Cluster algorithms for the Ising model

Monte Carlo methods which overcome the problem of critical slowing down close to second order phase transitions

Cluster algorithms for the Ising model

Monte Carlo methods which overcome the problem of critical slowing down close to second order phase transitions

Contents

- Ising model
- Single spin flip metropolis
- Autocorrelation, Binning Analysis
- Wolff algorithm
- Swendsen-Wang algorithm

Contents

- Ising model
- Single spin flip metropolis
- Autocorrelation, Binning Analysis
- Wolff algorithm
- Swendsen-Wang algorithm

Ising model

- Phase transitions
- One of the simplest statistical models that show a phase transition
- Magnetic systems, opinion models, binary mixtures

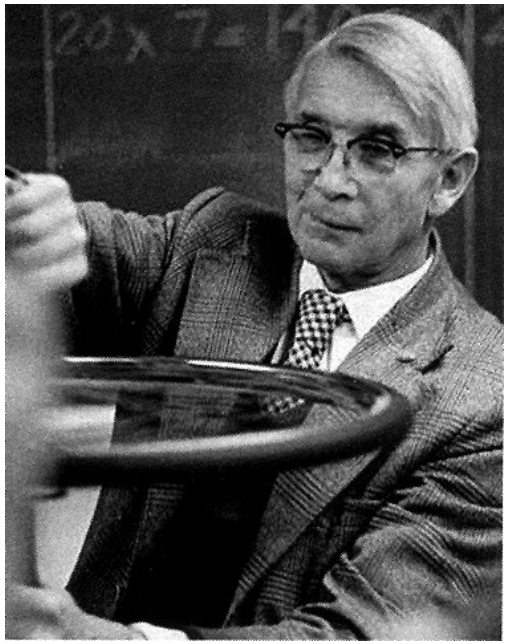


Figure : Ernst Ising (1900 - 1998)

Ising model

- Phase transitions
- One of the simplest statistical models that show a phase transition
- Magnetic systems, opinion models, binary mixtures

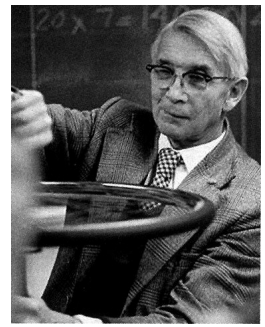


Figure : Ernst Ising (1900 - 1998)

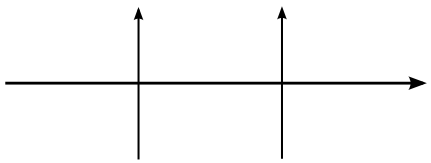
Figure: Ernst Ising, lived from 1900 to 1998

1. Phase transitions are important
2. The Ising model is one of the simplest statistical models that shows a phase transition.
3. Ising model can be used to simulate magnetic systems (ferromagnetic and antiferromagnetic), opinion models and binary mixtures.
4. Ising model invented by Wilhelm Lenz (1888 - 1957) (the same as the Lenz in the Laplace-Runge-Lenz vector) in 1920, his student Ernst Ising solved it in the one-dimensional case 1924.
5. Wolfgang Pauli (1900 - 1958), at whom the road outside is named after, was an assistant of Lenz.
6. Also Otto Stern (1888 - 1969) from the Stern-Gerlach experiment was an assistant of Lenz.

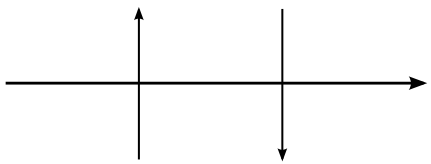
Ising model - Definition

- Discrete integer spins $\sigma_i = \pm 1$ on each lattice site

$$\mathscr{H}(\sigma) = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j - H \sum_i \sigma_i$$



-J

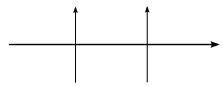


+J

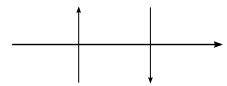
Ising model - Definition

- Discrete integer spins $\sigma_i = \pm 1$ on each lattice site

$$\mathscr{H}(\sigma) = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j - H \sum_i \sigma_i$$



-J



+J

---

1. Lattice with N sites.
2. Lattice of arbitrary geometry
3. Place a spin $\sigma_i = \pm 1$ on each lattice site
4. Hamiltonian:
5. Sum over nearest neighbours.
6. J: interaction, H: external field
7. Two aligned spins give a contribution of -J
8. Two antialligned spins give a contribution of J
9. J_{ij} in general case, $J > 0$: ferromagnetic, $J < 0$: antiferromagnetic
10. For J¿0, spins like to be aligned!
11. For positive J: More favorable for the spins to be aligned!

Ising model - Canonical ensemble

$$p(\sigma, T) = \frac{e^{-\beta \mathcal{H}(\sigma)}}{\mathcal{Z}(T)}, \qquad \beta = \frac{1}{k_B T}$$

$$\langle M \rangle_T = \sum_{\sigma} M(\sigma) p(\sigma, T)$$

Ising model - Canonical ensemble

$$p(\sigma, T) = \frac{e^{-\beta \mathscr{H}(\sigma)}}{\mathscr{Z}(T)}, \quad \beta = \frac{1}{k_B T}$$

$$\langle M \rangle_T = \sum_\sigma M(\sigma) p(\sigma, T)$$

---

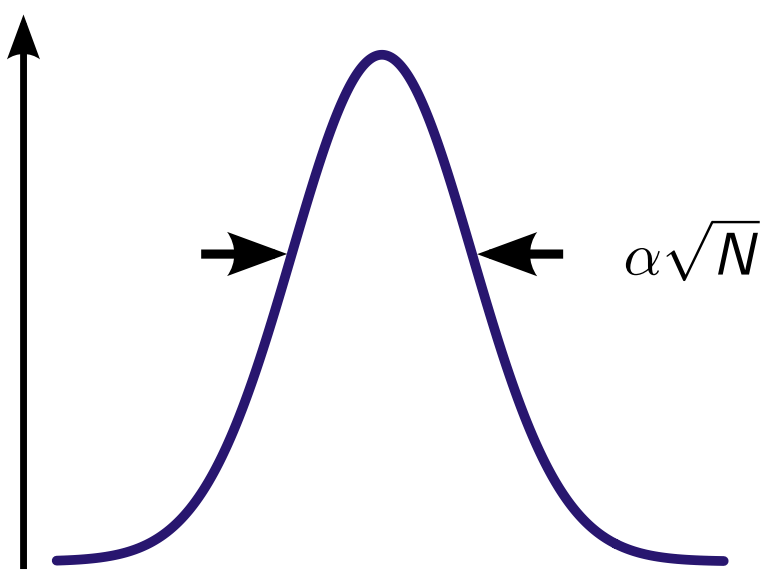
1. Configuration probability given by boltzmann distribution, Z partition function, given as

$$\mathscr{Z}(T) = \sum_\sigma e^{-\beta \mathscr{H}(\sigma)}, \quad \beta = \frac{1}{k_B T}$$

2. So we are looking at a canonical system with constant temperature T.
3. Measurement value of a function, e.g. magnetization, is given by the sum over all states of the measurement value at the configuration times the configuration probability.

