transition phenomenon that our Computable Information Density analysis also captures. Both in form of the CID that acts like Energy and Q-measure which displays a likeness to the discontinuity displayed by either specific heat or susceptibility. But herein lies the promising part of the method, in that it reveals additional structures, evident by the appearance of additional phase behaviour around the "shoulder" to the left of the phase transition as seen on figure 10. These are promising results, however, the nature of the Q measure ensures that we know something is causing the results, but not necessarily what is causing it explicitly. The degrees of freedom of the system has been reduced due to the shuffling action of the Q-measure. We know only that the Ising model consistently experiences an increased disorder around this point related to the temporal dimension. As with the spatial indicators, this could be an indicator for a temporal action as the cause for spontaneous clustering away from the phase transition.

The shoulder moves away from the critical point of the spatial second-order phase transition with an increase in `n`, but we need a significant amount of `n` to even start seeing it. Even at a reasonable value of `n`64 it is not a distinguished enough shape, hence we have to move to `n`512 to even start seeing a proper shape. Going beyond this for even higher values of `n` we continuously see the shoulder move further away from the phase transition but at an increasingly slowed rate and start reaching a stable range between `n`4096 and `n`16384.

An alternate route we did not investigate was that of increasing the system dimension, L, where we did initially see similar behaviour unfold for lower `n`, while increasing L. This however uses significantly more system resources. Concerning compute time, an increase in L corresponds to an increase in the Ising simulation time of $O(n^2)$, but an increase in the CID calculation of $O(n^3)$. This is due to the temporal slices filling a 3D dimensional Hilbert curve. Whereas keeping L fixed and increasing n only increase CID calculation time by adding additional slices in a linear form.

In performing the temporal decimation for a pseudo stable value of `n`4096, most of the movement dissipates and the peak values line up for the used resolution, all shown in figure 16. However, the resolution of the data makes it difficult to differentiate between the later n increments as the uncertainty takes over. Further, an increase in resolution would enable us to calculate a numerical gradient with stable gradient peak values that would allow us to see deeper into what is happening around the shoulder. i.e the resolution of the data ensures that the fluctuations are too severe to achieve a proper gradient. This could however be alleviated by increasing the samples of the temperature range, it would allow for a similar analysis of the gradient peaks that could hold further information.

Collecting the peak values on figure 16 gives the results of figure 17. Here we fit all the peak values as a function of n against a function of $\frac{b}{x^a}$, which yields the results of $a = 0.83 \approx \frac{4}{5}$ and some constant. Plotting the results on a log-log scale reveals a linear trend, indicating the existence of a power-law relation
These results indicate that there is an emergence of organisation on long enough time scales that are independent from spatial ordering.
A speculation on why these results have been eluded so far could be that performing additional n in ordinary Ising simulations seem superfluous when the system has reached an equilibrium state. Further with the introduction of the used compression based entropy technique, the separation of the temporal part is more pronounced with an increase in n, but even so n has to be increased

significantly to observe it.
For this project, further steps could be taken in the direction of the Kawasaki hybridization approach to gauge how the system acts for lower n. Since the regular phase transition is quenched and the system remains constant due to the Kawasaki approach, interesting cluster behaviour could be observed.

As we are using a method involving statistical compressors[12] it is pertinent to address Kolmogorov Complexity[5], [9]. What this means for us is that for our algorithmic approach, we can never know if a file can be compressed or not and if it can be compressed, we may never know if it can be further compressed, i.e. whether we have reached the lowest possible compressed entropy. This will add to the uncertainty of the results, the higher the entropy of the source, the less it can be compressed and ultimately after the phase transition occurs, we get further towards the upper bound of the transition to complete disorder and compression becomes impossible.

5 Conclusion

With the introduction of compression based entropy, we find the so-called computable Information density and its relative metric of Q-measure. Using this method we have reduced the 2D Ising spin-$\frac{1}{2}$ model to its base degrees of freedom, revealing new unseen phase behaviour, where a collection of spins cluster together prior to the well-known phase transition. A sort of pseudo critical exponent that depends on the long scale temporal range of the 2D spin-$\frac{1}{2}$ Ising model is indicated to exist and we find a reasonable value of said exponent to be close to $\frac{4}{5}$. These results are centred around a "shoulder" that moves away from the central phase transition as more temporal steps are taken. Using the decimation concept from coarse-graining on a temporal scale instead of spatial yields promising results.

