

Ising Model and the Metropolis Algorithm

Yutan Zhang

University of California, Davis, California, USA

(Dated: Oct/2020)

This article is mainly a summary of the simulation result of the 2D Ising model I obtained by the Metropolis algorithm. A summary of the history of the 2D Ising model is presented. The low and high temperature expansions of the model are extensively discussed. The Metropolis algorithm and its convergence are briefly discussed. In the end, the simulation results are cross compared with the expansions' results.

Keywords: Ising Model, Metropolis Algorithm, Monte-Carlo

I. ISING MODEL: BACKGROUND AND DEFINITION

The Ising model is a mathematical model of ferromagnetism in statistical mechanics. It is named after the physicist Ernst Ising, who first came up with this model and investigated its properties in one dimension.

In 1924, Ising solved the 1D Ising model in his PhD thesis and discovered that there is no phase transition in the 1D Ising model. That led Ising to abandon his model and to believe that his model would not be of any significance to the development of the statistical physics in any dimensions. However, that turned out to be incorrect as shown by Rudolf Peierls in 1936. Rudolf showed that the 2D Ising model would be ordered at very low temperature and disordered at very high temperature. So there has to be a phase transition in between. That means the Ising model would be of interest to people because it might help people understand the behavior of physical systems near critical points, be it the magnetic systems, alloy systems or even gas-liquid systems.

However, despite the growing interest in the model, people weren't able to find the analytical solution of 2D Ising model for a long time. Although the 1D Ising model is not very hard to solve using the transfer matrix approach. The 2D Ising model turned out to be very difficult to solve analytically. It was not until 1944 that Lars Onsager finally figured out the analytical solution of it.

Because of the mathematical difficulty faced by the analytical method, my project focus on numerical method in solving the 2D Ising model. In particular, I used the Metropolis-Hastings Algorithm, which will be introduced and explained in the coming sections. The particular model I investigated is defined like this:

The system consists of a infinite 2D square lattice. On each site of the lattice lives one spin which takes only one of the two possible values: 1 or -1. The interaction happens only between the neighboring spins through the same coupling constant J between any pairs of neighbours. The Hamiltonian of the system can thus

be written as:

$$H = - \sum_{\langle i,j \rangle} J \sigma_i \sigma_j$$

I focused specifically on the case where $J > 0$, which is also referred to as the *ferromagnetic* system.

II. HIGH AND LOW TEMPERATURE EXPANSION

However difficult it is to obtain a full analytical solution, it is not too hard to gain some insight to the physics of the model with some approximation.

Our ultimate goal while solving the model is to calculate the partition function or, equivalently, the free energy. We shall see below how this is achieved with series expansion at high and low temperature.

We know that the partition function of our model can be expressed like:

$$Z = \sum_{S_1} \dots \sum_{S_N} \prod_{\langle i,j \rangle} e^{\beta J S_i S_j} \quad (1)$$

To evaluate this expression, we confine ourselves to be at sufficiently high and low temperature, respectively.

A. High Temperature Expansion

At a sufficiently high temperature $T \rightarrow \infty$, $\beta \rightarrow 0$. Realizing that $S_i S_j = \pm 1$, we can make use of the identity below to rewrite the Boltzmann weight factor:

$$e^{\lambda x} = \cosh \lambda [1 + x \tanh \lambda] \quad (2)$$

This identity holds true as long as $x = \pm 1$.