Ising model - Monte Carlo

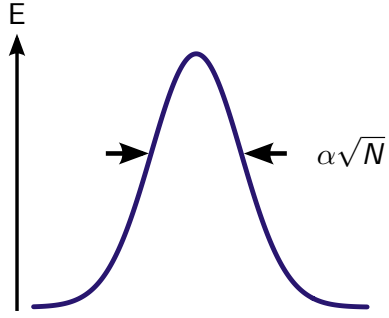
- We can't compute all configurations (2^N)
- We can't sample uniformely distributed over energy



- Solution: Importance sampling using Metropolis algorithm

Ising model - Monte Carlo

- We can't compute all configurations (2^N)
- We can't sample uniformely distributed over energy



- Solution: Importance sampling using Metropolis algorithm

---

1. Why not? $\rightarrow 2^N = 2^{L^d}$ e.g. $L = systemSize = 15$ in 2 dimensions: $2^{15^2} = 2^{225} = 5 \cdot 10^{67}$
2. We can't compute the exact expectation value of an observable. But that's what we're interested in.
3. Because the distribution of the average energy gets sharper with increasing size $\left(\alpha \sqrt{L^d}\right)$.

Single spin flip metropolis - Algorithm

$$A(X \to Y) = \min\left(1, \frac{p(Y)}{p(X)}\right)$$
$$= \min\left(1, e^{-\beta[E(Y)-E(X)]}\right)$$

1. Choose one site (uniformly randomly)
2. Generate new trial configuration Y by flipping spin
3. Accept new configuration with transition probability above

Single spin flip metropolis - Algorithm

$$A(X \to Y) = \min\left(1, \frac{p(Y)}{p(X)}\right)$$
$$= \min\left(1, e^{-\beta[E(Y)-E(X)]}\right)$$

1. Choose one site (uniformly randomly)
2. Generate new trial configuration Y by flipping spin
3. Accept new configuration with transition probability above

---

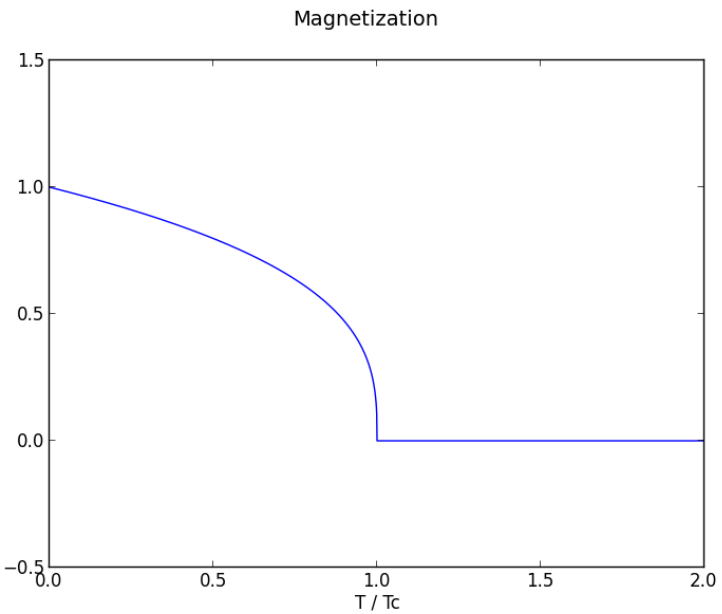
If energy decreases, always accept. If energy increases, accept with probability $e^{-\beta\Delta E}$.Blazingly fast (Troells: 2 flips/ns?), easy to implement.

1. New state given by spinflip at this site.

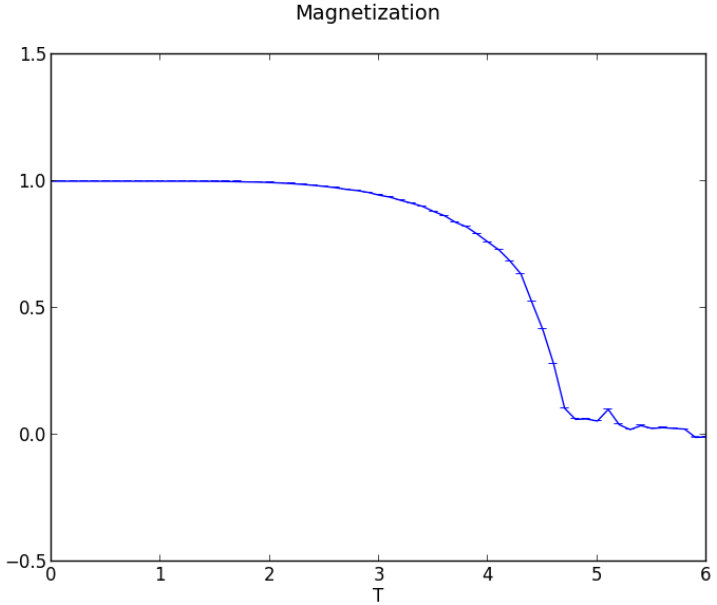
Single spin flip metropolis - Results

Magnetization $M(T) = \left\langle \frac{1}{N} \sum_{i=1}^{N} \sigma_i \right\rangle_T$

Expected



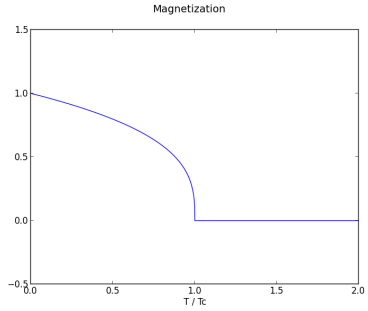
Obtained



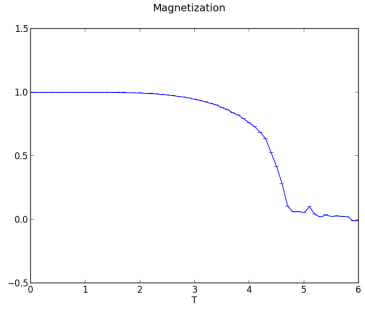
Single spin flip metropolis - Results

Magnetization $M(T) = \left\langle \frac{1}{N} \sum_{i=1}^{N} \sigma_i \right\rangle_T$

Expected



Obtained



---

1. I implemented this single spin flip metropolis algorithm and would like to show you some results.
2. Let's look at the spontaneous magnetization.
3. In theory, we expect a spontaneous magnetization below T_C. We expect a symmetry breaking (either positive or negative Magnetization in this case) and a discontinuity in the derivative of the magnetization at T_C. We want something linear to $(T - T_C)^{\nu}$.
4. Well, this is what we get. It looks perfect - right? We see a spontaneous magnetization in the simulation. The discontinuity in its derivative is not there, but this is okay, since we are looking at a finite sized system. Well, life is not so easy. This result is in fact not the truth. Let me show you what we expect for the spontaneous magnetization.

