

Laboratory of Computational Physics

BY LUCA CASSIA

Dipartimento di Fisica, Università di Milano-Bicocca
I-20126 Milano, Italy

Email: l.cassia@campus.unimib.it

Table of contents

1 Ising Model 2d	1
1.1 Thermalization	1
1.2 Autocorrelation Times	4
1.3 Binning Analysis	6
1.4 Observables	9
1.5 β Critical	11
1.6 Probability Distribution Functions	16
1.7 Spatial Correlations	19
1.8 Finite Size Scaling	22

1 Ising Model 2d

In this section we study the statistical properties of a 2d Ising Model. We consider a system of $L \times L$ spins situated on the points of a regular square lattice with periodic boundary conditions (PBC) in both directions. Each spin interacts with its nearest neighbours inside the lattice, with an Hamiltonian:

$$H = - \sum_{\langle i,j \rangle} \sigma_i \sigma_j \quad (1)$$

where the sum is taken only over the set of unordered pairs $\langle i, j \rangle$ such that σ_i and σ_j are nearest neighbours.

In order to obtain expectation values for physical quantities of the system we should be able to sum over the space of all the spin configurations, or at least to sample configurations from this space with probability:

$$P(\{\sigma\}) \sim e^{-\beta H(\{\sigma\})} \quad (2)$$

which reproduces the integration measure of the functional integral. The obvious choice for numerical simulations is the second one, which we implement through Monte Carlo (MC) algorithms. In particular we focus on the Metropolis-Hastings (MH) and the Swendsen-Wang (SW) algorithms.

1.1 Thermalization

Since we do not know where to start in our Markov process, we first initialize the system in a disordered configuration (hot start) and then evolve for a certain Markov time until the system reaches equilibrium. This process is usually called *thermalization*.

After a few run we decided to opt for a cold start approach (i.e., all the spins are initially aligned). The reason for this choice is that, when using MH in 2-dimensional finite-size systems with periodic boundary conditions, the non trivial topology of the lattice allows the existence of stable configurations of the type depicted in (Fig.1), that often arise when the system is rapidly cooled from a disordered configuration. Those configurations would invalidate the mixing process and the sampling of relevant configurations at low temperatures. Therefore, by using a cold start, we exclude this possibility.

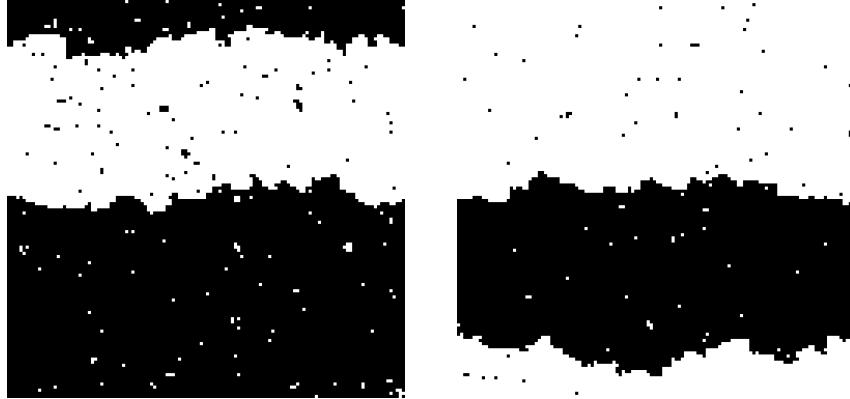


Figure 1. Illustration of two topologically non-trivial configurations obtained with the MH algorithm on a lattice 128×128 .

We expect the MH algorithm to be slower to thermalize because of the local update it employs as opposed to the cluster update of the SW algorithm.

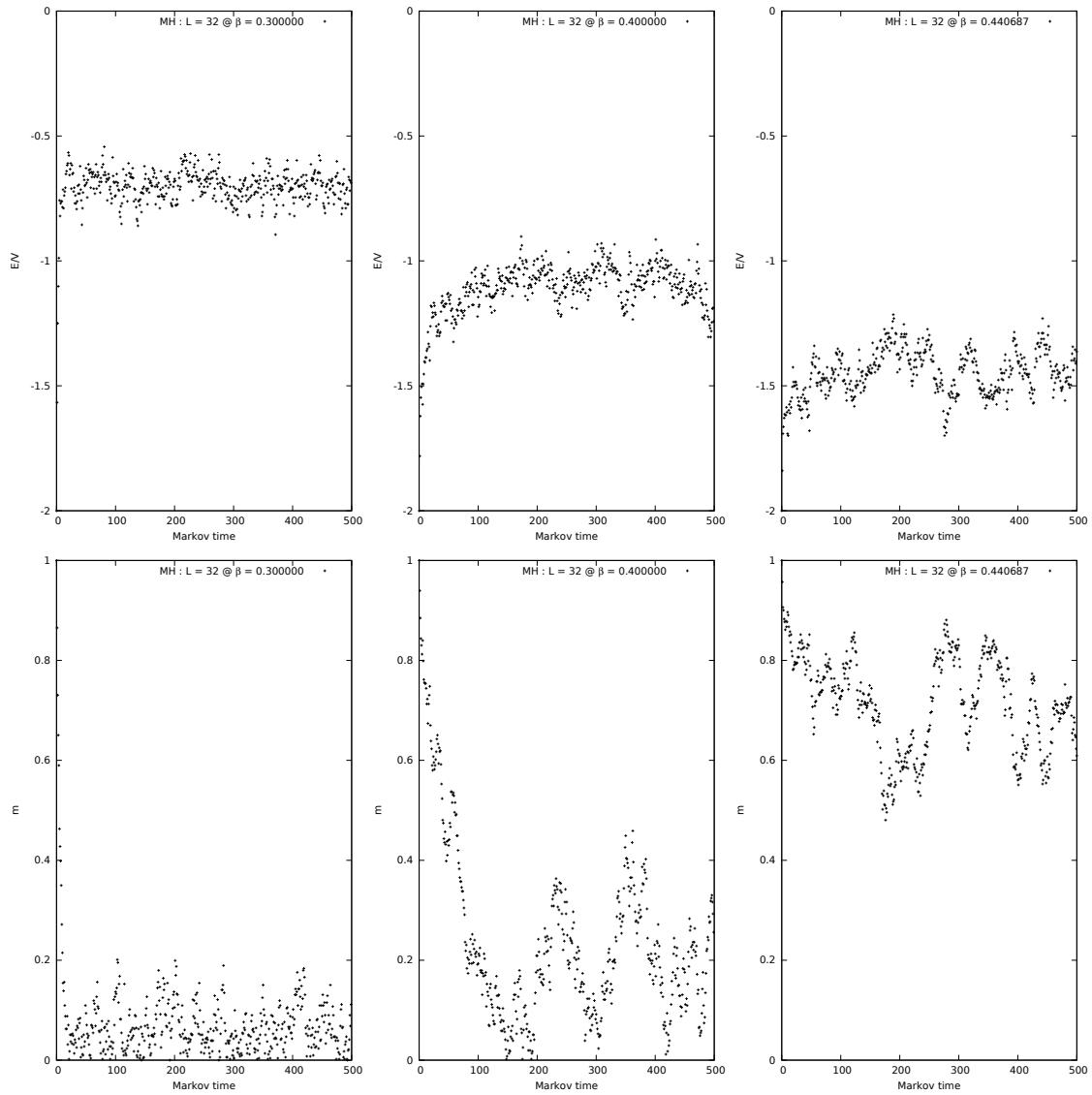


Figure 2. Plot of the energy density and magnetization for the MH markov process at different values of β .