With $\lambda = \beta J \rightarrow 0$ ($\tanh \lambda \rightarrow 0$ as well), the partition

function (1) can be rewritten and expanded as:

$$\begin{aligned} Z &= \sum_{S_1} \dots \sum_{S_N} \cosh^N \lambda \prod_{\langle i,j \rangle} \left(1 + S_i S_j \tanh \lambda \right) \\ &= \sum_{S_1} \dots \sum_{S_N} \cosh^N \lambda \left(1^N + 1^{N-1} S_1 S_2 \tanh \lambda + \dots \right. \\ &\quad \left. \dots 1^{N-1} S_{N-1} S_N \tanh \lambda + o(\tanh^2 \lambda) \right) \end{aligned} \quad (3)$$

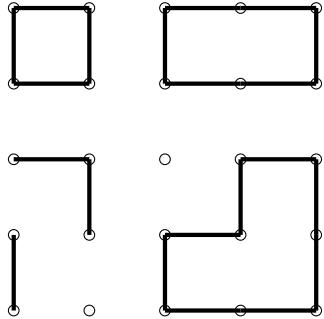
At high temperature we are only interested in the first few orders of $\tanh \lambda$.

When calculating the first few coefficients of this expansion, an interesting fact about this sum should draw our attention: We find that due to the sum over all spin variables, any term in the expansion with an odd power of S_i (i being any indices of the lattice) would sum up to zero. For example:

$$\begin{aligned} &\sum_{S_1} \dots \sum_{S_N} \cosh^N \lambda S_1 S_2 \tanh \lambda \\ &= \cosh^N \lambda \tanh \lambda \sum_{S_1} S_1 \sum_{S_2} S_2 \sum_{S_3} \dots \sum_{S_N} 1^{N-2} \\ &= 0 \end{aligned}$$

That tells us, only the terms whose powers of spin variables are all even numbers contributes.

A graphical representation below of this expansion can be used to help us make sense of this expansion:



in the lattice above, each connected piece represents one piece in the summation. The lattice in the graph represents the actual lattice on which the spins sit. Each stroke connect adjacent dots represents the term $S_i S_j \tanh \lambda$, where i, j are indices of the two connected dots. For example, the square at the bottom left corner

of the lattice:

$$\begin{array}{c} \dots \quad \dots \quad \dots \\ \circ \quad \circ \quad \circ \quad \dots \\ \square \quad \circ \quad \dots \\ \circ \quad \dots = S_{(1,1)}^2 S_{(2,1)}^2 S_{(1,2)}^2 S_{(2,2)}^2 \tanh^4 \lambda \end{array} \quad (4)$$

After summation over all S_i , all S_i^2 sums to 1 and this term just represents $\tanh^4 \lambda$. Similarly, any graph with an unclosed figure would be zero after the sum. However, it is noticeable that disconnected pieces (for example, two separated squares) does have contribution.

Another thing to notice is that because we can put the same square graph on different sites in the lattice, we have to count how many places we can put the square. Thus the algebraic calculation process becomes a business of counting figures.

The closed figure with the least number of points is square. So the lowest order expansion is to the 4th order of $\tanh \lambda$. We can theoretically do this expansion to an arbitrary order. Here we will show the expansion up to the 8th order:

$$\begin{aligned} Z = & \left(\square + \left(\square - \square + \square \right) \right. \\ & + \left(\square - \square + \square + \square \right. \\ & \quad \left. \left. + \square - \square + \square + \square \right) \right. \\ & + \left(\square - \square + \square \right. \\ & \quad \left. \left. + \square - \square \right) \right) \end{aligned} \quad (5)$$

For each graph, there is a corresponding coefficient, which is the number of that particular graph in the summation. Taken this into account, we can convert

the graph representation of Z to a algebraic expression:

$$Z = 2^N \cosh^N \beta J \left(1 + N(\tanh^4 \beta J) + \{N + N\}(\tanh^6 \beta J) + \left\{ \frac{N(N-5)}{2} + 4N + N + N + N \right\} \tanh^8 \beta J \right) \quad (6)$$

Using this expansion, we could write down any physical quantity of the system we want at high temperature. For example, we could look at the internal energy E :

