Cluster Monte Carlo for Ising model from 2D to 5D

Yijing Huang

March 8, 2017

Abstract

I used the cluster Monte Carlo Method for Ising spin systems, calculated the critical temperature for 2D, 3D, 4D, 5D, got the correlation function and tried to extract the critical exponent for correlation length, magnetization, susceptibility, and heat capacity. The codes and all the output files can be found at the following link, and will not be part of the appendix. https://github.com/2huangyj/cluster-monte-carlo-for-Ising-model

1 Introduction

Ising model is the simplest model that can be used to reproduce the behavior of much more complex macroscopic systems. In order to understand phase transitions, and the behavior of systems in higher dimension, Ising model will be a good start point. For 1D and 2D systems(Onsager 1944), there are exact solutions for the model, so they are also good tests for the applicability of numerical methods for investigating phase transition behaviors. The simplest Ising model is two-spin models, and it shows a spontaneous symmetry breaking below a critical temperature. The Hamiltonian of the systems is :

$$\beta H = J \sum_{<i,j>} s_i s_j + h \sum_i s_i + g$$

Here the s_i, s_j takes the value of either 1 or -1, representing the two difference spins,and i ,j are adjacent lattice spins. J is the dimensionless coupling constant, h is the magnetic field applied to the lattice system, and g is a constant energy. Only the coupling between the nearest neighbors are considered. In my calculation I assumed no magnetic field and set the free energy constant to zero, and without loss of generality, I set $J = 1$. So the free energy is simply:

$$\beta H = J \sum_{<i,j>} s_i s_j$$

One way to look at the problem is to calculate the partition function. Since there is not yet an analytical solution to the partition function and all the thermodynamic quantities, we can use computational approach. Iterating over the configurations is very computational expensive. I tried a very small lattice 5*5 in 2D Ising system using this brute force method and never got the result because it took so long to run. It is impractical to solve the simplest Ising model with this enumeration.

Monte Carlo simulation made it possible.The idea of the Monte Carlo simulation is to produce enough samples at equilibrium condition and average over them. Every step in a Monte Carlo simulation is a transfer from one state to the other, and the rate of transition is determined a transfer matrix. This reduces the computational task to sampling a few samples from the phase space to accurately estimate the physical quantities. We spare the effort to obtain the partition function itself. The downside of it is that apart from the fluctuation inherent to the thermodynamics quantities, we add more statistical fluctuations to the system, due to the very limited sample size.

In my calculation I only consider the thermal equilibrium states, and focused particularly on the system behavior near the phase transition temperature and the mean field theory tendency at higher dimensions. The critical exponents are important because they describe the general phase transition behavior regardless of the details of systems, which is the generality of universality class. Therefore I described the critical behavior with these critical exponents. They are the follows:

$$\xi \approx |t|^{-\nu}$$

$$M \approx |t|^{1/\delta}$$

$$\chi \approx |t|^{-\gamma}$$

$$C \approx |t|^{-\alpha}$$

In dimension lower than 3, the Ising model has a different behavior than the mean field theory solution, in dimensions higher than 4, the mean field theory is correct because each spin has more neighbors.

2 The Wolff algorithm

A valid Monte Carlo simulation must obey detailed balance condition as well as ergodicity. The former is entailed by the system's time reversal symmetry, which should be true for every system described by standard quantum mechanics or classical mechanics. If the detailed balance condition

is not true, then in equilibrium we can always find an evolution path from some state A to some state B, but when we reverse the time, we need to find another evolution path to bring state B back to state A, which contradicts the time reversal symmetry. The ergodicity means that there is a path in phase space from every state to every other state. If the samples are chosen in a way that certain states cannot be reached, the estimated thermodynamics observables will be significantly different from the real expectation. The easiest Monte Carlo method is the Metropolis algorithm, which flips the spin one by one. This works fairly well at the non critical temperatures but at T_c, but close to T_c this method suffers from critical slowdown, which is induced by both the critical fluctuation and the increase of the number of steps we need to go through to sample one independent configuration. The Wolff algorithm solve this problem because the it flips the whole cluster of spins, which approaches the equilibrium state much faster, and it produces independent samples fast enough for practical calculation even as correlation length approaches infinity near the critical temperature.