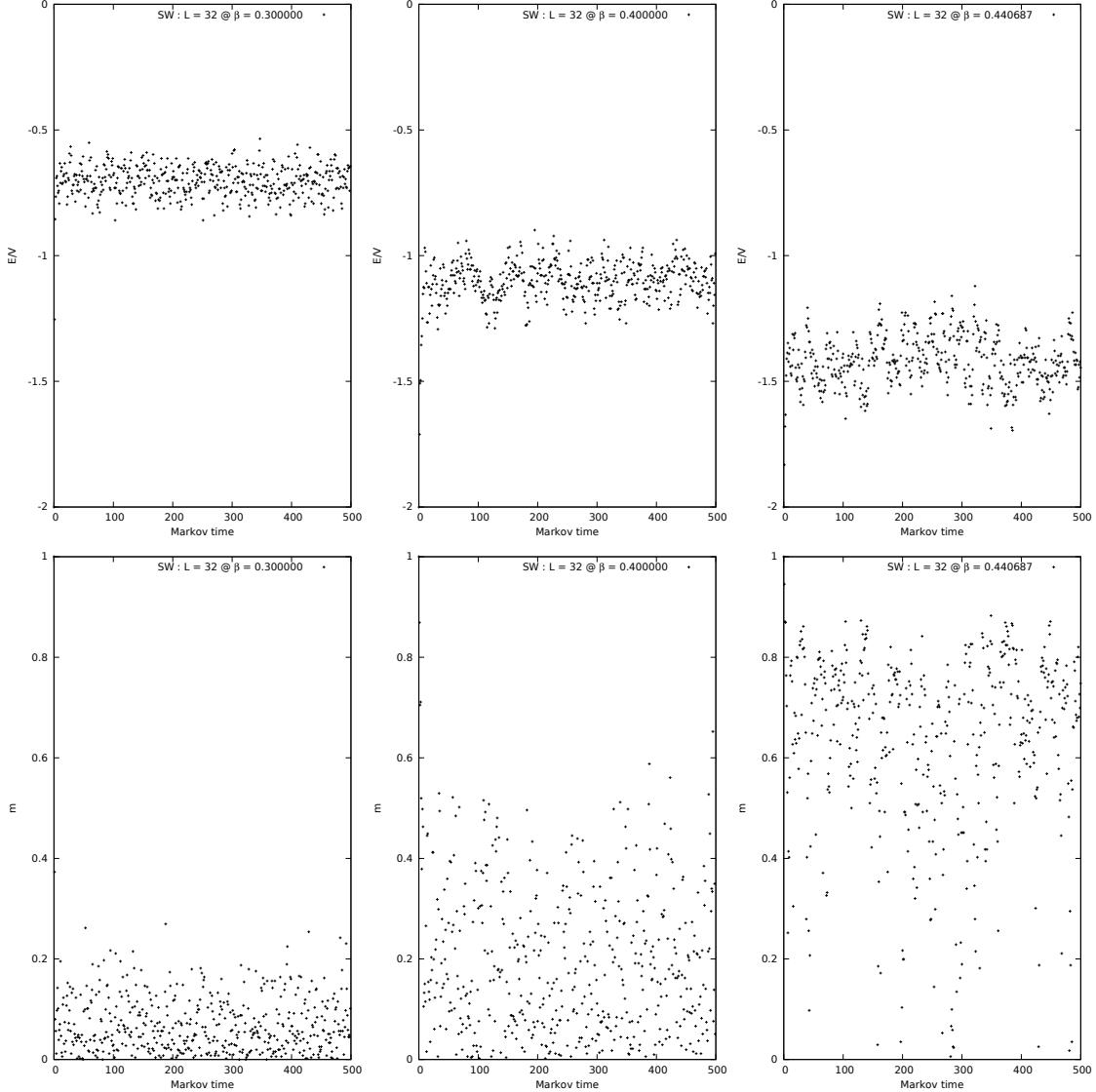


Figure 3. Plot of the energy density and magnetization for the SW markov process at different values of β .

The MH algorithm reaches thermalization very rapidly for inverse temperatures β far from the critical value β_c , while for values of β near the phase transition it becomes much slower (Fig.2).

This effect is called *critical slowing down* and it tells us that, when using local update algorithms, successive steps in the Markov chain are strongly correlated and therefore only approximately every L^2 sweeps through the lattice, a statistically independent measurement can be taken. Close to a critical point, in fact, the autocorrelation time (in the infinite-volume limit) typically scales as:

$$\tau_{\mathcal{O}, \text{exp}} \propto \xi^z \quad (3)$$

where $z \geq 0$ is the so-called dynamical critical exponent. The spatial correlation length ξ would diverge at the critical point, but in a finite system it is limited by the size L of the system and the scaling law (3) becomes:

$$\tau_{\mathcal{O}, \text{exp}} \propto L^z \quad (4)$$

where $z \approx 2$ for local dynamics (MH) and $z \approx 0$ for SW.

We also remark that for smaller values of β , the energy of the system can have larger fluctuations around its average and therefore thermal equilibrium is reached faster.

On the other hand the SW algorithm mixes very fast even at the phase transition (Fig.3) being a multicluster update algorithm (Fig.3).

For the rest of the simulations we considered the system to be at equilibrium after a thermalization time $t \approx 1000$ both for MH and for SW.

1.2 Autocorrelation Times

We now study the correlations present between consecutive configurations sampled by the two algorithms in order to obtain the characteristic time after which we can consider two configurations to be statistically independent. This is called *integrated autocorrelation time* and is computed in the following way:

$$\tau_{\mathcal{O},\text{int}} = \frac{1}{2} + \sum_{k=1}^{k_{\max}} R(k) \quad (5)$$

where \mathcal{O} is the physical observable for which we compute autocorrelation and $R(k)$ is defined as:

$$R(k) = \frac{1}{(n-k)\sigma^2} \sum_{t=1}^{n-k} (\mathcal{O}_t - \mu)(\mathcal{O}_{t+k} - \mu) \quad (6)$$

here n is the total number of samples of \mathcal{O} and μ, σ^2 are the mean and variance of the process. For large time separations k , $R(k)$ decays exponentially:

$$R(k) \xrightarrow{k \rightarrow \infty} r e^{-k/\tau_{\mathcal{O},\text{exp}}} \quad (7)$$

which defines the exponential autocorrelation time $\tau_{\mathcal{O},\text{exp}}$.

We immediately see from (Fig.4) that the MC process is strongly correlated in time particularly near the phase transition:

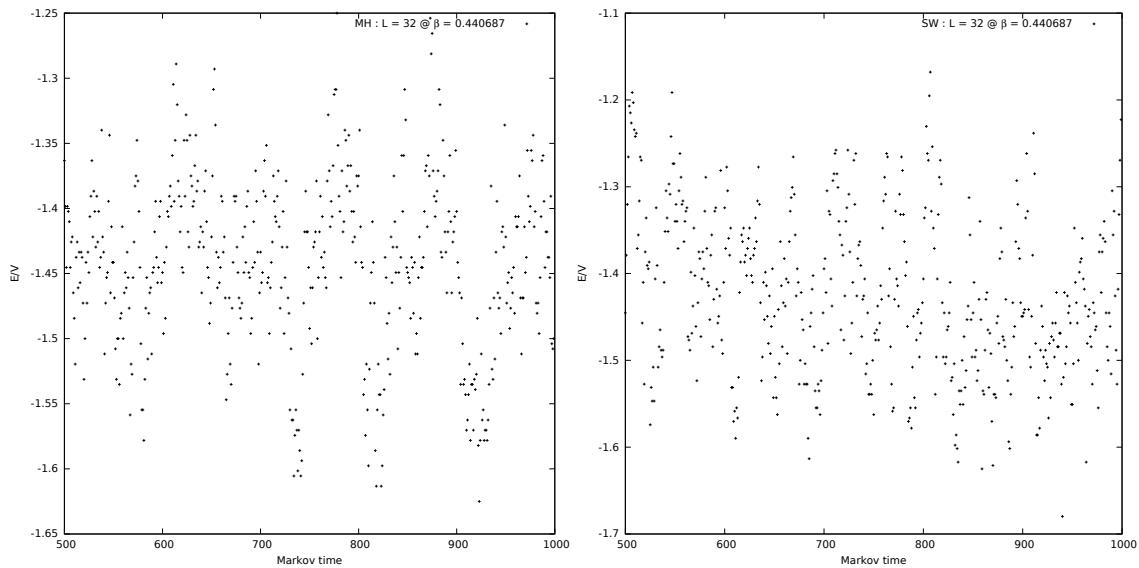


Figure 4. Comparison of the energy for the two algorithms at $\beta = \beta_c$ and $L = 32$. (Left) MH (Right) SW.

We now compute the autocorrelation time τ_{int} for the observable $e = E / V$ using formula (5):

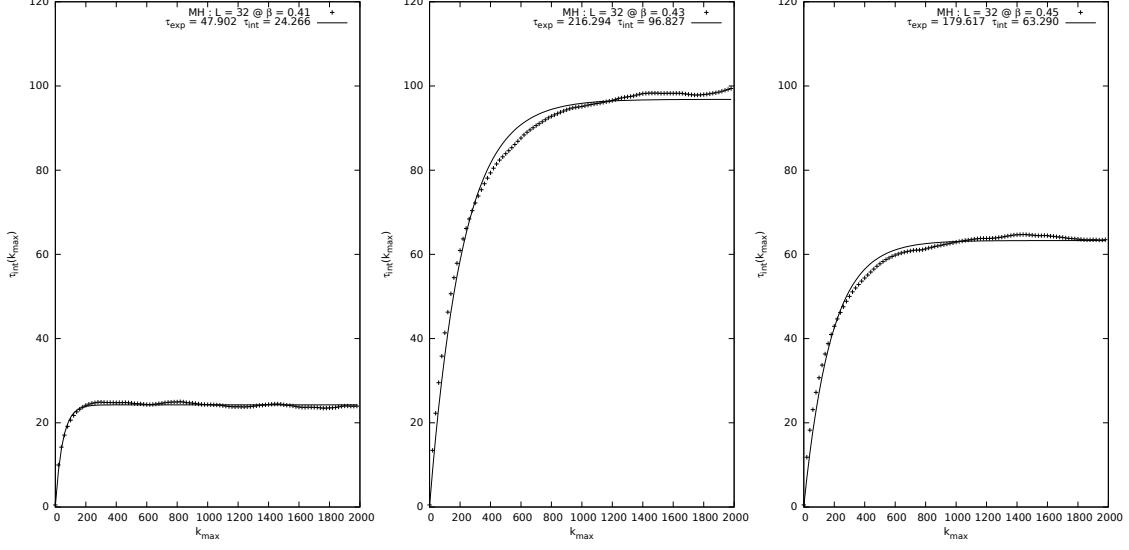


Figure 5. Plot of the integrated autocorrelation time near the phase transition for the MH algorithm. The raw data was obtained from a simulation of 10^6 measurements on a lattice 32×32 . The fit in the center is not very accurate because near β_c the autocorrelation time becomes very large and the interval $[0, 2000]$ becomes too small for a reliable estimate of τ_{int} .