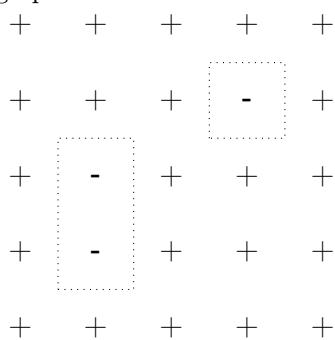
$$\begin{aligned} \langle E \rangle &= -\frac{\partial}{\partial \beta} \ln Z \\ &= -\frac{\partial}{\partial \beta} \left\{ N \ln 2 + N \ln \cosh \beta J + N \tanh^4 \beta J \right. \\ &\quad + 2N \tanh^6 \beta J + \left(\frac{N^2}{2} + \frac{9}{2}N \right) \tanh^8 \beta J \\ &\quad \left. - \frac{N^2}{2} \tanh^{10} \beta J + \dots \right\} \end{aligned} \quad (7)$$

We can see the coefficient correspond to $N^2 \tanh^8 \beta J$ cancels. In fact, the coefficients of $N^\alpha (\alpha \neq 1)$ of all orders of $\tanh \beta J$ should cancel. This is to ensure that energy E is an extensive property, which always has to be true. This provides a good check for your expansion. If calculated correctly, extensive properties would indeed be proportional to N .

B. Low temperature expansion

When the temperature T is low ($T \rightarrow 0$), we take another approach. Interestingly, this approach we will find eventually relates to the high temperature expansions.

We begin this expansion by disturbing the ground state. When at ground state ($T = 0$), the system (with $J > 0$) stays in a ferromagnetic state. When T gets a little bit higher, excited states would start to be filled. However, we know that the few lowest energy states are those with only a few spins flipped from the ground state. Again, we can understand this with the help with a graph:

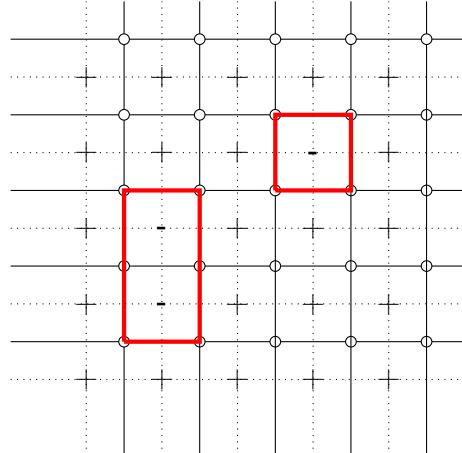


Ground state of the system is the state where all spins line up. While at low temperature, as shown in the

graph above, little Islands of opposite spins appear. The energy penalty for those little Islands is proportional to their circumferences (The length of the dashed line). Suppose the length of the dashed line is r :

$$\Delta E = 2J \times r \quad (8)$$

This reminds us of the high temperature expansion we just did. Actual we can make explicit connection between them by means of the dual lattice. That is, we bisector each bound of the original lattice to create its



dual:

The dual lattice in our case has the same structure as itself. And we can see that the circumference of our little island becomes the graphical representation of terms in the high temperature expansion.

Let's write out the partition function Z with this observation in mind. Latter we shall return to this point and construct a one on one correspondence between the high and low temperature expansion.

Considering that the total number of bonds is $2N$, the ground state energy would thus be $2NJ$. In low temperature expansions, the partition function would be:

$$\begin{aligned} Z &= \sum_r n(r) e^{(2N-2r)\beta J} \\ &= e^{2N\beta J} \sum_r n(r) e^{-2r\beta J} \end{aligned} \quad (9)$$

Here $n(r)$ represents the number of graphs with circumference r . Because of the duality of the lattice, each island in the low temperature expansion has a one on one correspondence to a graph in the high temperature expansion of its dual lattice - which is the same lattice system. So we realize that this $n(r)$ factor for (9) is exactly the same as the coefficients in (6). The expansions (9) and (6) thus have exactly the same structure, except that the variables to be expanded are different, as well as an overall coefficient outside the sum symbol.