The justification for this algorithm is as follows: The detailed balance conditioncan be fulfilled when the transition rate form state μ to state ν is the same. If we choose one spin, and get the parallel spins around it to have the probability P of becoming part of the cluster, and apply the same procedure for the newly identified cluster spins until the cluster size no more grows, then flip the whole cluster, and define it as a transition from state μ to ν then the detail balance condition can be expressed as:

$$\frac{(1-P)^m A_\mu}{(1-P)^n A_\nu} = 1$$

where the m is the bonds to be broken from state μ to state ν and n is the number of bonds broken for the other way around. A_μ is the probability of the system to be at the state μ, and the ratio $\frac{A_\mu}{A_\nu}$ is obtained by the Boltzman weights, so there is the equation:

$$(1-P)^{m-n} = e^{-\beta(E_\nu - E_\mu)}$$

where it requires that

$$E_\nu - E_\mu = 2J(m-n)$$

here m and n add up to the whole length of the cluster boundary, and the energy difference between the two configurations is given by either the bonds connecting parallel spins changing into bonds connecting non parallel spins, or the inverse.So if we let

$$P = 1 - e^{-2\beta J}$$

the detailed balance condition will be satisfied.[NB99]

So the whole process of simulating towards equilibrium state in the real algorithm is:

1. generate a random configuration, select a spin at random, and add it into a queue
2. add into the queue its parallel neighbors with the probability P as defined above
3. take out the first element from the queue, and repeat step 2, and then 3
4. end the procedure when there is no more element in the queue and then return to 1 to generate the next equilibrium sample

The algorithm can be paraphrased in another way[Lui06], for every equilibrium sample, we divide the system into many clusters with the same rule as 2, and then choose one of the cluster, flip it, then wipe off the division and redo the process. We in principle allow any kind of cluster division, which is equivalent to saying that there is no forbidden state of this algorithm.

For the details of the implementation, I thought of using a stack before but after some application I found a queue may be quicker, because a stack in this case is a "depth prioritized search" and if the initial configuration is not favorable then the approach to the equilibrium state will take very long, the extreme case is when the initial configuration contains only one kind of spin and with a stack the cluster just expands along a line.

The initial configuration should be "random" for this algorithm to be fast. Since we always flip the whole cluster, it is easy for the system to order than to disorder.

3 Simulation

3.1 Correlation length

The correlation function in my calculation is between that of the two spins on one "lattice line", so that Fast Fourier Transform for each of the D dimensions combined with averaging over all the possible lattice line stretched between two boundary "surface", will be fast enough to provide a decent result.

The justification for this method can be described as follows:

$$\langle s_0 s_r \rangle = \int_{-\infty}^{+\infty} dr' s(r+r') s(r')$$

which is based on the approximation of an infinite translational invariant system. According to the convolution theorem, the Fourier Transform of the above equation on both sides yields:

$$F\{\langle s_0 s_r \rangle\} = |s(q)|^2$$

So do a FFT, and take the square modulus after that and FFT back to the real space will give us what we need. Although the above formulation is in the continuous space, it can be re-formulated in the discrete form, with the same assumption of translational invariance. This is faster than doing the summing over lattice and can be averaged over enough samples to give a more representative correlation function, which is better than just randomly picking two sites with given spacing in between.

The relation between correlation length and correlation function are as follows:

$$\langle s_0 s_r \rangle = e^{-\frac{r}{\xi}}$$

there is fundamental reason for why the exponential relation should stand. It is related to the fact that the correlation function in real space is a propagator. Correlation function of a certain kind of interaction, can obtain exponential behavior as long as under renormalization group this kind of interaction is irrelevant. The correlation length describes how big a cluster can contain correlated spins, and is a kind of gauge of the interaction scale of the system.

The above definition is the only way to calculate the correlation length, it can not be obtained from the cluster size. Because in fact, the cluster size at low temperature is very large(expand throughout the whole lattice) but the correlation length is very small. The cluster size has more to do with the magnetization and susceptibility, (although under the condition H=0), in fact at temperature above T_c there is the relation that $\chi = \beta\langle n \rangle$, [NB99] where $\langle n \rangle$ is the average size of cluster in terms of n the number of spins in contains. Correlation length is the property of interaction.