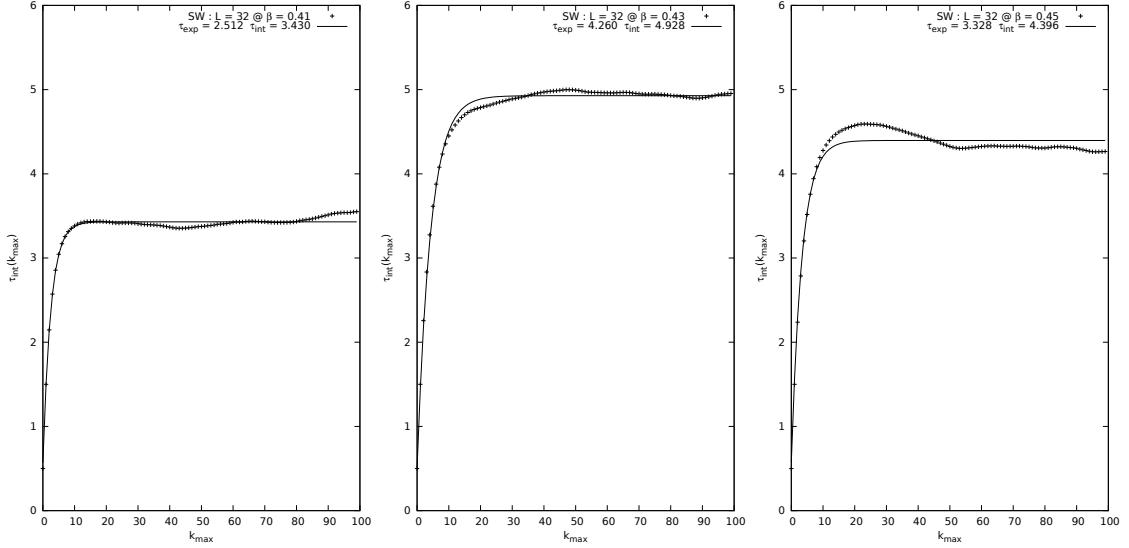


Figure 6. Plot of the integrated autocorrelation time near the phase transition for the SW algorithm. The raw data was obtained from a simulation of 10^5 measurements on a lattice 32×32 .

The data is fitted using the function:

$$\tau_{\text{int}}(k_{\text{max}}) = \tau_{\text{int}} \left[1 - \frac{2 \tau_{\text{exp}}}{2 \tau_{\text{exp}} + 1} e^{-k_{\text{max}}/\tau_{\text{exp}}} \right] \quad (8)$$

which is the exact solution for a bivariate gaussian process.

As we can see, near the phase transition the autocorrelation time for the MH process becomes very large compared to the autocorrelation time of the SW process (Fig.7). In fact due to the

critical slowing down of the MH algorithm, we have $\tau_{\text{int}}^{\text{MH}} \sim 100$ compared to an autocorrelation time $\tau_{\text{int}}^{\text{SW}} \sim 5$ for SW.

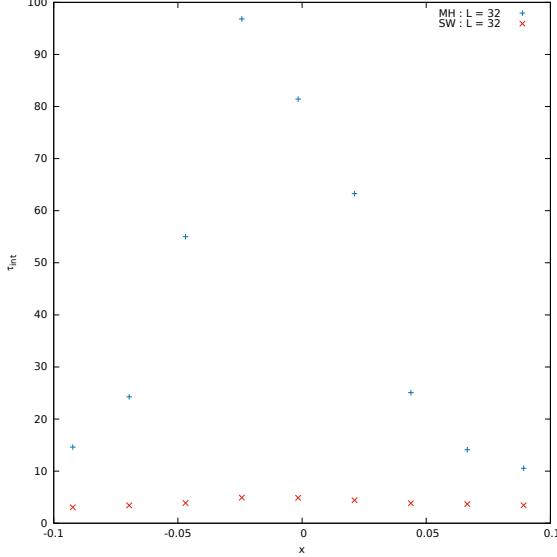


Figure 7. Plot of the integrated autocorrelation time as a function of $x = \frac{\beta - \beta_c}{\beta_c}$ for $L = 32$.

We remark that the peak is slightly off centered because of the finite size of the lattice (pseudocritical point $\beta_c(V) < \beta_c(\infty)$).

1.3 Binning Analysis

In order to have a statistical ensemble of uncorrelated data we need to address the issue of autocorrelation in the Markov process. An easy and efficient way to do so is to split the data in non-overlapping blocks of the same size and then average over each block to obtain (almost)uncorrelated data for the observable \mathcal{O} of interest. The blocks are also called *bins* and this method is called binning.

Consider an ensemble of $N = N_B \cdot k$ samples divided in N_B blocks of size k . For each block n we take the average of the observables \mathcal{O}_i it contains and then obtain a block-observable $\mathcal{O}_{B,n}$:

$$\mathcal{O}_{B,n} = \frac{1}{k} \sum_{i=0}^{k-1} \mathcal{O}_{nk+i} \quad n = 0, \dots, N_B - 1 \quad (9)$$

From a simple calculation we obtain that the error estimate on the mean value is:

$$\epsilon_{\mathcal{O}}^2 \equiv \sigma_{\mathcal{O}}^2 = \frac{\sigma_B^2}{N_B} = 2 \tau_{\mathcal{O},\text{int}} \frac{\sigma_{\mathcal{O}_i}^2}{N} \quad (10)$$

hence:

$$2 \tau_{\mathcal{O},\text{int}} = k \sigma_B^2 / \sigma_{\mathcal{O}_i}^2 \quad (11)$$

We study the dependence of the variance σ_B^2 on the block size k . The observable we consider is the energy density e .

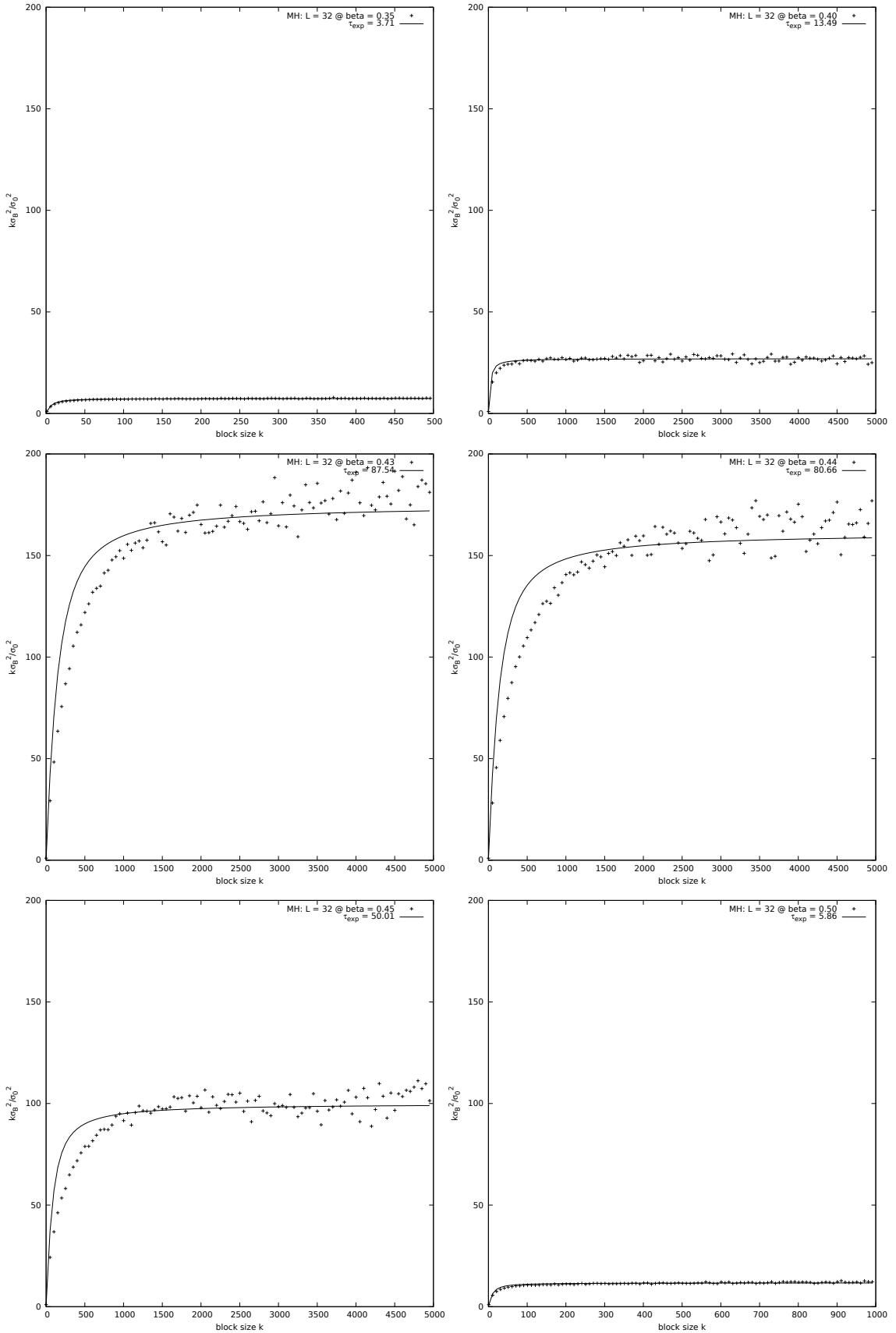


Figure 8. Binning analysis for the MH algorithm at various values of β ($L = 32$).