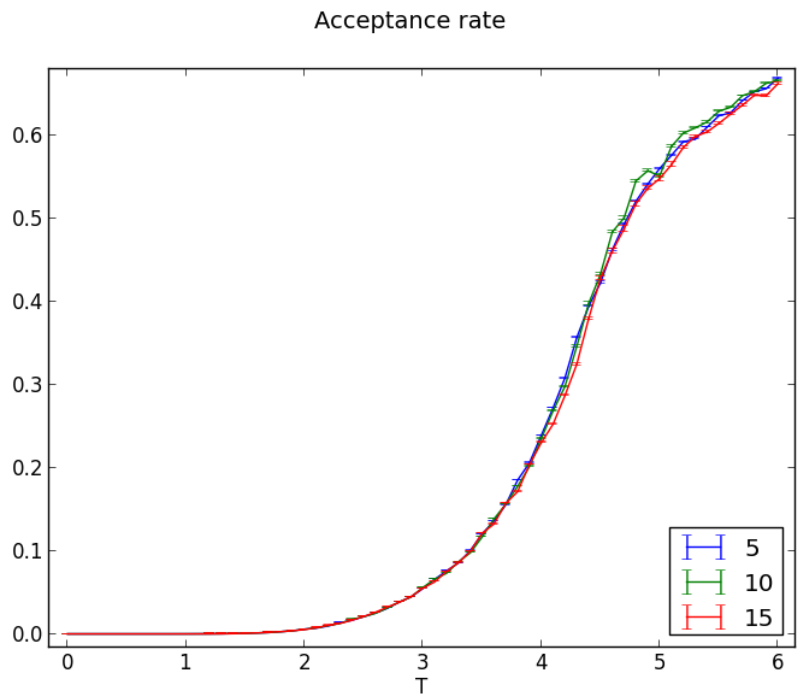
Ising model - 2 spins

Ising model - 2 spins

Show that we should expect the magnetization to vanish.

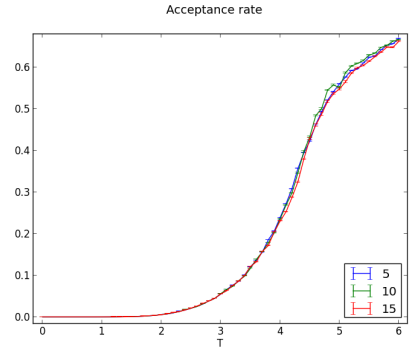
Single spin flip metropolis - To flip or not to flip

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



Single spin flip metropolis - To flip or not to flip

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

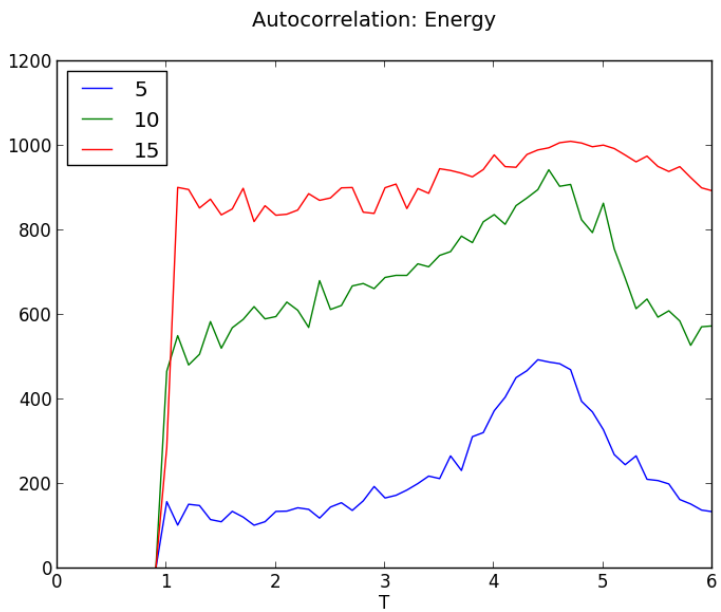


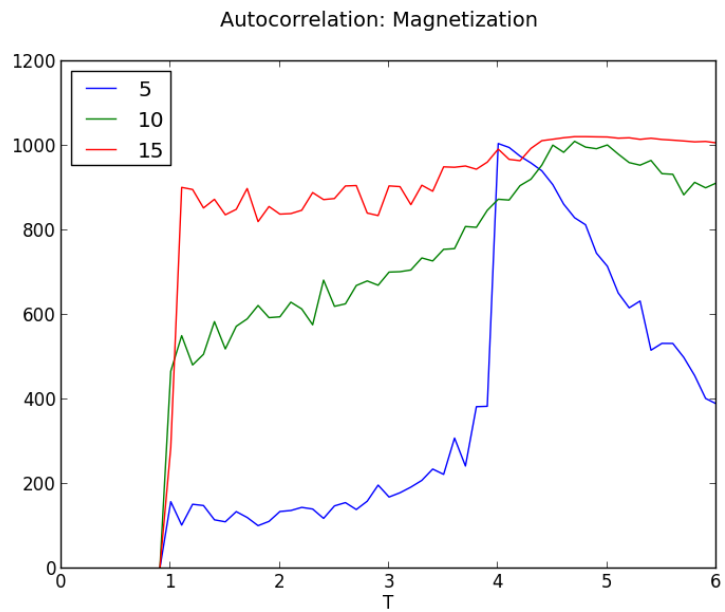
---

1. Now that we know that the magnetization should vanish, let's explore why it hasn't done so.
2. The hamiltonian of the system states clearly that the spins like to be aligned. They were all $+1$ at the beginning, and then thermalized (brought to thermodynamic equilibrium) in 3 sweeps (abbrevation for: 3 times number of sites spinflips).
3. The point is that the spins can also be aligned when they are all -1. But we didn't see that happen. Why?
4. I took 0.5 million measurement values, after every single spin flip. This means that the spin on every site could have flipped around 150 times. But it didn't. This is why:
5. What we see here is the acceptance rate of a spin flip. It is one if the suggested new configuration through the single spin flip was accepted. Well, it's actually almost zero for low temperatures. So we move ultra slow through phase space, which is no fun, because we want to do an ergodic sampling.
6. $$\frac{N}{a(T)} = \#\text{ trials}$$
7. We want to change flip the magnetization of the system completely at least 100 times to have good statistics. For uncorrelated measurements we then have an error of 10%.

Autocorrelation time

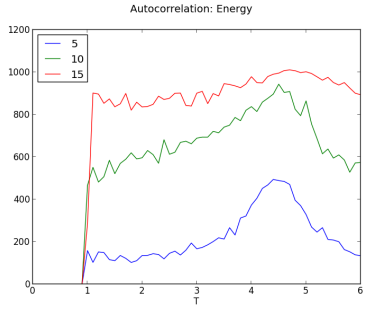
$$\phi_A(t) = \frac{\langle A(t_0)A(t)\rangle - \langle A\rangle^2}{\langle A(t_0)^2\rangle - \langle A\rangle^2} \propto e^{-\frac{t}{\tau_A}}$$

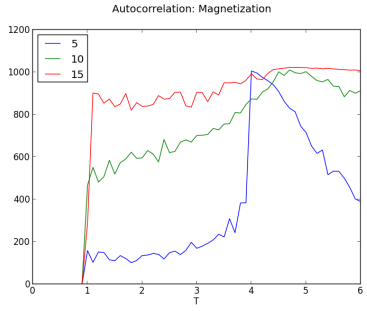




Autocorrelation time

$$\phi_A(t) = \frac{\langle A(t_0)A(t)\rangle - \langle A\rangle^2}{\langle A(t_0)^2\rangle - \langle A\rangle^2} \propto e^{-\frac{t}{\tau_A}}$$





Definition and plots from measurements, magnetization and energy.

1. There is another approach to describe what's going on. Not accepting a spinflip or only flipping one spin at a time means that the new configuration is highly correlated to the old one.
2. We define the linear autocorrelation function of a measurement value as a ratio of the covariance of two observables at a given point in time and at time t_0.
3. We see that the energy converges much faster than the magnetization. I read that people often focus on the energy, but here you really have to focues on the magnetization to be sure.