C. Duality Transformation of the Partition Function

It turns out that we the low temperature expansion of Z at temperature T should equal to the high temperature expansion of it at T^* with the exception of an overall factor. If we let:

$$\sinh 2\beta J = \sinh 2\beta^* J \quad (10)$$

Then it follows that $e^{-2\beta J} = \tanh \beta^* J$. So we find the low temperature expansion of $Z = Z_L$ becomes (Z_H is the high temperature expansion of Z):

$$\begin{aligned} Z_L(\beta) &= e^{N\beta J} \sum_r n(r) e^{-2r\beta J} \\ &= e^{N\beta J} \sum_r n(r) \tanh^r \beta^* J \\ &= (\sinh 2\beta^* J)^{-N} \sum_r n(r) \tanh^r \beta^* J \\ &= (\sinh 2\beta^* J)^{-N} Z_H(\beta^*) \end{aligned} \quad (11)$$

We have now related Z at low temperature with itself at high temperature. Because the coefficient $(\sinh 2\beta^* J)^{-N}$ does not have singularity at finite temperature, any singularity Z has at $T = \frac{1}{\beta}$ corresponds to a singularity at $T^* = \frac{1}{\beta^*}$. If there could only be one singularity in finite temperature (as proven by the theorems of Yang and Lee(1952)), the singularity has to be at the point T_C where $T_C = T_C^*$. Thus it follows:

$$\sinh 2\beta J = \sinh 2\beta^* J \quad (12)$$

which gives us $T_C = 2.269$ (I have set the Boltzmann constant K_B to 1), the critical point of our model.

The expansions also give us good approximation of physical quantities at sufficiently high and low temperature, which we shall compare to our Monte Carlo simulation result.

III. METROPOLIS-HASTING ALGORITHM

The Metropolis-Hastings algorithm was brought up by Nicolas Metropolis in 1953 for symmetrical proposal functions and was extended by W.K.Hasting in 1970 to a more general case. We will define this algorithm and provide a short proof of convergence in our case of use.

The Metropolis-Hastings algorithm is a way of drawing samples from a probability distribution $P(\vec{X})$ as long as we know a function $f(\vec{X})$ that is proportional to $P(\vec{X})$. This method is extremely useful when \vec{X} is very high dimensional where the normalization of $f(\vec{X})$ becomes hopelessly hard. However, implementing this method requires that $f(\vec{X})$ is calculable at each point \vec{X} .

A. The Metropolis Algorithm

We'll be using only symmetrical proposal functions, so the algorithm we used is just the original algorithm brought up by Metropolis.

The Metropolis consists of the following steps:

Algorithm 1: Metropolis Algorithm

Result: Sample of system states (a chain of states) generated proportional to $P(\vec{X})$ by some function $f(\vec{X}) \propto P(\vec{X})$

initialization:

1. Choose an arbitrary starting state \vec{X}
 2. Choose a proposal function $g(\vec{X}|\vec{Y})$ from which we draw the next move \vec{Y} , given the current state \vec{X} . This function has to be symmetric in the sense that $g(\vec{X}|\vec{Y}) = g(\vec{Y}|\vec{X})$.
- for** n steps **do**
1. Suggest the next move(state) \vec{X}'_{t+1} by drawing from the distribution $g(\vec{X}_t|\vec{X}'_{t+1})$
 2. Calculate the ratio $r = \frac{f(\vec{X}'_{t+1})}{f(\vec{X}_t)}$
 3. Draw a random number u from the interval $[0, 1]$
- if** $r \geq u$ **then**
- | Accept the new proposed state \vec{X}_{t+1} and set $\vec{X}_{t+1} = \vec{X}'_{t+1}$;
- else**
- | Decline the new state and set $\vec{X}_{t+1} = \vec{X}_t$;
- end**
- end**
-

In our case we can set $f(\vec{X}) = e^{-\beta E(\vec{X})}$

B. Convergence of Metropolis Algorithm

In our model, we could prove that the Metropolis Algorithm should converge after sufficient numbers of iteration.