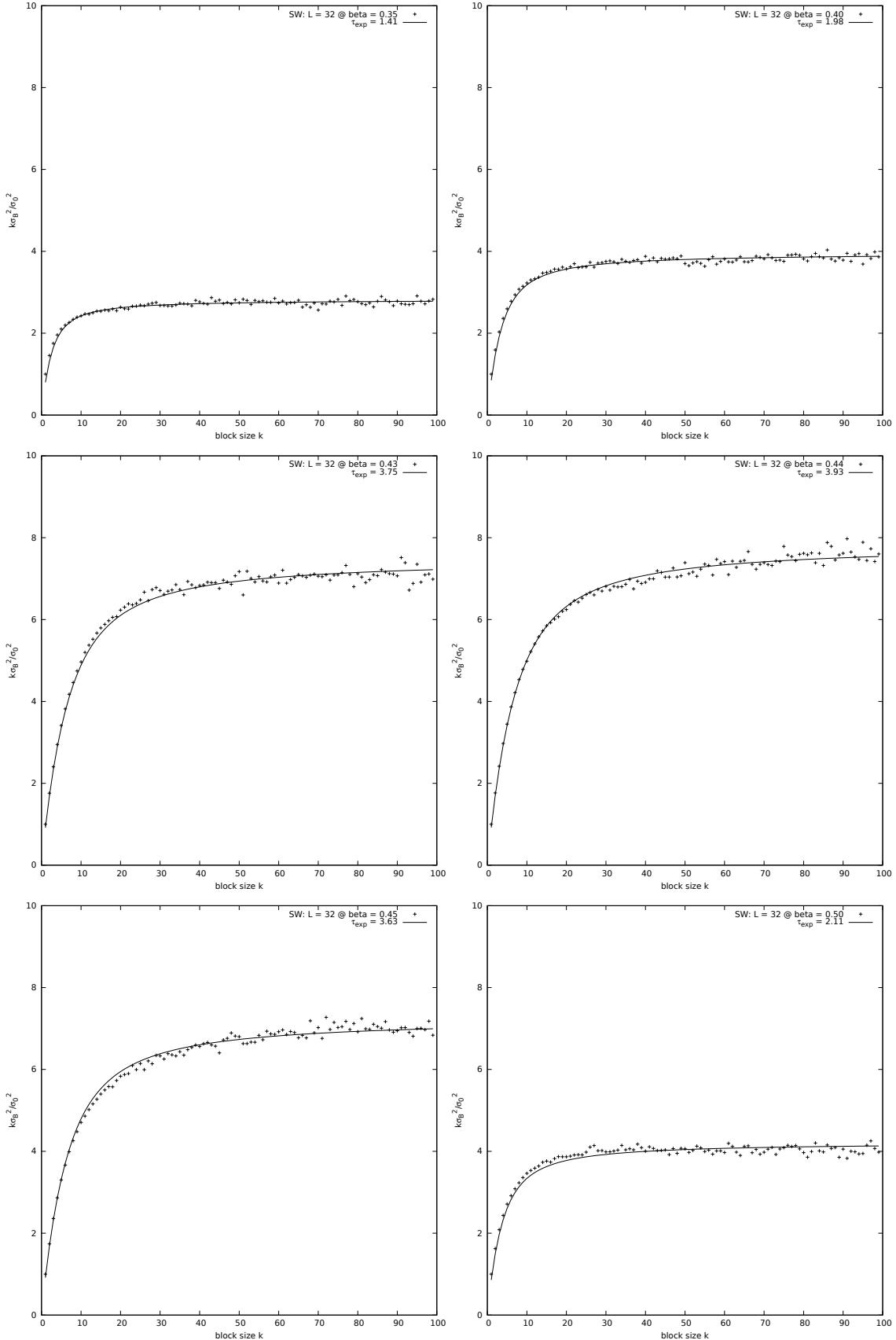


Figure 9. Binning analysis for the SW algorithm at various values of β ($L = 32$).

Where we fitted the data using the formula:

$$k \sigma_B^2 \approx 2 \tau_{\text{exp}} \left(1 - \frac{\tau_{\text{exp}}}{k} (1 - e^{-k/\tau_{\text{exp}}}) \right) \quad (12)$$

The lowest value of k for which we can consider the binned data to be uncorrelated is obtained by looking at the point for which the signal for $k \sigma_B^2 / \sigma_{\mathcal{O}_i}^2$ stabilizes and its first derivative tends to zero. We call this point $k^*(\beta)$. Since τ_{exp} has a local maximum at the critical point, we take $k^*(\beta_c)$ as the block size for the following simulations.

With the block sizes set to:

$$\begin{aligned} k_{\text{MH}}^* &= 1000 \\ k_{\text{SW}}^* &= 50 \end{aligned}$$

we should now able to sample data without having to worry about autocorrelation in the signal.

1.4 Observables

The main observables of interest for this system are the energy density e and the magnetization m :

$$e = E/V, \quad E = \langle \mathcal{H} \rangle, \quad V = L^d \quad (13)$$

$$m = M/V, \quad M = \langle \mathcal{M} \rangle, \quad \mathcal{M} = \left| \sum_i \sigma_i \right| \quad (14)$$

Thanks to the analytical solution of the 2d Ising model, first obtained by Onsager, we are able to compare the estimators computed by numerical simulations with their exact values obtained analytically.

The samples and the errors are collected employing the binning procedure of the previous section. For each inverse temperature we compute the estimate of the generic observable \mathcal{O} from

10^5 measurements for MH and 10^4 for SW.

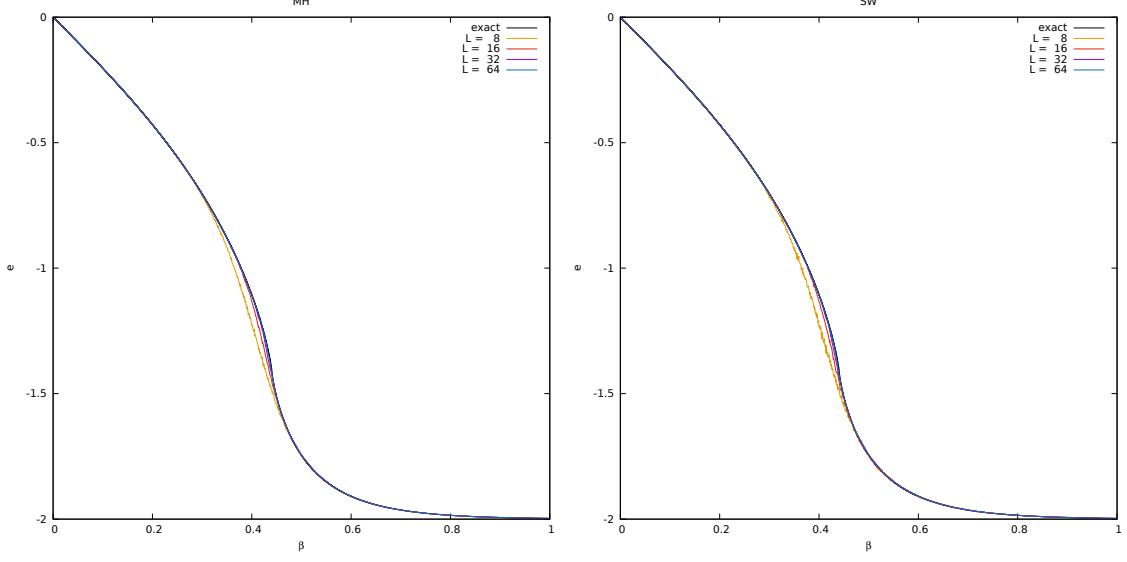


Figure 10. Energy density as a function of β . MH on the left and SW on the right.

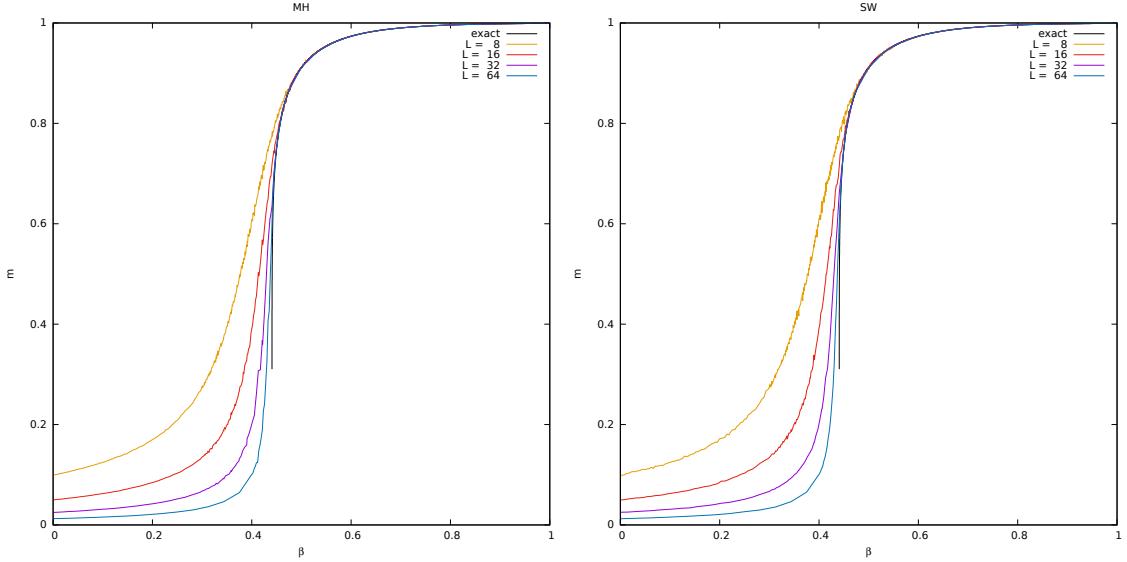


Figure 11. Magnetization as a function of β . MH on the left and SW on the right.

We find that the simulation data are remarkably close to the exact solution both for the energy density and for the magnetization especially for large sizes of the lattice. The only departure from Onsager's solution is in the paramagnetic phase of the magnetization plot: there we observe that the exact solution goes abruptly to zero before the critical point, while the numerical data seems to interpolate a smooth function. This is imputable to the finite size of the lattice used for the simulation. It is indeed a well known fact that discontinuities and divergences only appear in the thermodynamic limit of infinite size, which is precisely the premise of Onsager's solution. In fact, the numerical solution approaches more and more the exact one as the size of the lattice increases.