3.2 Heat capacity and susceptibility

In order to calculate the heat capacity and the susceptibility, we can always take the derivative of the partition function, which corresponds to the free energy. For example, the susceptibility can be expressed as

$$\chi = -\frac{1}{\beta}\frac{\partial^2 logZ}{\partial H^2}|_{H=0}$$

but the partition function obtained from the Monte Carlo method, due to the limited sample space, is going to suffer from a large fluctuation or noise, which is extremely unfavorable for doing derivative. But the advantage of Monte Carlo simulation is that we can spare the effort of calculating the partition function, and therefore we should apply the linear response theorem, the susceptibilities are the fluctuations of certain variables, so the ξ can be expressed as follows:

$$\chi = \beta N(\langle m^2 \rangle - \langle m \rangle^2)$$

and in the same we we can calculate the heat capacity.

$$C = \frac{\beta^2}{n}(\langle E^2 \rangle - \langle E \rangle^2)$$

We only need to average over samples instead of taking derivatives.

The calculation of the magnetization and the energy takes place every $\triangle t$ apart, which will be defined in the following section 2.5.I set one array (or matrix) to describe spin of every lattice sites, and another array to describe the connection property of every lattice bond, updated at every cluster flip: if it connects two non parallel spins, it is given the value 1, otherwise given the value -1. The magnetization is just summation of all the lattice sites, and the energy is just summation over all the lattice bonds.

At very low temperature, the clusters grow to the size of the system, and each cluster flip is essentially flipping the whole lattice, this will cause the magnetization to oscillate between -1 and +1 for every step of simulation, if I average it out, the magnetization will be zero at very low temperatures, which is wrong because spontaneous symmetry breaking at low temperatures corresponds to a finite value of magnetization. In mathematical language, the magnetization should be defined under the limit of a vanishing training field, but in real simulation, I simply take the absolute value of the magnetization at each step, and then average over the ensemble.

3.3 Size of Ising spin system

For each dimension I chose a difference lattice size. For 2D, it is 100, for 3D, 20, for 4D, 10, and for 5D, 5. The system size should neither be too large nor too small.

If we define the correlation of the magnetization in time[NB99]

$$\chi(t) = \int dt'[m(t')m(t'+t) - \langle m \rangle^2]$$

where the time unit is one step in the Monte Carlo simulation. It can be proved that the in Monte Carlo the correlation function exponentially decays with a time constant of τ. The time constant is a measure of how long it takes to sample an independent configuration. τ also scales as the other thermodynamic quantities:

$$\tau \approx |t|^{-z\nu}$$

But since the correlation length is

$$\xi \approx |t|^{-\nu}$$

we have essentially

$$\tau \approx \xi^z$$

where z is another kind of critical exponent. At limited lattice size ξ can't grow over the system size because the correlation length describes the typical dimension of a cluster of correlated spins, L. So at limited systems size, the time constant scales like:

$$\tau \approx L^z$$

This means that if we pick smaller system size, we go through the sample space faster, and since for the actual computer CPU, the time it uses for one cluster flip is approximately L^d, small sized systems produce the equilibrium states much faster. But at the same time we are not able to produce the real critical exponent because the time constant scales with L instead of $|t|$. The same thing happens to the other thermodynamic quantities as well, and therefore for a real computational model nothing really diverges at the critical temperature.
For my case I would want the system size to suffice for giving a reasonable critical temperature, and see clearly the trends of critical behavior.

3.4 Time step

There are two things that I should determine before I set the program on a run. One is the Maximum time stepτ_{max}, the other is the sampling rates, or the time separation between two sampling actions. Deciding those two is again the trade off of the calculation time and the quality of the result. First of all, the simulation has to at least reach the equilibrium state, and for a simulation we actually can never reach the ideal equilibrium, so the longer we go the more we are possible of being in an equilibrium state. Second, we have to wait some steps to generate one independent sample, which is the correlation time as defined before.Third, we have to get enough number of independent samples, since the number of independent measurements is

$$n = \frac{t_{max}}{\triangle t}$$

where the $\triangle t$ is the inverse of sampling rate, which should correspond to the time period we wait between two calculations of the thermodynamic quantities, in a real algorithm. The statistical error is just

$$\sigma = \sqrt{\frac{2\tau}{t_{max}}(\langle m^2 \rangle - \langle m \rangle^2)}$$

This is because Monte Carlo is essentially "taking measurement of a system". The general equation for evaluating the standard deviation of measurements applies here, and the more samples we take

and average, the smaller error we are subject to.