The process of generating samples in the Metropolis Algorithm is the process of generating a Markov Chain. We only have to prove that the transition matrix of the Markov Chain acting on any initial states for enough time give us the desired probability distribution of states.

If we represent the probability distribution of states with a column vector, then each step in the Metropolis

Algorithm is equivalent to multiplying a matrix T_{ij} to the current state.(The summation of repeated indices is indicated if not explicitly stated otherwise)

$$P'_i = T_{ij}P_j \quad (13)$$

And our T_{ij} can be written as

$$T_{ij} = \begin{cases} g(i,j)\alpha(i,j) & \text{if } i \neq j \\ g(i,i) + \sum_{k \neq i} g(i,j)(1 - \alpha(i,j)) & \text{if } i = j \end{cases} \quad (14)$$

Where $\alpha(i,j) = \min\left\{1, \frac{f(i)g(j,i)}{f(j)g(i,j)}\right\}$

This matrix T_{ij} is constructed exactly abiding to the law of the Metropolis Algorithm.

Our goal is just to prove that $\lim_{n \rightarrow \infty} T^n V = P$, where V is an arbitrary starting state and P is the actual distribution function of our physical system.

To prove it, we could use a few definitions and theorems:

Definition 1. If matrix T satisfies the following conditions:

1. $\sum_i T_{ij} = 1$
2. $\forall i, j T_{ij} > 0$

Then the matrix T is called a stochastic matrix.

Definition 2. If a stochastic matrix T and a vector P satisfies $T_{ij}P_j = T_{ji}P_i$ (here we do not sum over repeated indices), we say that state P is reversible. This is called the **detail balanced condition**.

There are a few theorems about stochastic matrices that we can use to prove the theory:

Theorem 1. If T is a stochastic matrix and P is reversible, P is the eigen vector of matrix T with eigen value 1.

Proof. Because

$$T_{ij}P_j = T_{ji}P_i$$

We simply sum over j on both sides:

$$\sum_j T_{ij}P_j = \sum_j T_{ji}P_i$$

Realizing that $\sum_i T_{ij} = 1$, we arrive at:

$$\sum_j T_{ij}P_j = P_i \iff TP = P$$

Theorem 2. Let λ_i be eigen values of stochastic matrix T . Then $|\lambda_i| \leq 1$ for $\forall i$.

Proof. Let's assume V to be one of left eigenvectors of T , such that $VT = \lambda V$. We can assume that the k th entry of V has the largest module in vector V .

We then have:

$$\begin{aligned} |\lambda||V_k| &= |\lambda V_k| \leq \left| \sum_j T_{jk}V_j \right| \leq \sum_j T_{jk}|V_j| \leq \sum_j T_{jk}|V_k| \\ &\implies |\lambda||V_k| \leq \sum_j T_{jk}|V_k| = |V_k| \\ &\implies |\lambda| \leq 1 \end{aligned}$$

Where the positivity of T_{ij} is used. \square

Theorem 3. The probability distribution $P(i)$ of the states is the reversible state that is the eigenvector of T of eigenvalue 1.

Proof. We can prove this theorem by calculation.

$$\begin{aligned} T_{ij}P(j) &= g(i,j) \frac{f(i)g(j,i)}{f(j)g(i,j)} P(j) = g(i,j) \frac{P(i)g(j,i)}{P(j)g(i,j)} P(j) \\ &= g(j,i)P(i) = T_{ji}P(i) \end{aligned}$$

That means T and P satisfies the detail balanced condition. Then from Theorem 1, we conclude P is the eigenvector of T with eigenvalue 1. \square

With the help of the above theorems, we can conclude that: The eigenvector of the stochastic matrix that represents the Metropolis Algorithm is the probability distribution that we wish to converge to. Furthermore, because the real probability distribution has the largest eigenvalue of the stochastic matrix, a repeated application of the matrix to any initial state should converge the state to our desired eigenstate. Thus we have proved the convergence of the Metropolis Algorithm.