Binning analysis I

$$\mathrm{Var}X := E\left[X^2\right] - E\left[X\right]^2$$

$$(\Delta X)^2 = \frac{\mathrm{Var}X}{N}\left(1 + 2\tau_X\right)$$

$$\tau_X := \frac{\sum_{i=1}^{N}\sum_{t}\left(E\left[X_i X_{i+t}\right] - E\left[X\right]^2\right)}{\mathrm{Var}X}$$

Binning analysis I

$$\mathrm{Var}X := E\left[X^2\right] - E\left[X\right]^2$$

$$(\Delta X)^2 = \frac{\mathrm{Var}X}{N}\left(1 + 2\tau_X\right)$$

$$\tau_X := \frac{\sum_{i=1}^{N}\sum_{t}\left(E\left[X_i X_{i+t}\right] - E\left[X\right]^2\right)}{\mathrm{Var}X}$$

For uncorrelated samples we know:

$$E\left[X_i X_j\right] = E\left[X_i\right] E\left[X_j\right] \text{ for } i \neq j$$

Binning analysis II

$$A_i^{(l)} = \frac{1}{2}\left(A_{2i-1}^{(l-1)} + A_{2i}^{(l-1)}\right)$$

$$\Delta^{(l)} = \sqrt{\mathrm{Var}A^{(l)}/M^{(l)}} \overset{l\to\infty}{\longrightarrow} \Delta = \sqrt{(1+2\tau_A)\mathrm{Var}A/M}$$

$$\tau_A = \lim_{l\to\infty}\left(\frac{2^l \mathrm{Var}A^{(l)}}{\mathrm{Var}A^{(0)}} - 1\right)$$

Binning analysis II

$$A_i^{(l)} = \frac{1}{2}\left(A_{2i-1}^{(l-1)} + A_{2i}^{(l-1)}\right)$$

$$\Delta^{(l)} = \sqrt{\mathrm{Var}A^{(l)}/M^{(l)}} \overset{l\to\infty}{\longrightarrow} \Delta = \sqrt{(1+2\tau_A)\mathrm{Var}A/M}$$

$$\tau_A = \lim_{l\to\infty}\left(\frac{2^l \mathrm{Var}A^{(l)}}{\mathrm{Var}A^{(0)}} - 1\right)$$

Draw graphics with the bins and the levels.

What really happens in the System

What really happens in the System

Draw domains on board and show single spin flip only is useful at borders of domains.

Cluster algorithms: Wolff algorithm

$$P(\sigma_x, \sigma_y) = 1 - \min\left(1, e^{-2\beta\sigma_x\sigma_y}\right)$$

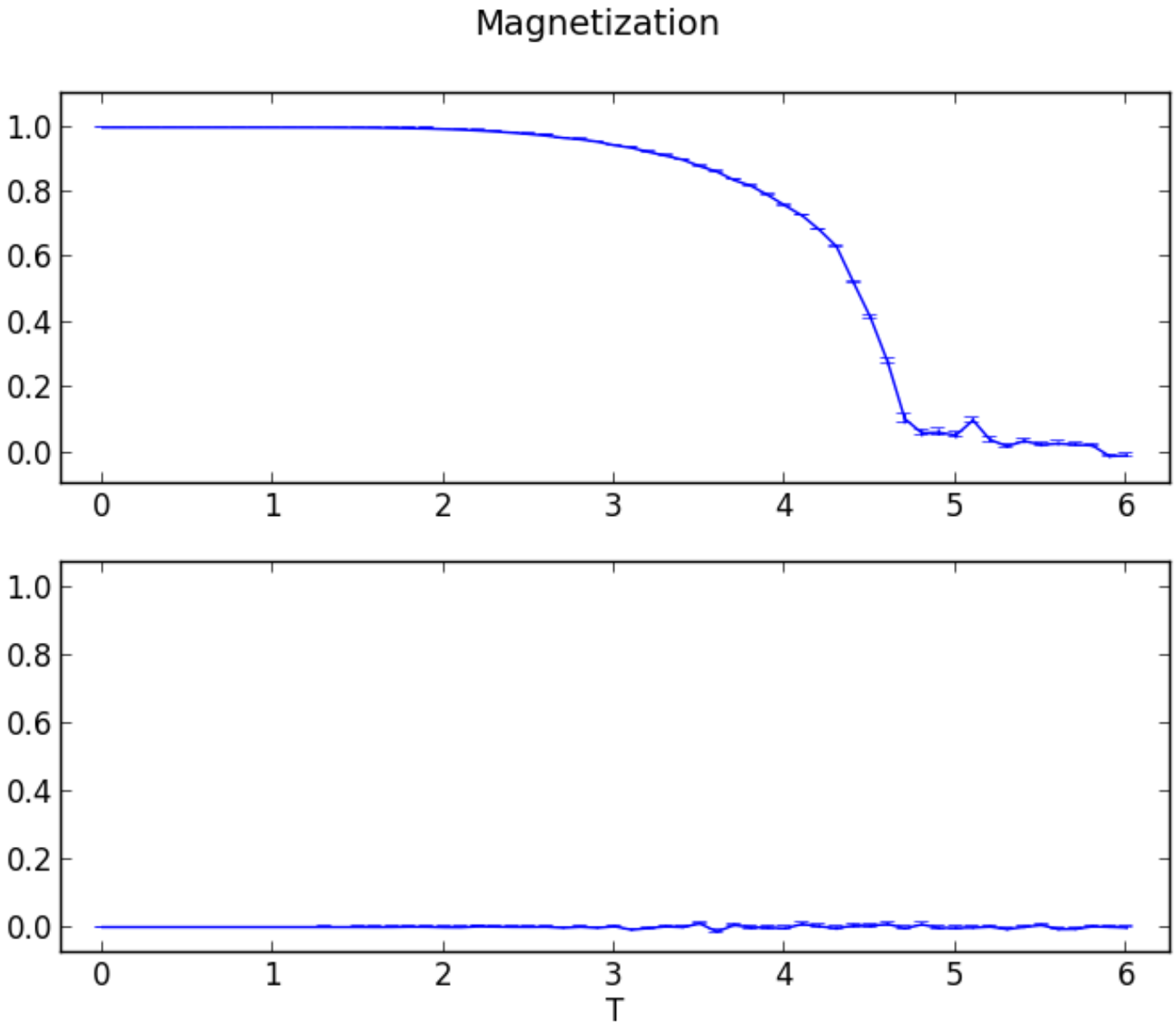
1. Choose one site (uniformly randomly)
2. Flip its spin and add it to the cluster
3. For all sites in the cluster:
   3.1 Visit every unknown neighbour, flip its spin and add it to the cluster with probability given above