We also plot the heat capacity:

$$C = \frac{\beta^2}{V} \langle (\mathcal{H} - \langle \mathcal{H} \rangle)^2 \rangle \quad (15)$$

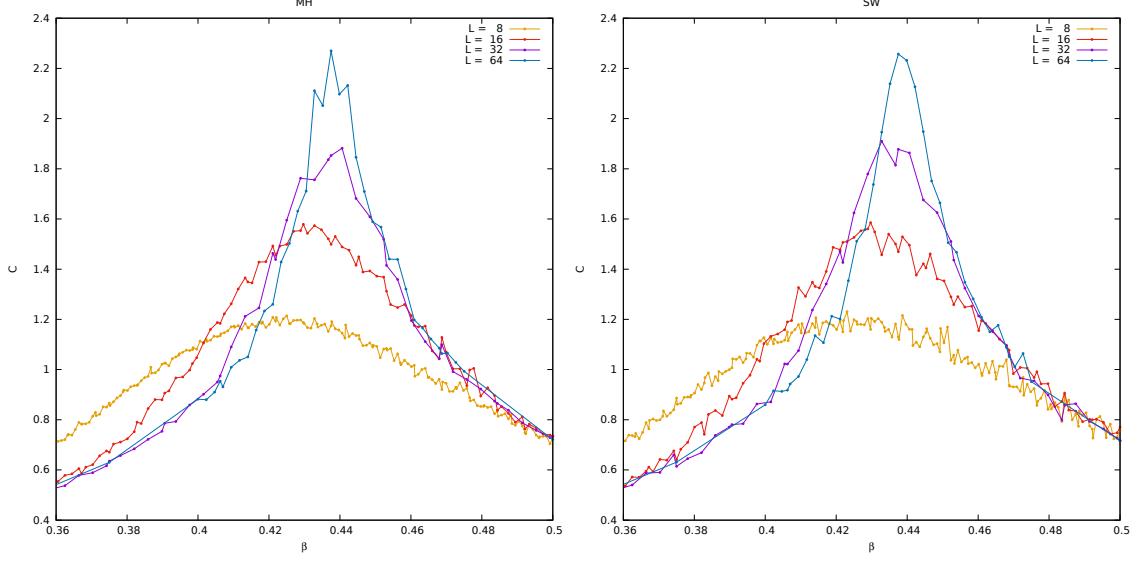


Figure 12. Heat capacity as a function of β . MH on the left and SW on the right.

and the magnetic susceptibility:

$$\chi = \frac{\beta}{V} \langle (\mathcal{M} - \langle \mathcal{M} \rangle)^2 \rangle \quad (16)$$

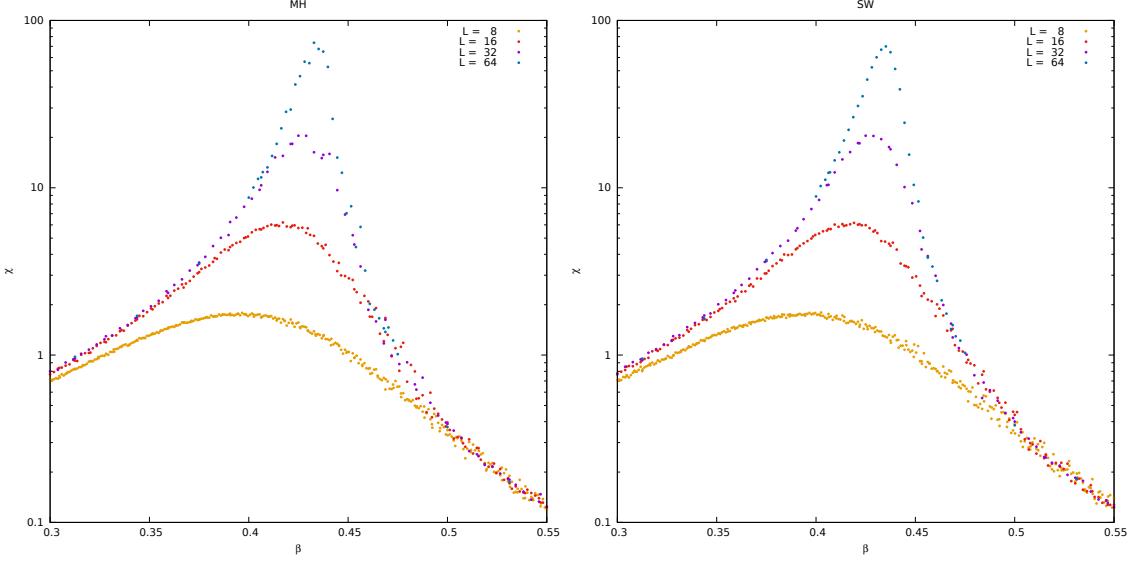


Figure 13. Magnetic susceptibility as a function of β . MH on the left and SW on the right.

As we immediately see from (Fig.12) and (Fig.13) the peak of the curve is progressively shifted to the infinite-size limit value β_c as we approach larger and larger sizes of the lattice:

$$\lim_{L \rightarrow \infty} \beta_{\max}(L) = \beta_c = \ln(1 + \sqrt{2})/2 \quad (17)$$

1.5 β Critical

From a power law fit of the points near the peak of each dataset we obtain estimates for the pseudocritical β -values at finite sizes $L = 8, 16, 32, 64$.

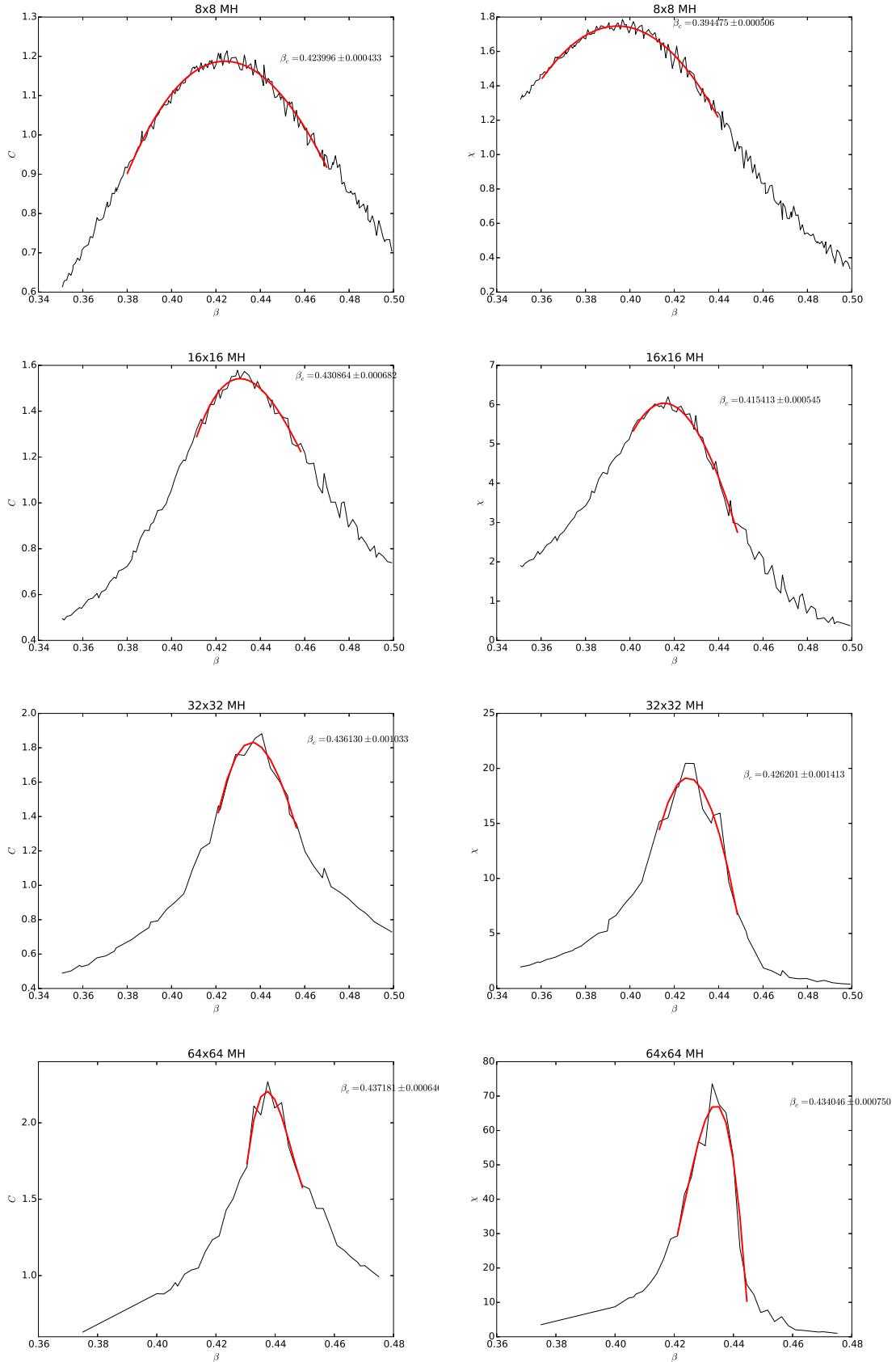


Figure 14. Fit of the heat capacity and susceptibility peaks (MH).

L	$\beta_{\max}(C)$	$\beta_{\max}(\chi)$
8	0.42400 ± 0.00043	0.39448 ± 0.00051
16	0.43086 ± 0.00086	0.41541 ± 0.00055
32	0.43613 ± 0.00103	0.42620 ± 0.00141
64	0.43718 ± 0.00065	0.43405 ± 0.00075

Table 1. Pseudocritical β -values obtained by polynomial fit of the peaks of C and χ (MH).

The polynomial function we used for the fit is:

$$f(\beta) = A + B(\beta - \beta_{\max})^2 + C(\beta - \beta_{\max})^3 \quad (18)$$

where we omitted the term of order one since it should be zero near the maximum of the function.