IV. RESULT

I used the Metropolis Algorithm mentioned above for my Monte Carlo simulation. Specifically, I have set the Boltzmann Constant $k_B = 1$, the coupling constant $J = 1$.

I will present several results from the simulation in compare with the high/low temperature expansion to the first few orders in the high/low temperature region. I have ran the simulation on 8×8 , 12×12 and 16×16 lattices with periodic boundary condition. \square

We have calculated a few physical quantities of the system, all of them are divided by the size of the lattice so that they are size invariant: the energy per site E_{site} , the heat capacity per site C , the polarization per site S_{site} , the susceptibility per site χ_{site} and the Binder's ratio B . The simulation of all of those quantities on the three different sizes of lattices will be presented.

First let's see the result for the energy per site E_{site} . On the figure the high and low temperature expansions are also drawn. We can see that low temperature expansion fits pretty well at the low temperature part until the critical point where the energy shows a steep jump. The high energy expansion up to the eighth order fits well on the high temperature range and start to curve down near the critical point as well. The result does not show much difference with different size, which justifies our simulation on a relatively small lattice.

From the graph of the heat capacity (per site) C and the susceptibility (per site) χ_{site} , we should be able to find the critical temperature by locating their peaks. The peaks for C on the 8×8 , 12×12 and 16×16 lattices are at $T = 2.42$, $T = 2.35$ and $T = 2.34$ respectively. And the peaks for the susceptibility χ_{site} on corresponding lattices are at $T = 2.54$, $T = 2.43$ and $T = 2.40$. This is very close to the result of calculation with dual lattices ($T = 2.26$) in the previous section, considering the finite size of our system.

Another quantity we might look at is the polarization per site S_{site} . It shows that below the critical temperature, the system is at a ferromagnetic phase while above it the system arrives at a paramagnetic phase. The system exhibit the simultaneous symmetry broken at low

temperature away from the critical point. However, it begins to oscillate vigorously around the critical point. If we force the result to all point up, a more continuous phase transition shows up. In our simulation, if the simulations are very long, the system would tunnel between its two symmetrical polarization states and results in a zero polarization in average. The oscillation around the critical point in my opinion shows that the free energy gaps for the two simultaneous symmetry broken states are getting sufficiently smaller around the critical point, so it is much more likely to get tunneled through with the same steps simulated. The overly simplified picture of this could be seen similarly in the Landau's mean field theory, though which does not work properly in our problem(in 2 dimensions).

Finally we present the Binder's ratio $B = \frac{1}{2} \left[\frac{\langle M^4 \rangle}{\langle M^2 \rangle^2} - 1 \right]$, which is also another very important quantity to show phase transition. We can see from the three figures that B goes from 0 to 1 around $T = 2.5$, which agrees with our observation of C and χ that the phase transition happens around that temperature. From the Binder's ratio we do see that the transition becomes more and more fast as we increase the lattice size.

ACKNOWLEDGEMENTS

Thanks Professor Richard Scalettar for his lectures and collaboration.

-
- [1] S. G. Brush, *Reviews of modern physics* **39**, 883 (1967).
 - [2] Wikipedia contributors, “Ising model — Wikipedia, the free encyclopedia,” https://en.wikipedia.org/w/index.php?title=Ising_model&oldid=983711682 (2020), [Online; accessed 25-October-2020].
 - [3] “Metropolis-hastings and markov chains,” (2017), <https://ermongroup.github.io/cs323-notes/probabilistic/mh/>.
 - [4] P. P. (https://math.stackexchange.com/users/414530/pietro_paparella), “Proof that the largest eigenvalue of a stochastic matrix is 1,” Mathematics Stack Exchange, URL:<https://math.stackexchange.com/q/3174116> (version: 2020-05-17), <https://math.stackexchange.com/q/3174116>.
 - [5] “Eigenvalue of stochastic matrix,” (2020), <https://yutsumura.com/eigenvalues-of-a-stochastic-matrix-is-always-less-than-or-equal-to-1/>.