Cluster algorithms: Wolff algorithm

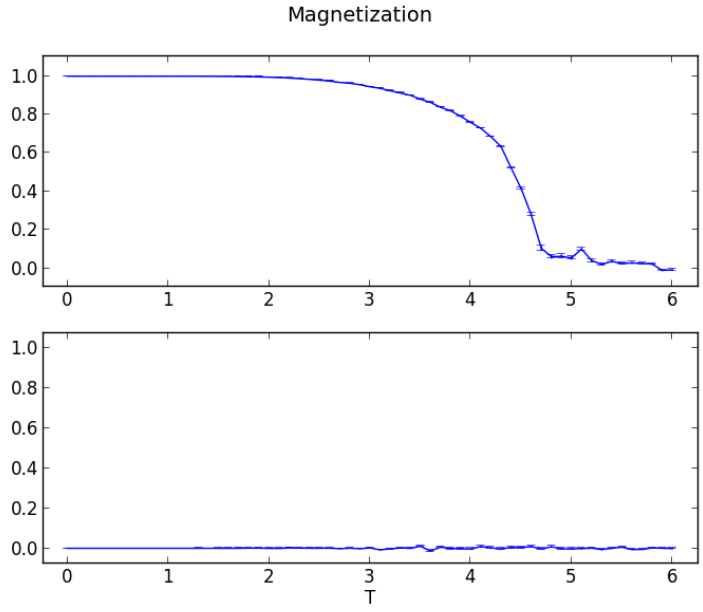
$$P(\sigma_x, \sigma_y) = 1 - \min\left(1, e^{-2\beta\sigma_x\sigma_y}\right)$$

1. Choose one site (uniformly randomly)
2. Flip its spin and add it to the cluster
3. For all sites in the cluster:
   3.1 Visit every unknown neighbour, flip its spin and add it to the cluster with probability given above