In practice, I don't know either the equilibrium time or the correlation time, so I picked several temperature points not far from the critical temperature known from theory, then use the real time plots of the domains to see how they evolve, and at the same time outputting the magnetization against the number of step. If the domains distribution start to exhibit stable features, and the magnetization converges to some constant, then I can make a rough estimate of the equilibrium time. And I estimate the correlation time to be one tenth of the equilibrium time. Of course this method becomes hard when the temperature is set really close to the critical temperature, because it is hard to see the magnetization converges. I did this estimation for 2D case, and applied the same time constant setting to 3D to 5D, because the 2D case has the largest correlation time(near T_c), (for it has the largest L), in this way I can be sure that the time constant setting suffices to yield reasonable results.

For each dimension, I took 140 "independent" equilibrium samples to average through. This is surely not as good as averaging over 1,000 samples but it can still reveal some important trends of critical behaviors.

3.5 Temperature points and curve fitting

Of course the more temperature points I put into the simulation, the more smooth the curve will be, and the better the fitting will be. But since I only care about behaviors near the critical point, I only need to sample more points near the critical temperature. In fact, in real fitting, I need only fit one side of the critical temperature.

At low temperatures, there are very few lattice spins as excitations, which means that they are different from the majority. According to the Wolff Algorithm, if we randomly pick one site, it is very likely that this site won't be one of those excitation sites. And in the expansion of the cluster, it includes almost all of the spins inside the original cluster, this searching procedure is time consuming. In fact, in 2D and very low temperature, for every single spin there is only a probability of $(1-P)^4$ that it will be rejected from the cluster, which is $e^{-8\beta J}$, very small. Therefore the simulation becomes really slow under low temperatures. Of course at very high temperature, the algorithm is also low efficient because it degrades to the single spin flip, because the cluster does not grow larger than 1 lattice spin. But since the high temperature is so far away from the critical temperature, I would not worry about the time cost induced by the high temperature side. The Wolff algorithm is actually best suitable for calculating the system near phase transition, since it was invented to solve the problem of critical slow down near T_cI would rather avoid taking too many points from the low temperature side, but take more points from the high temperature side for a single side fit.

	2D	3D	4D	5D
Theory	2.27	4.5	6.68	*
Simulation	2.32	4.5	6.7	8.9

Table 1: Critical exponents, comparison between theory(T) and simulation (S).

The limited system size can actually distort the thermodynamic quantity vs T curve quite a bit. So the fitting principle is to focus on only one side of the "divergence" and prioritize the fitting accuracy near the critical temperature over the accuracy for points away from the critical temperature. First, the behavior of the system far away from the critical temperature is not a power law, second, for a really small lattice system, there seems very likely an "offset" added to the power law at higher temperatures, so I fitted the curve with an offset of the fitting function, this should not affect the asymptotic behavior near the critical temperature, because the analysis of system behavior near singular points mainly focuses on the singular behavior.

4 Results

I plotted and fitted everything here : https://github.com/2huangyj/cluster-monte-carlo-for-Ising-model.

4.1 Critical temperature

I used the M-T curve for each dimension to obtain the critical temperature, because the M-T curve gives tidies behavior among all the thermodynamic quantities I computed, and I believe that as the order parameter of this second order phase transition, a singularity at the critical temperature for other thermodynamic quantities are harder to fit than a simple non-zero to zero change. The result is shown in the graphics. Clearly the critical temperature has to increase with the dimensionality, because higher dimensions add more neighbors to each single spin.

The critical temperature obtained by my fairly small system size is quite close to the other theoretical/simulation results.The comparison is shown in Table 1, the star in the 5D theory means I did not find an accurate number from literature. The critical behavior is quite instinct and can't be missed when we search for the critical temperature.