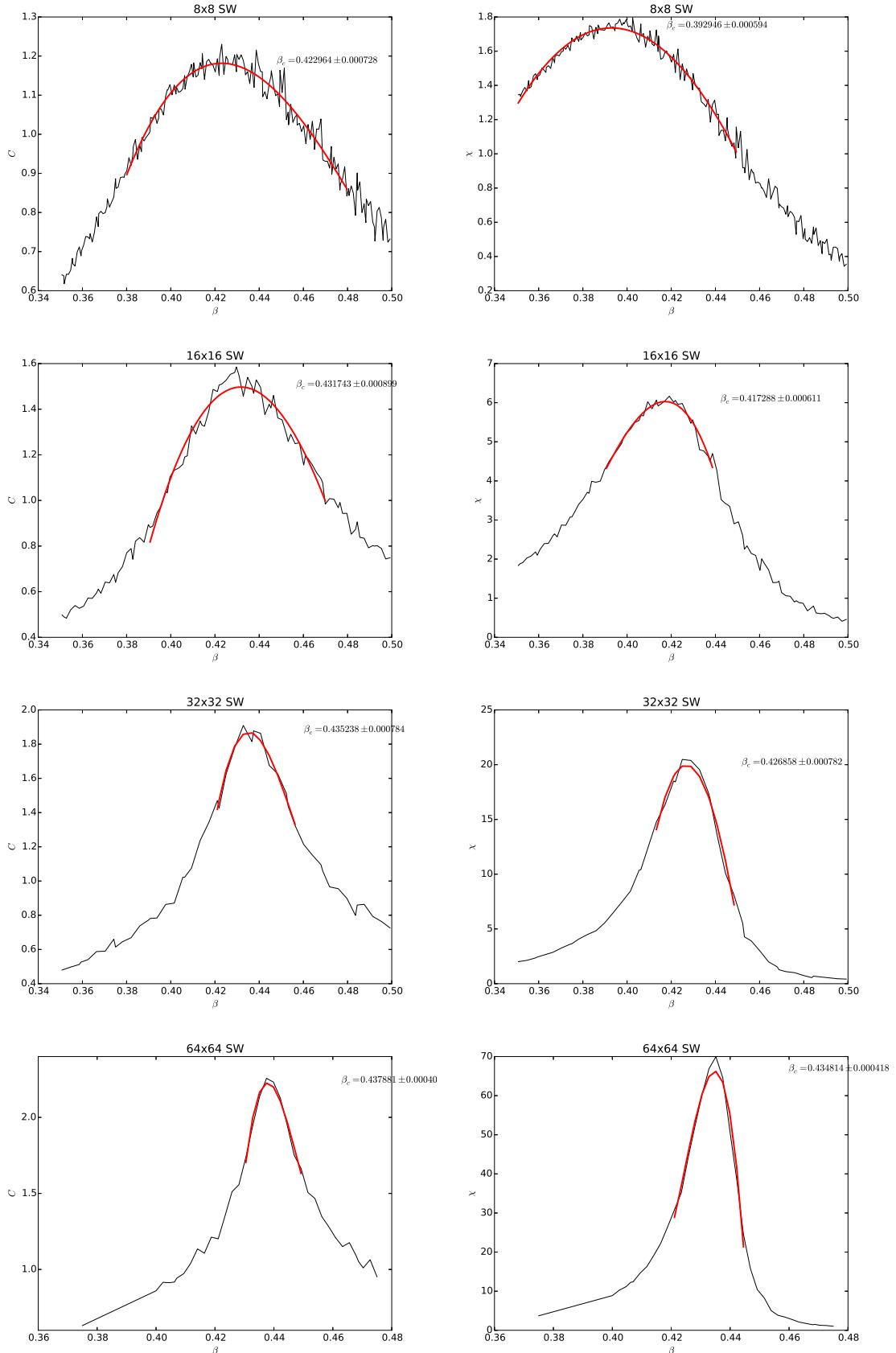


Figure 15. Fit of the heat capacity and susceptibility peaks (SW).

L	$\beta_{\max}(C)$	$\beta_{\max}(\chi)$
8	0.42296 ± 0.00073	0.39295 ± 0.00059
16	0.43174 ± 0.00090	0.41729 ± 0.00061
32	0.43524 ± 0.00078	0.42686 ± 0.00078
64	0.43788 ± 0.00040	0.43481 ± 0.00042

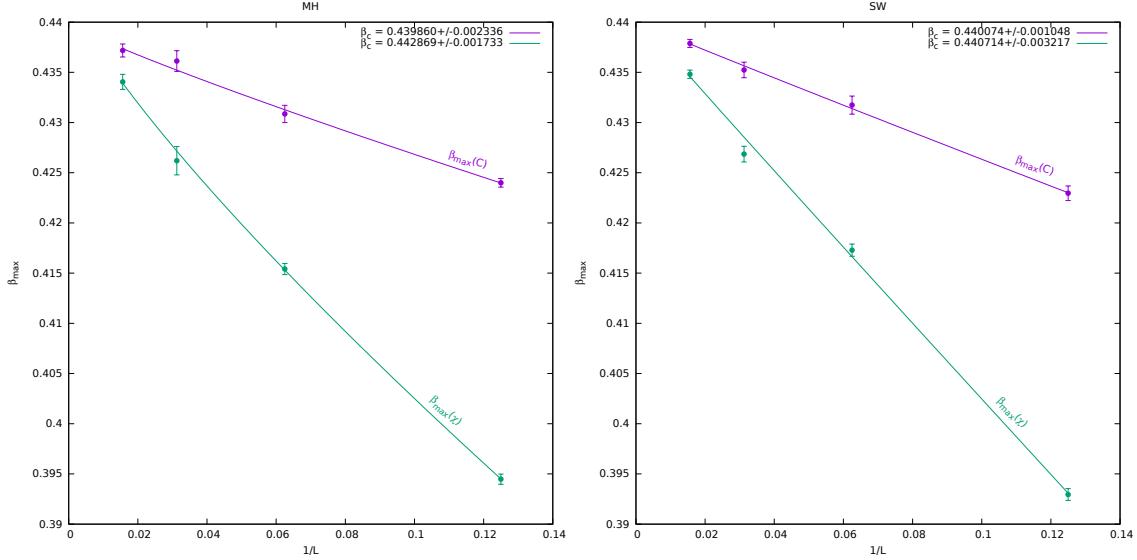
Table 2. Pseudocritical β -values obtained by polynomial fit of the peaks of C and χ (SW).

Remark: the estimates of the heat capacity and magnetic susceptibility of the MH algorithm are very accurate for small values of L where we employed a large dataset of 10^5 measurements and the autocorrelation time is relatively small. For large values of L the autocorrelation time grows very large compared to the size of the dataset and the estimates become less accurate.

The situation for the SW algorithm is different since the autocorrelation time in this case is always under control. However the smaller size of the dataset results in less accurate estimates especially at small lattice sizes where the MH algorithm is most efficient.

We can now obtain an estimate of β_c from a power law fit to the location of the maxima β_{\max} :

$$\beta_{\max} = \beta_c - c L^{-\nu} \quad (19)$$

**Figure 16.** Pseudocritical inverse temperature fit. MH on the left, SW on the right. The purple lines are relative to the heat capacity C while the green ones to the magnetic susceptibility χ .

The result of the fit are:

	$\beta_c(C)$	$\beta_c(\chi)$
MH	0.43986 ± 0.00234	0.44287 ± 0.00173
SW	0.44007 ± 0.00105	0.44071 ± 0.00322

These results are very close to the exact solution but still they could be greatly improved by taking more points especially at larger values of L for which C and χ are more peaked.

1.6 Probability Distribution Functions

We study the probability distribution of the magnetization for a lattice of size $L = 8$. The energy and magnetization levels of a discrete system are quantized. In particular, for an Ising model on a square lattice of size L^2 with PBC, there are exactly $L^2 + 1$ magnetization levels equidistributed in the interval $[-1, +1]$.

With this consideration one can construct the probability distribution function of the magnetization just by counting the number of samples in each level (for L not too large). In (Fig.17) and (Fig.18) we show the PDFs of m for both of the algorithms and for various values of β :

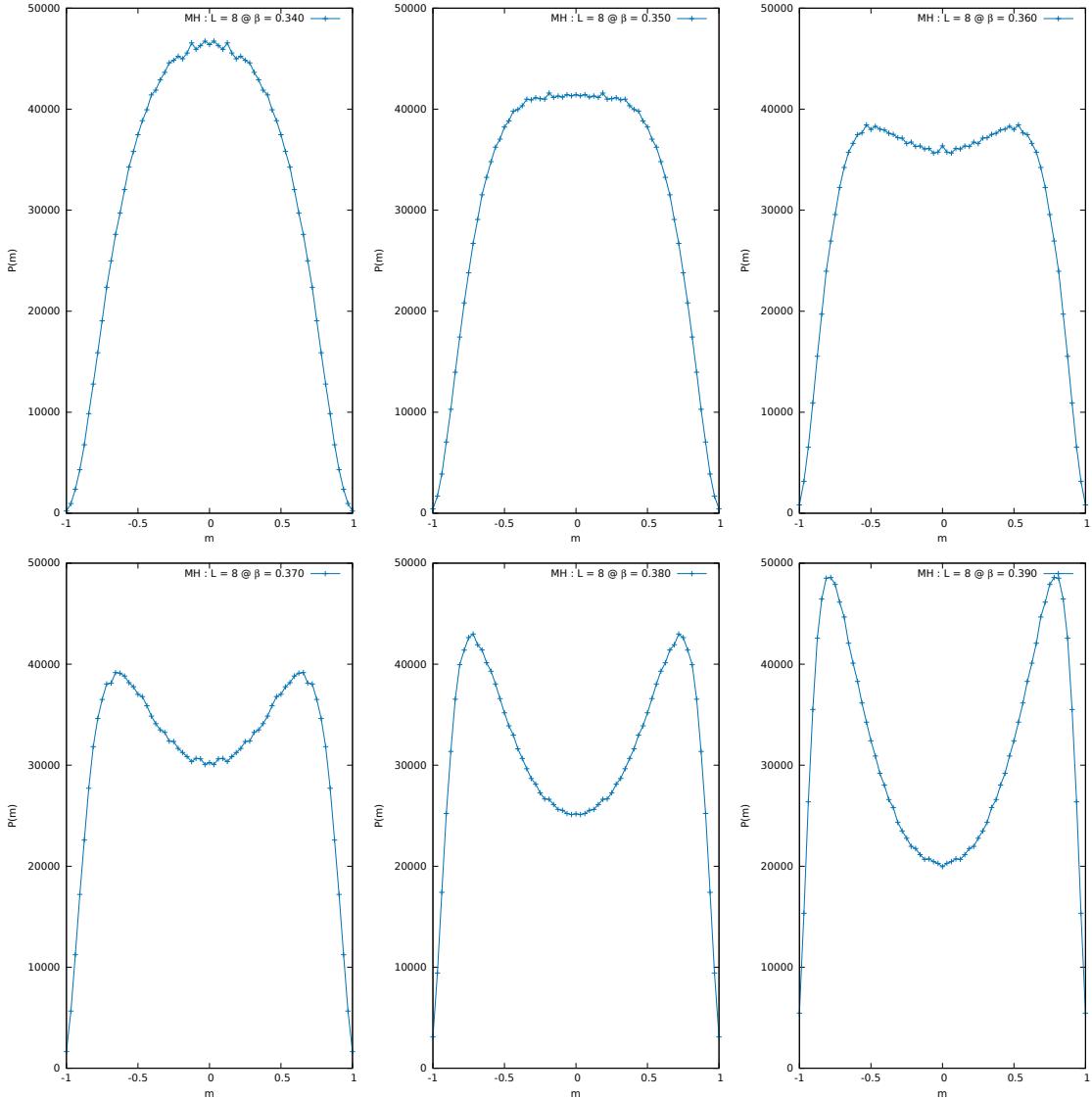


Figure 17. Probability distribution functions of m obtained from 10^6 sweeps of MH on a lattice 8×8 .