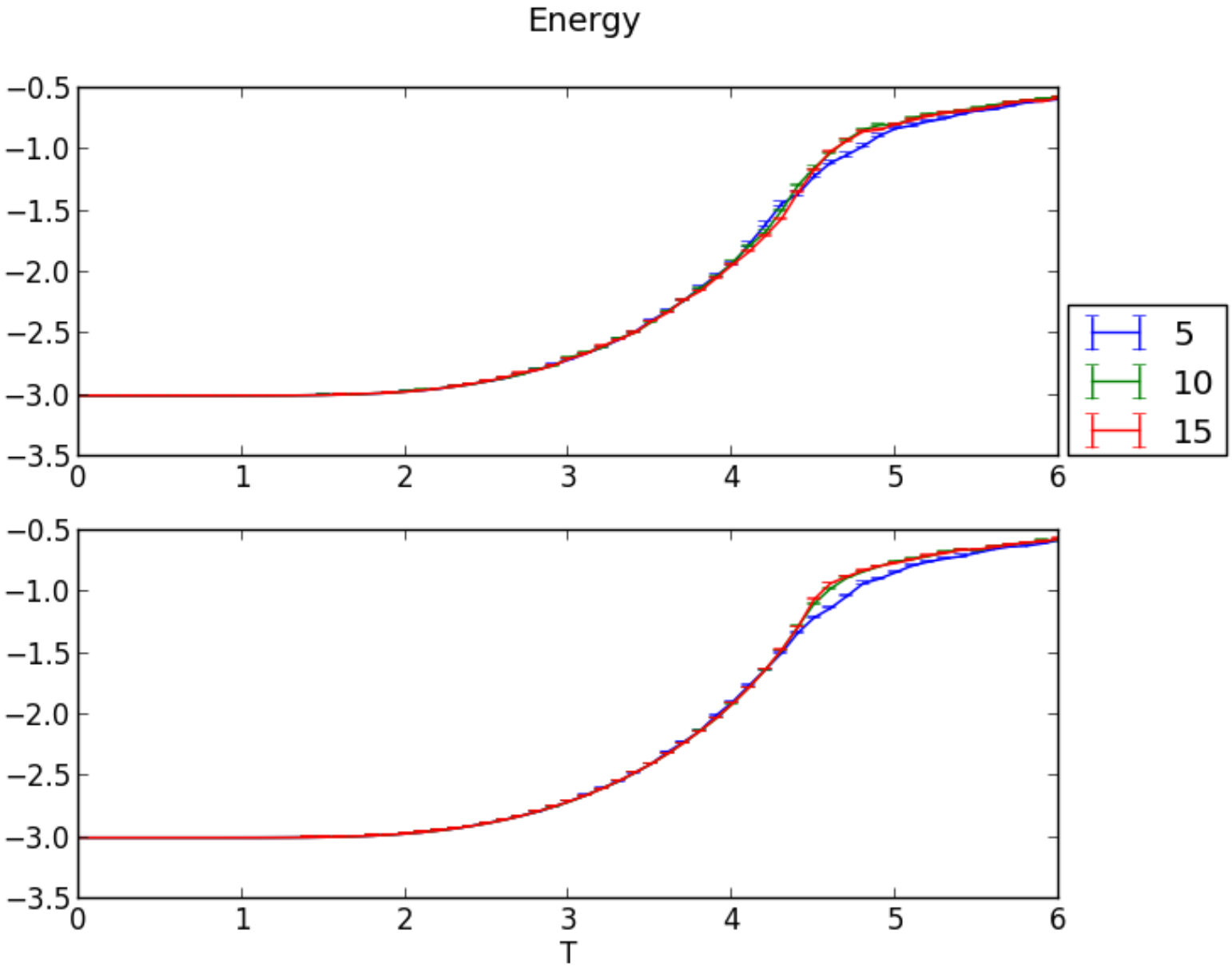
Why cluster algorithms are better



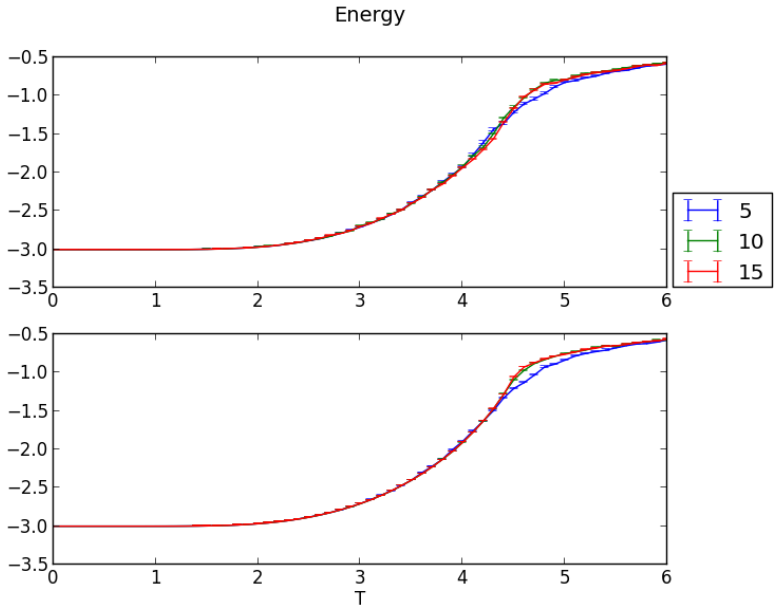
Why cluster algorithms are better



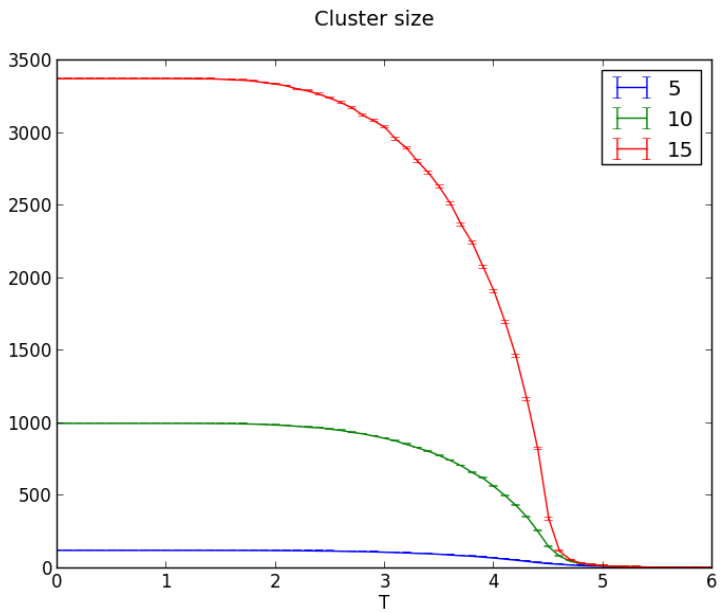
Energy

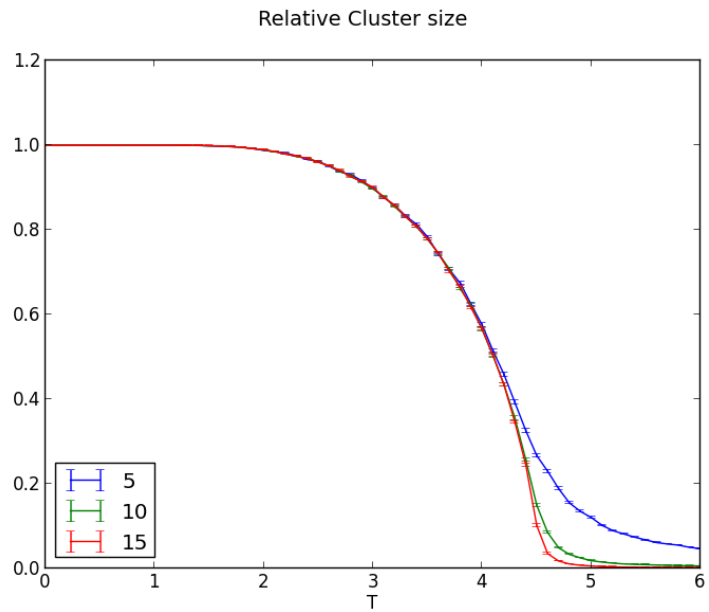


Energy

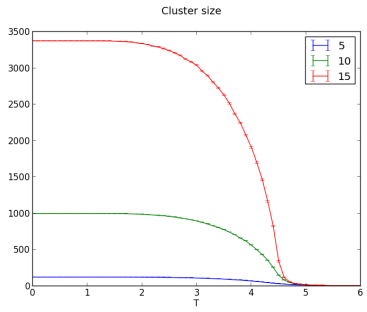


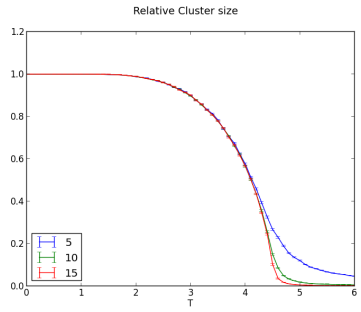
Cluster size



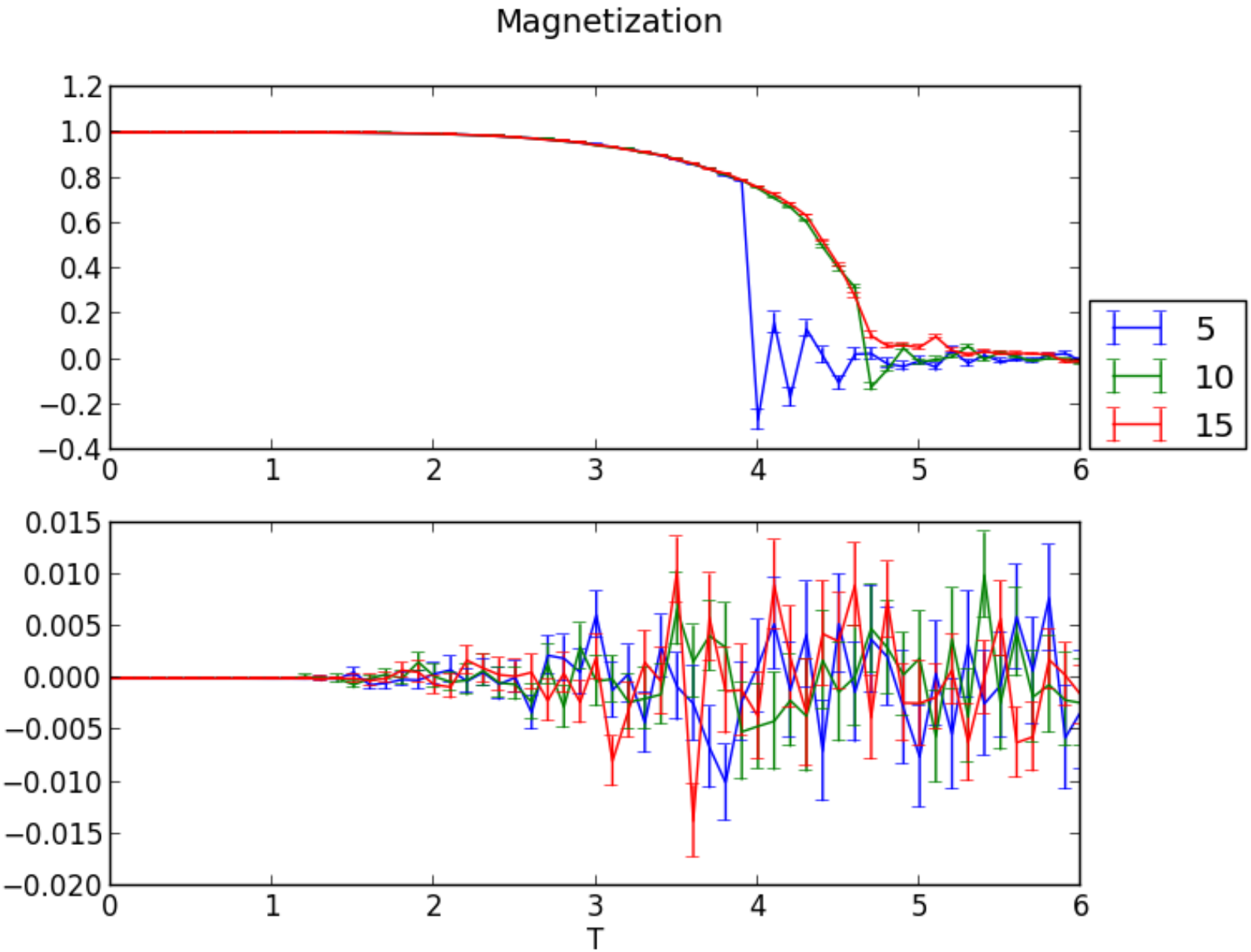


Cluster size

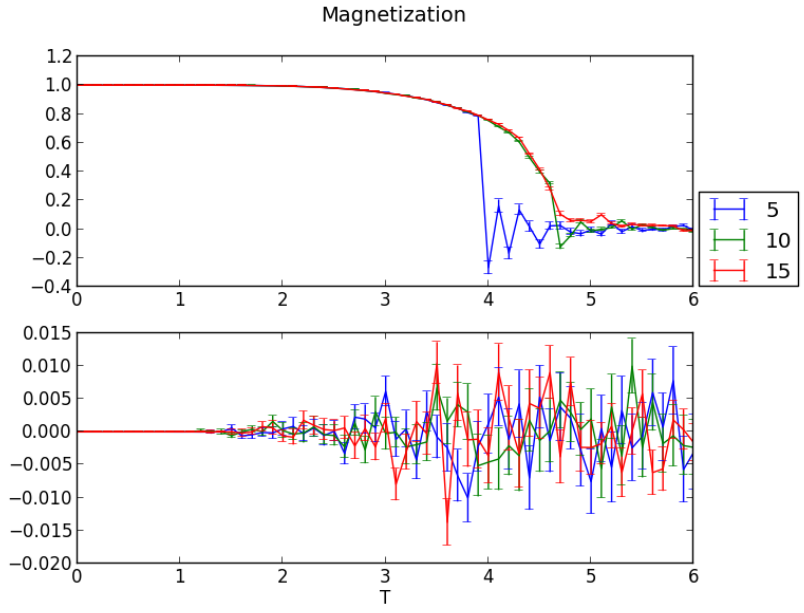




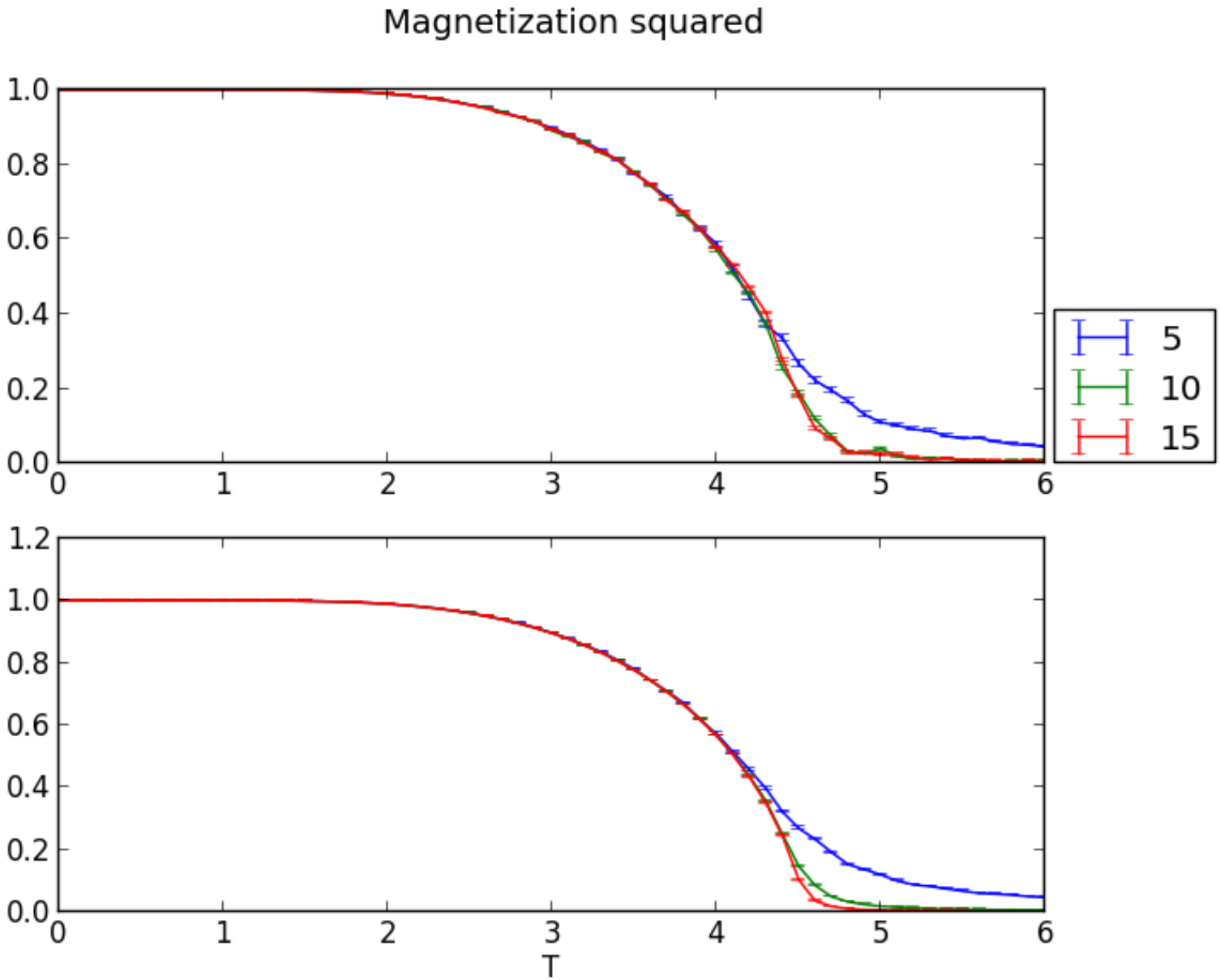
Magnetization



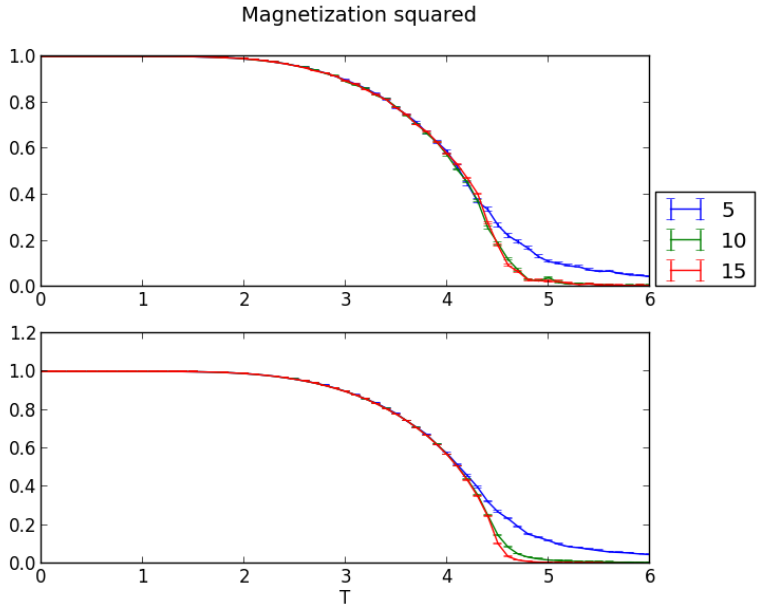
Magnetization



Magnetization squared

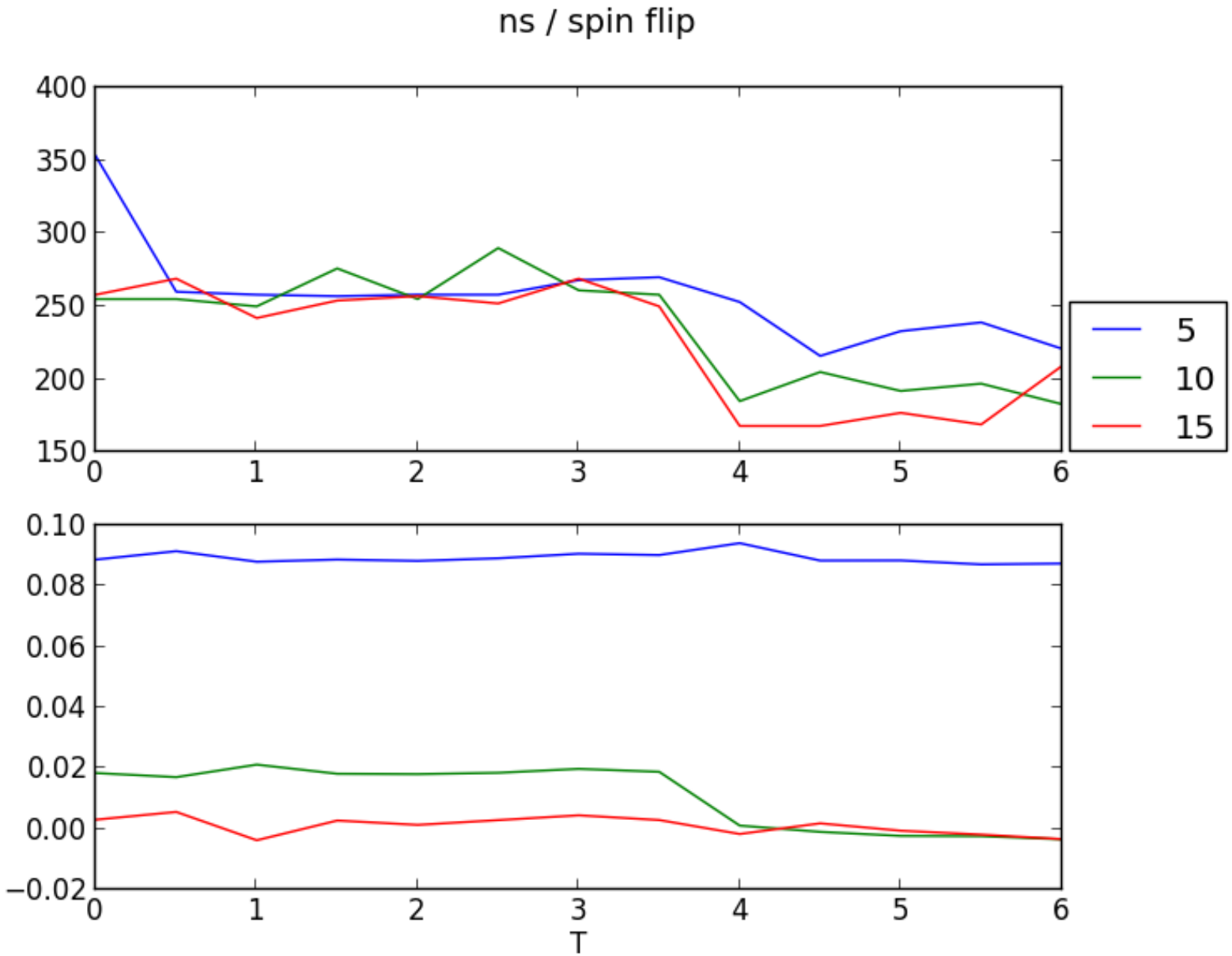


Magnetization squared

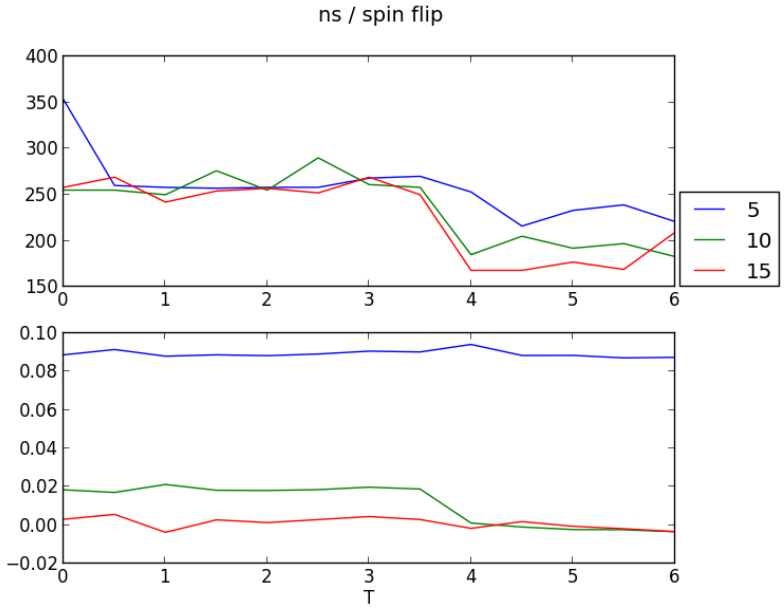


Magnetization itself is a bad order parameter, since it can have two signs so it cancels. Magnetization squared is a good order parameter, because it consists of sums of spatial spin correlations. It shows the phase transition.

Computation time per spin flip