4.2 Critical exponents

The critical exponents obtained by fitting each curve is displayed in the Table 2, and the critical exponent for each physical quantity has been defined before in the Introduction part. The two stars in the table means the failure to produce a decent fit. The susceptibility and the heat capacity actually shows nice approximation to the theory, and especially a very obvious trend to Mean Field

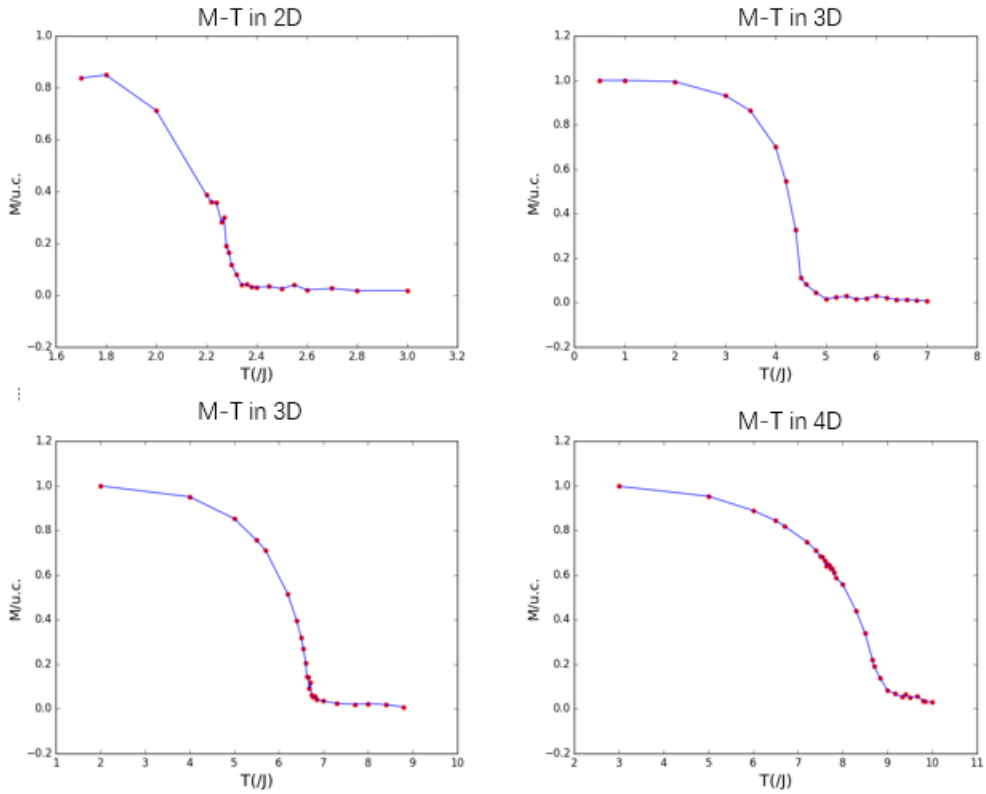

Figure 1: Ising model M-T curve and determination of T_c

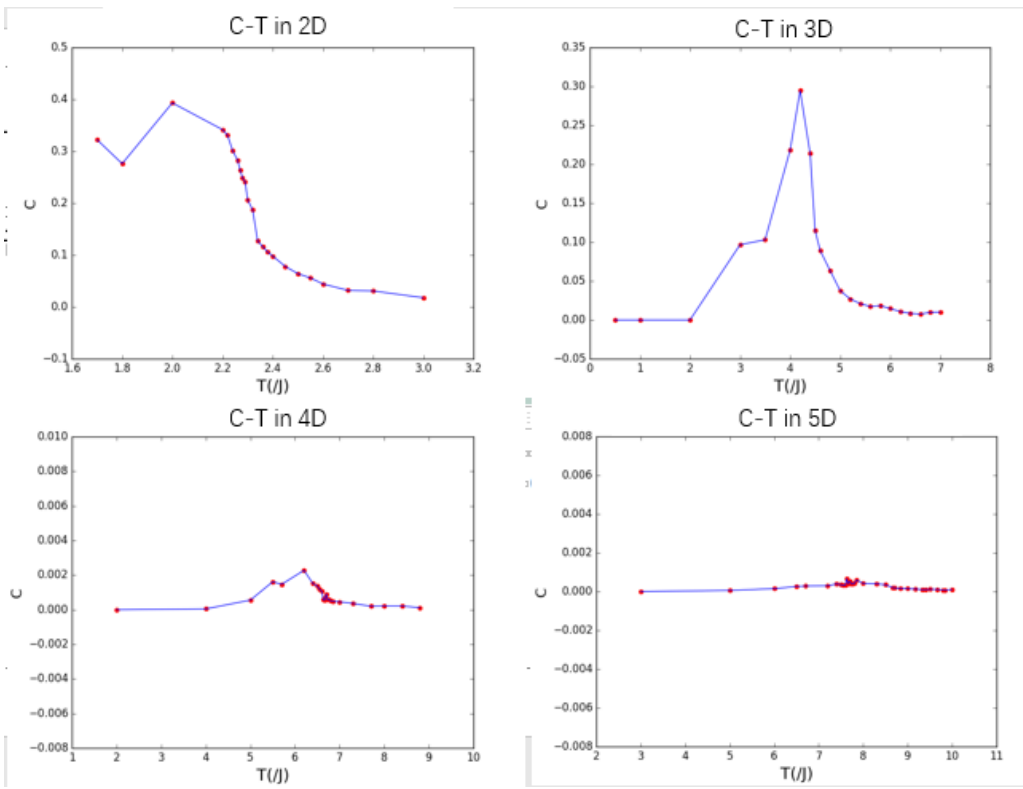

Figure 2: Ising model C-T curve and trend toward MFT

	2D(T)	2D(S)	3D(T)	3D(S)	4D(T)	4D(S)	5D(T)	5D(S)
α	0	*	0.11	0.6	0	0	0	0
δ	15	3	4.7	3	3	3	3	3
γ	7/4	1.3	1.24	1.24	1	1	1	1.1
ν	1	0.2	0.63	0.1	0.5	0.09	0.5	*

Table 2: Critical exponents, comparison between theory(T) and simulation (S).

Theory behavior. Although I used a pretty small system size, some of the result result is not that bad, the size and the time steps I chose was reasonable.

But the magnetization all produces the behavior like in the Mean filed theory, which is supposed to be correct only at high dimensions. I am not sure if it is due to the too small number of points on the curve.

The susceptibility and heat capacity shows a clear trend toward mean field theory as we go into higher dimensions. And for 4D and 5D, all the critical exponents(except for correlation length, which I will discuss later) are very consistent with the known theory, which is mean field theory. Given that I used a much smaller system size than an accurate solution would intuitively require, (for 5D I used only 5*5*5*5!) this is no coincidence- the mean field theory is valid in dimensions higher than 4, and the mean field theory itself means that we can see each spin as affected by only a "local" field, which is effective field for the overall interactions that take into account every lattice site of the system. For our case, this is nothing other than saying that a "local field" picture can reasonably spare the effort of calculating through a large spin system, so the critical behavior can be accurately predicted by just choosing a very small