[12] Also known as Statistical data compression or from signal processing simply as data compression or bit-rate reduction. The LZ77 compression is one such statistical compressor.

6 Appendix

6.1 Quantum mechanics and the Ising model

The Ising model is a great toy model to understand the physics behind ferromagnetism and going beyond the large scale simulations, we take a look at the physics behind the toy model.
In its simplicity, the model investigates the quantum mechanical property of the spin of electrons and the simple rules for how they interact. Such a spin measured along any axis has a value of $\tilde{s} = \pm\hbar/2$ depending on its spin direction[13].
Let us start by considering the electrons themselves, they repel each other electrostatically as they have the same charge and the Pauli exclusion principle states that no two electrons can be in the same quantum mechanical state [8]. Generally, we can have either aligned or anti-aligned spin pairs, where when anti-aligned they are close together, but in different quantum mechanical states due to the Pauli exclusion principle and experience an electrostatic repulsion. When aligned they never get close to each other as it would violate the Pauli exclusion principle and the electrostatic repulsion is weak.
An electrons spin is related to its magnetic moment as two electrons close together to create the electrostatic force between them and will take a preference to line up anti-parallel. From this follows that we can also expect the two election spins would also be anti-aligned. It is, however, energetically favourable to be in a parallel spin state and the difference in energy is mostly electrostatic. Considering a regular lattice of electron spins, these will tend to align in the same direction. Where the combined magnetic moment of each electron adds up to a large net magnetic moment constituting the ferromagnetic properties which are simulated using the Ising model.
From the formulation of the many-body hamiltonian we can consider two electron spins interacting and can construct their coupled representation as the Heisenberg hamiltonian from $H = h_1 + h_2 + V_{12}$[8],[6]. From the Pauli principle, we know that the energy is affected by spin, even when the hamiltonian does not explicitly depend on it. Here represented respectively by the singlet and triplet state we can find their energies using Dirac notation[14] as

$$< S = 0|H|S = 0 > = E_s \tag{20}$$

$$< S = 1|H|S = 1 > = E_t. \tag{21}$$

The difference in energies are then $E_s - E_t$. The difference between single and triplet states can be parametrized as $s_1 \tilde{\cdot} \tilde{s_2}$ and we further consider the following operator

$$\sum{}_{12} = \tilde{s_1} \cdot \tilde{s_2} = \frac{1}{2}S^2 - \frac{3}{4} \tag{22}$$

Using the same dirac notation we construct the equation which is diagonal on the coupled basis with the following eigenvalues

$$< S = 0|\sum_{12}|S = 0 > = -\frac{3}{4} \tag{23}$$

$$< S = 1|\sum_{12}|S = 1 > = \frac{1}{4} \tag{24}$$

The Hamiltonian can be written in the form of an effective Hamiltonian $H = \frac{1}{4}(E_s+3E_t)-(E_s-E_t)_{12}$. Consisting of two terms, the first term is a constant term and the second is the spin-dependent term. Considering only the spin dependent part we write the spin dependent hamiltonian as $H_{spin} = -(E_s - E_t)_{12}$ and finding the integrals using dirac as

$$< S = 0|H_{spin}|S = 0 > - < S = 1|H_{spin}|S = 1 >= E_s - E_t \tag{25}$$

[13] noting that $2\pi\hbar$ is Plancks constant.
[14] The full integral form can be found in [3, Sec. 1]

from which we define the exchange constant as $\tilde{J} = \frac{E_s - E_t}{2}$. The spin dependent term in the model Hamiltonian can be written as