Computation time per spin flip



Swendsen-Wang

Similar to Wolff, but:

- Partitions the whole lattice in clusters
- Randomly chooses a new (same) spin for every cluster
- Always touches every spin

Swendsen-Wang

Similar to Wolff, but:

- Partitions the whole lattice in clusters
- Randomly chooses a new (same) spin for every cluster
- Always touches every spin

Summary

Summary

References

References

Questions

Questions

Binning analysis in detail

$$\mathrm{Var}X := E\left[X^2\right] - E\left[X\right]^2$$

$$\begin{aligned}
(\Delta X)^2 &= \frac{1}{N^2}\sum_{i,j=1}^{N}\left(E\left[X_i X_j\right] - E\left[X\right]^2\right) \\
&= \frac{\mathrm{Var}X}{N} + \frac{1}{N^2}\sum_{i\neq j}\left(E\left[X_i X_j\right] - E\left[X\right]^2\right) \\
&= \frac{\mathrm{Var}X}{N} + \frac{2}{N^2}\sum_{i=1}^{N}\sum_{t}\left(E\left[X_i X_{i+t}\right] - E\left[X\right]^2\right) \\
&:= \frac{\mathrm{Var}X}{N}\left(1 + 2\tau_X\right)
\end{aligned}$$

Binning analysis in detail