system in 4D and 5D, and of course in higher dimensions. To show this the C-T relations are shown in Figure 2, because the MFT behavior $\alpha = 0$, corresponding to a horizontal line, can be identified without fitting. And it is clear that for 4D and 5D, $\alpha = 0$. For $D = 2$ the α is also supposed to be zero, but it is not, in my simulation, due to the small system size 100*100, although comparable to the 4D and 5D case(10*10*10*10 and 5*5*5*5*5), still does not suffice to produce a value close to the analytical solution, because the mean field theory is not valid for $D = 2$.

The large deviation of the correlation length behavior from the theory is not a surprise: It is much smaller than the theoretical results, because the system size is very limited, and henceforth the system size suppresses the "real" correlation length in the thermodynamic limit, which states that the system goes to infinity.On the one hand, the correlation length is less correlated to the temperature than the system size itself, on the other hand, the size of the systems is far from the thermodynamic limit, so it does not really diverge, actually even a very careful plot will show a flat maximum of the curve instead of a divergence. That explains why the critical exponent is smaller than it should be . I once thought of increasing the systems size and even started running the 2D

case for 500*500 but it seems no hope of yielding a result in time.

In general, the critical exponent for all thermodynamic quantities given by this simulation is not as close to theory as the critical temperature is, technically this is because the lack of points near the critical temperature, the physical reason for this has been stated before, the systems size is far from thermodynamic limit, and also the critical behavior is suppressed by the system size L.

References

[Lui06] E Luijten. Introduction to cluster monte carlo algorithms. In *Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology Volume 1*, pages 13–38. Springer, 2006.

[NB99] MEJ Newman and GT Barkema. *Monte Carlo Methods in Statistical Physics chapter 1-4*. Oxford University Press: New York, USA, 1999.