$$H_{spin} = -2\tilde{J}\tilde{s_1} \cdot \tilde{s_2}. \tag{26}$$

Here $J > 0$ indicates that $E_s > E_t$ and the triplet state $\tilde{s} = 1$ is favoured corresponding to the ferromagnetic case. Likewise for $J < 0$, means that $E_s < E_t$ and the singlet state, $\tilde{s} = 0$ is favoured. For many interacting electrons in a grid we therefore have

$$H = -\sum_{i,j} \tilde{J}_{ij}\tilde{s_i}\tilde{s_j}. \tag{27}$$

In the interest of simplicity, we make a reduction of the physical dimensions and the units of spin used in the section 2, are related as $\tilde{s} = \pm\frac{\hbar}{2}s$ and it is therefore implied that $s = \pm 1$. Likewise for the exchange constant $\tilde{J} = \tilde{J}(\frac{\hbar}{2})^2 s^2 = Js^2$. The total energy due to electron interactions can then be written as

$$E = \sum_{i=1} E_i = -\frac{J}{2} \sum_{j=i\pm 1} s_i s_j. \tag{28}$$

Here the factor 1/2 is due to the double sum where we do not count each pair of neighbouring electrons twice. The energy of the lattice depends on whether spins are mostly aligned or random. If spins are random the resulting energy will be $E = 0$ and if spins are aligned $E = -2JN$ and in general we are interested in calculating the average energy per site, E/N.

6.2 Computing resources

Executing the model described in the Methods and Results section was done using a HPC system with the following specifications for a node of 2x 24-core Xeon 6248R @ 3.0GHz CPU and 192 GB with 4 GB / core DDR4-2933 MHz RAM. All the available system resources of a given node was used and up to 20 nodes ran at a time for up to 5 days. The resulting compute time usage was model system dependent, where for the Ising code, increasing as $O(n^2)$[15].

References

[1] Stephen G Brush. History of the lenz-ising model. *Reviews of modern physics*, 39(4):883, 1967.

[2] Andrea Cavagna, Paul M Chaikin, Dov Levine, Stefano Martiniani, Andrea Puglisi, and Massimiliano Viale. Vicsek model by time-interlaced compression: A dynamical computable information density. *Physical Review E*, 103(6):062141, 2021.

[3] Paul M Chaikin, Tom C Lubensky, and Thomas A Witten. *Principles of condensed matter physics*, volume 10. Cambridge university press Cambridge, 1995.

[4] A Colliva, R Pellegrini, Alessandro Testori, and Michele Caselle. Ising-model description of long-range correlations in dna sequences. *Physical Review E*, 91(5):052703, 2015.

[5] Kolmogorov complexity. https://www.sciencedirect.com/topics/computer-science/kolmogorov-complexity. Last visited on 20/12/2021.

[6] Christopher J Foot et al. *Atomic physics*, volume 7. Oxford University Press, 2005.

[7] David J Griffiths. Introduction to electrodynamics fourth edition. 2021.

[15] going from a system of L=32 to L=64 doubles the area

[8] David J Griffiths and Darrell F Schroeter. *Introduction to quantum mechanics*. Cambridge University Press, 2018.

[9] "Incompressibility - Kolmogorov complexity" Marcus Hutter. *http://www.scholarpedia.org/article/Algorithmic_randomness#MartinL.C3.B6f_randomness*. Last visited on 20/12/2021.

[10] Stefano Martiniani, Yuval Lemberg, Paul M Chaikin, and Dov Levine. Correlation lengths in the language of computable information. *Physical review letters*, 125(17):170601, 2020.

[11] M Newman and G Barkema. Monte carlo methods in statistical physics chapter 1-4. *New York, USA*, 1999.

[12] Cambridge University Press. Numerical recipes, 1999.

[13] Tsachy Weissman. Lecture notes - lempel-ziv compression, January 2019. *https://web.stanford.edu/class/ee376a/files/EE376C_lecture_LZ.pdf* - Last visited on 12/01/2021.

[14] Eric W. Weisstein. "wiener-khinchin theorem." From MathWorld—A Wolfram Web Resource. Last visited on 24/01/2022.

[15] Julia M Yeomans. *Statistical mechanics of phase transitions*. Clarendon Press, 1992.