$$\mathrm{Var}X := E\left[X^2\right] - E\left[X\right]^2$$

$$\begin{aligned}
(\Delta X)^2 &= \frac{1}{N^2}\sum_{i,j=1}^{N}\left(E\left[X_iX_j\right] - E\left[X\right]^2\right) \\
&= \frac{\mathrm{Var}X}{N} + \frac{1}{N^2}\sum_{i\neq j}\left(E\left[X_iX_j\right] - E\left[X\right]^2\right) \\
&= \frac{\mathrm{Var}X}{N} + \frac{2}{N^2}\sum_{i=1}^{N}\sum_{t}\left(E\left[X_iX_{i+t}\right] - E\left[X\right]^2\right) \\
&:= \frac{\mathrm{Var}X}{N}\left(1 + 2\tau_X\right)
\end{aligned}$$

For uncorrelated samples we know:

$$E\left[X_iX_j\right] = E\left[X_i\right]E\left[X_j\right] \text{ for } i \neq j$$

The correlation time of a variable is defined as:

$$\tau_X := \frac{\sum_t\left(E\left[X_iX_{i+t}\right] - E\left[X\right]^2\right)}{\mathrm{Var}X}$$

Wolff or Swendsen-Wang?

Swendsen-Wang better for parallelization because it touches the whole lattice.

Wolff or Swendsen-Wang?

Swendsen-Wang better for parallelization because it touches the whole lattice.