By looking at the PDFs of m we can witness the formation of a 2-fold degeneracy of the ground state of the system for $\beta > \beta_c$. Moreover, if we get close to the phase transition, we see that the distribution $P(m, \beta)$ becomes almost flat at $m=0$ before “splitting” into the sum of two PDFs for the two degenerate vacua. This implies that near β_c the distribution does not have a well defined second cumulant, which is precisely the magnetic susceptibility.

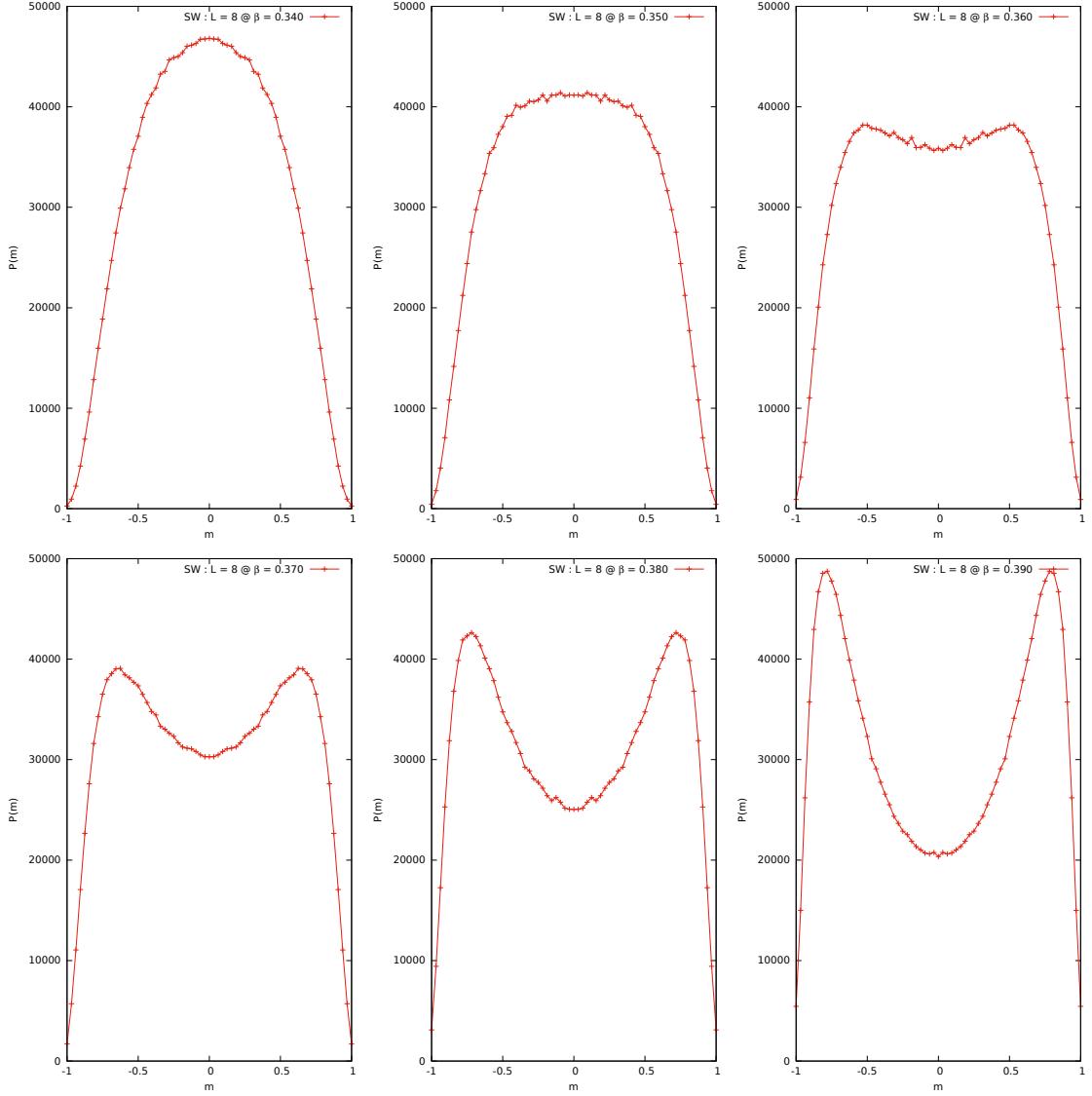


Figure 18. Probability distribution functions of m obtained from 10^6 sweeps of SW on a lattice 8×8 .

Since we used a cold start initialization, all our simulations of MH at temperatures lower than T_c spontaneously broke the \mathbb{Z}_2 symmetry of the ground state to some value $m_0 > 0$. Having used a hot start initialization instead, we would have ended up with $m_0 = \pm|m_0|$ with probability $\frac{1}{2}$ each. For the SW algorithm this is not an issue since it can easily tunnel between ground states by flipping large numbers of spins in a single sweep. In both cases, though, we measured the unsigned magnetization thus obtaining only histograms for the range $[0, 1]$. In (Fig.17) and (Fig.18) we computed the histograms for the negative range by assuming the \mathbb{Z}_2 symmetry:

$$P(-m) = P(m) \quad (20)$$

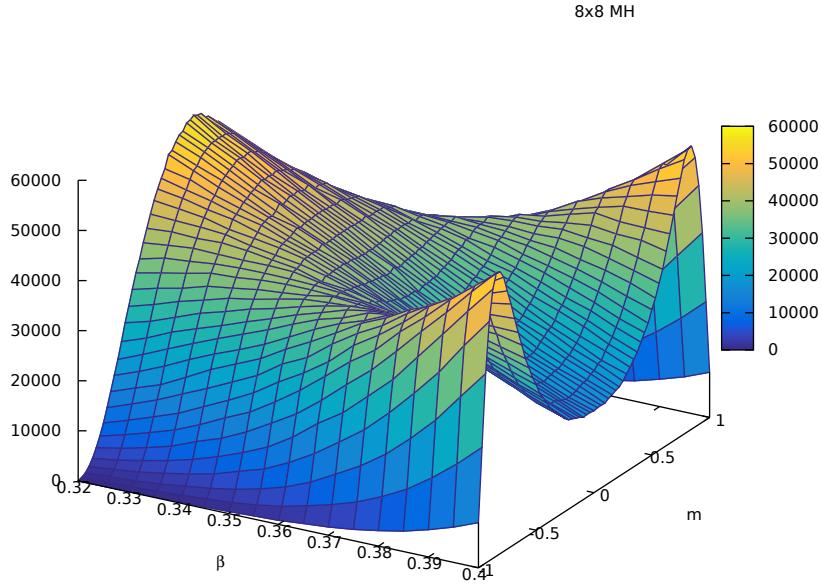


Figure 19. 3d plot of $P(m, \beta)$ as computed from 10^6 measurements of m obtained on a 8×8 lattice with the MH algorithm.

1.7 Spatial Correlations

The two-point correlation function is defined as:

$$G(\vec{r}_i - \vec{r}_j) = \langle \sigma_i \sigma_j \rangle \sim \exp(-|\vec{r}_i - \vec{r}_j|/\xi) \quad \text{for large } |\vec{r}_i - \vec{r}_j| \quad (21)$$

where ξ is the correlation length of the system:

$$\xi = -\lim_{|\vec{r}| \rightarrow \infty} (|\vec{r}| / \ln G(\vec{r})) \quad (22)$$