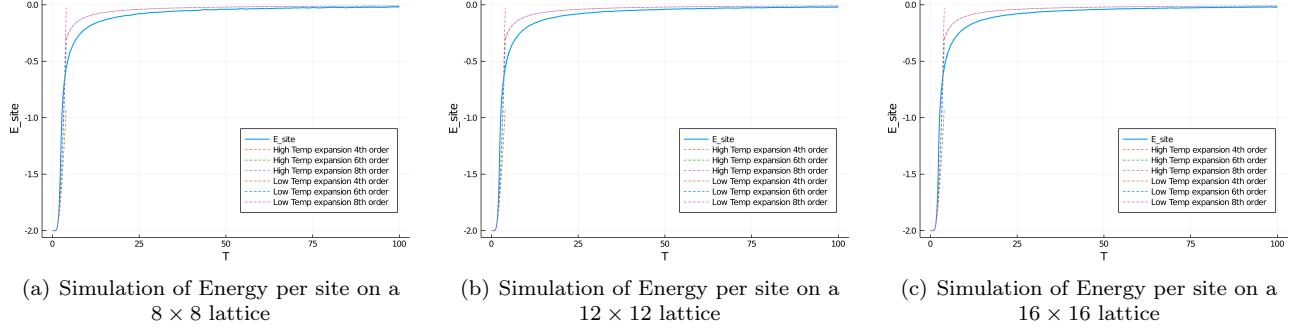


Figure 1. Simulations of energy per site E_{site} on lattices of different sizes. High temperature and low temperature expansion calculations are also shown on the graphs with dashed lines.

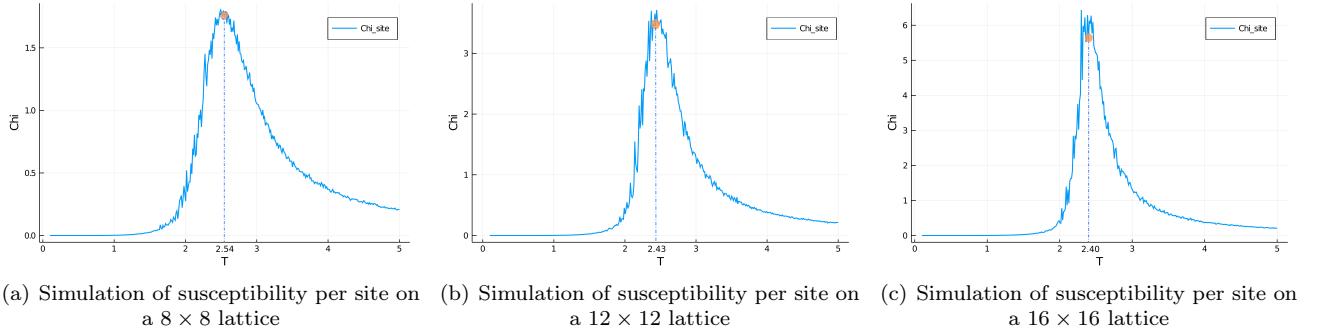


Figure 2. Simulations of susceptibility per site χ_{site} on lattices of different sizes. Peak positions are marked.

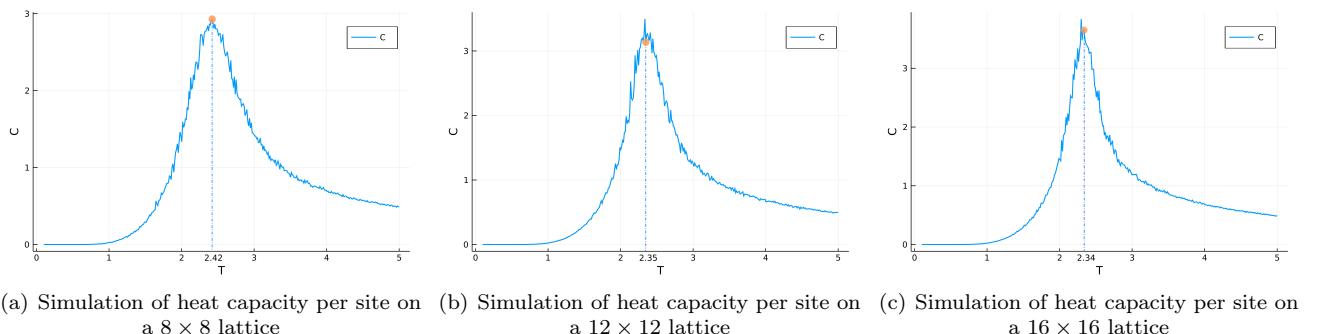


Figure 3. Simulations of heat capacity per site C_{site} on lattices of different sizes. Peak positions are marked.

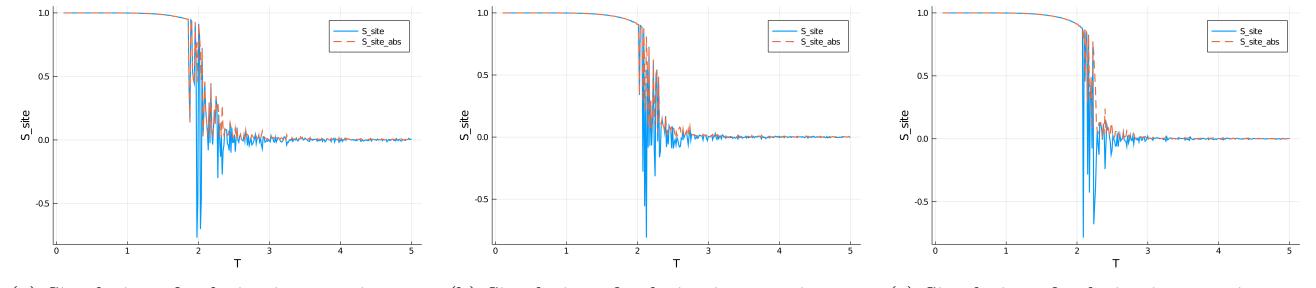


Figure 4. Simulations of polarization per site S_{site} on lattices of different sizes. Dashed line flips all down spins from our result to the symmetrical upper spins.

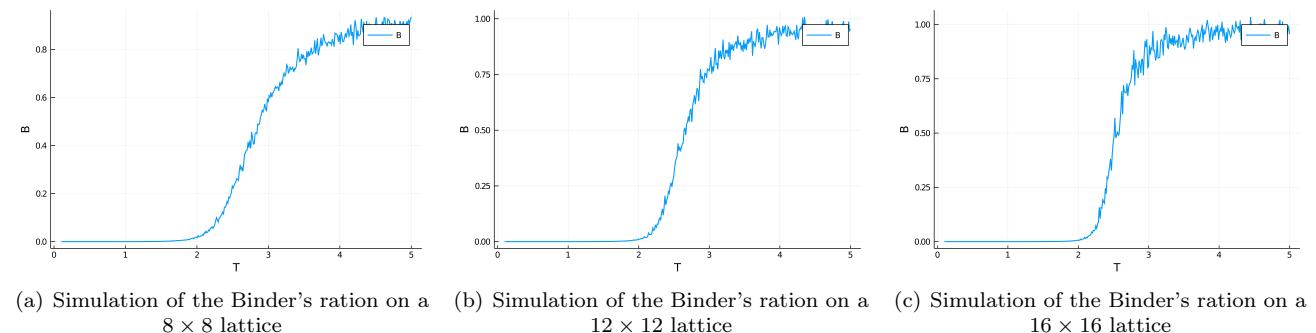


Figure 5. The Binder's ratio simulated on three different sizes of lattices