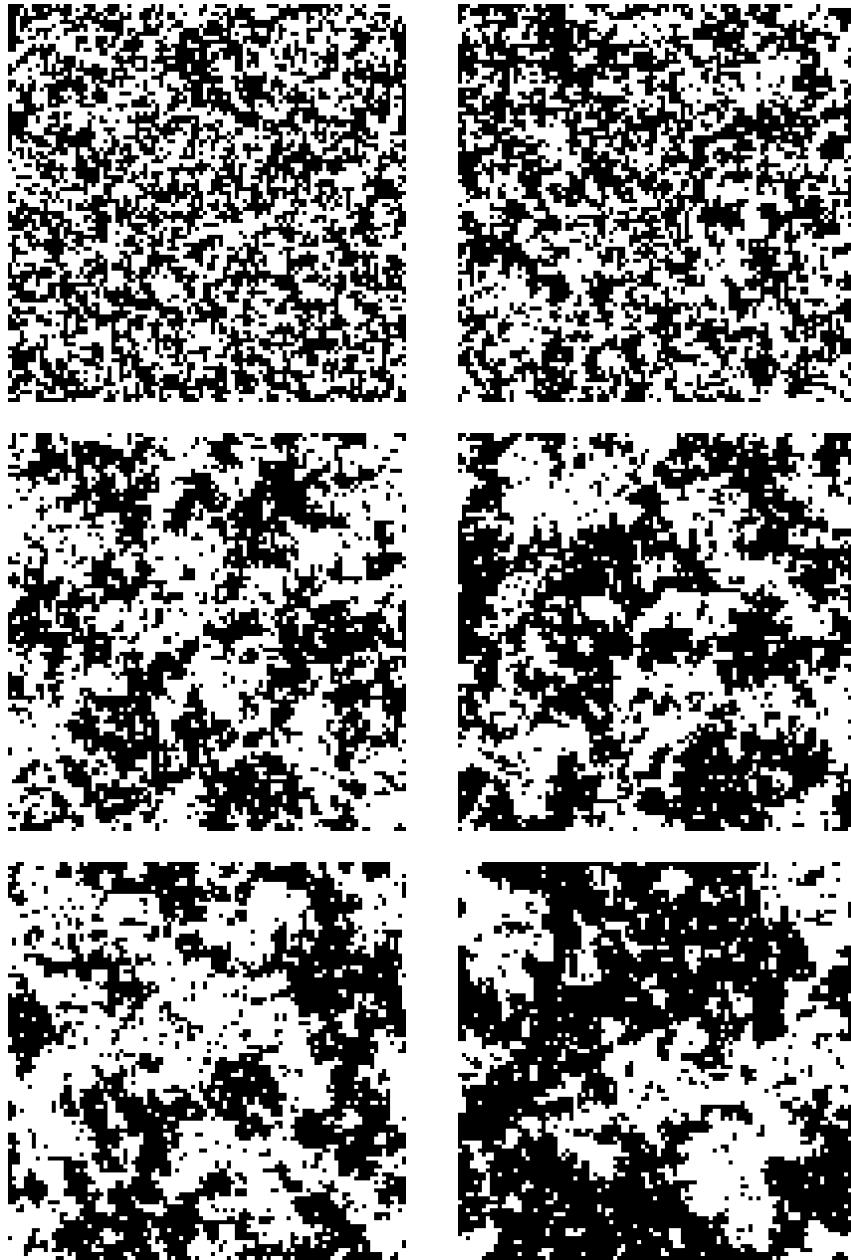


Figure 20. Illustration of the growth of spatial correlations when criticality is approached on a lattice 100×100 : $\beta = 0.22, 0.31, 0.37, 0.39, 0.42, 0.43$.

Since our system is finite and lives on a lattice, the rotational symmetry $\text{SO}(2)$ of \mathbb{R}^2 is broken down to the discrete subgroup \mathbb{Z}_4 of $\frac{\pi}{2}$ rotations. We take advantage of this and the translational symmetry of the system to compute an improved version of the correlation function $G(\vec{r})$ by defining the observables:

$$S_x \equiv \frac{1}{L} \sum_{y=1}^L \sigma(x, y) \quad (23)$$

$$S_y \equiv \frac{1}{L} \sum_{x=1}^L \sigma(x, y) \quad (24)$$

where x and y are the two orthogonal directions in the lattice.

We can compute the correlation function as:

$$G(r) = \frac{1}{2} \left(\frac{1}{L} \sum_{x=1}^L S_x S_{x+r} + \frac{1}{L} \sum_{y=1}^L S_y S_{y+r} \right) \quad (25)$$

Because of the periodic boundary conditions on the lattice we also have that:

$$\sigma(x+aL, y+bL) = \sigma(x, y) \quad \forall (a, b) \in \mathbb{Z}^2 \quad (26)$$

and therefore, the exponential form of G is modified to that of a hyperbolic cosine:

$$G(r) \sim \frac{1}{2} \left(e^{\frac{r}{\xi}} + e^{-\frac{r}{\xi}} \right) \sim \cosh \left(\frac{r - \frac{L}{2}}{\xi} \right) \quad (27)$$

In (Fig.21) we can see the exponential decay of the correlation function at large distances.

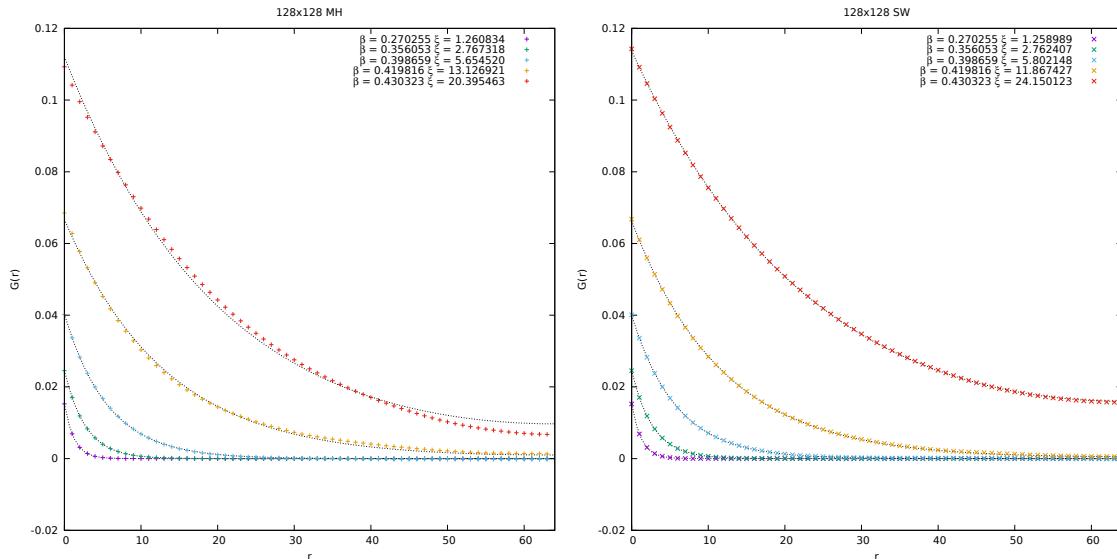


Figure 21. Correlation functions for various values of β . The lattice size used is 128×128 in order to reduce finite-size effects. As a precaution, we consider an interval in β -space such that the correlation length is much smaller than the size of the lattice (approximately one order of magnitude smaller).

The correlation length diverges at the critical point as:

$$\xi \sim A |x|^{-\nu} \quad (28)$$

where ν is an example of critical exponent of the model. We plot the data obtained from the previous fit using logarithmic scales on both axis. This way we can easily obtain the critical exponent ν from a linear fit:

$$\log(\xi) = \log(A) - \nu \log(|x|) \quad (29)$$

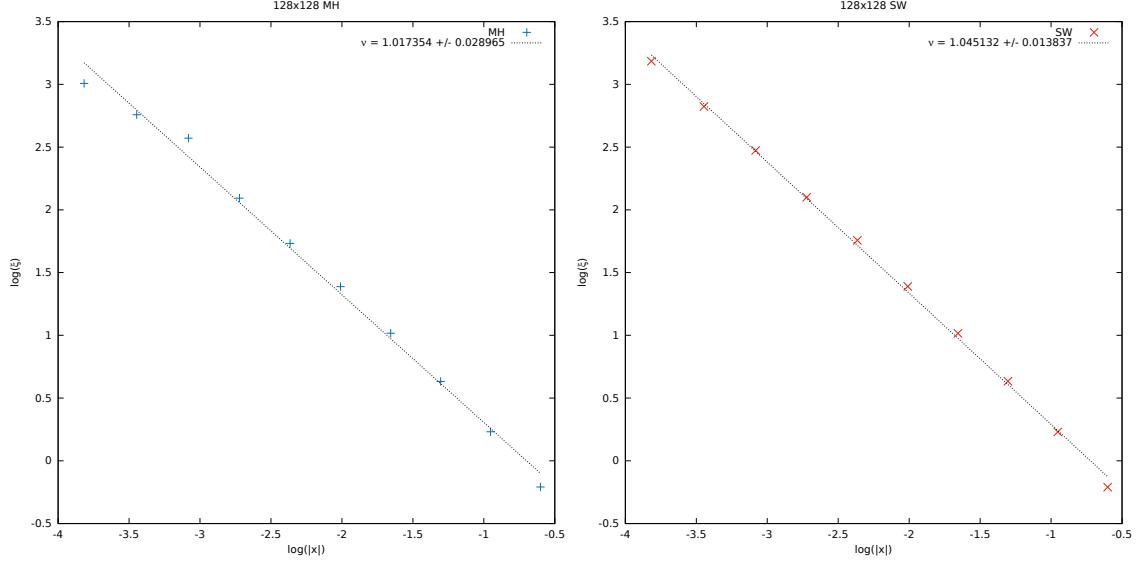


Figure 22. Plot of the correlation length dependence on the parameter $x = \frac{\beta - \beta_c}{\beta_c}$ for a lattice of size 128×128 . MH on the left and SW on the right. The errors are computed by jackknife binning of the fit data from (Fig.21) through (27).

The results are:

$$\nu_{\text{MH}} = 1.017 \pm 0.029 \quad (30)$$

$$\nu_{\text{SW}} = 1.045 \pm 0.014 \quad (31)$$

Both results are compatible with the known exact value $\nu = 1$.

1.8 Finite Size Scaling

The main results of the finite size scaling (FSS) study are usually estimates of the critical temperature and the critical exponents characterizing the universality class of the transition. In the infinite-volume limit most of the expectation values one can compute from e and m exhibit singularities at the transition point. In finite systems, though, the singularities are smeared out and the standard observables scale according to:

$$m = L^{-\beta/\nu} f_m(y) + \dots \quad (32)$$

$$C = C_{\text{reg}} + L^{\alpha/\nu} f_C(y) + \dots \quad (33)$$

$$\chi = L^{\gamma/\nu} f_\chi(y) + \dots \quad (34)$$

where α, β, γ and ν are the critical exponents, $f_i(x)$ are FSS functions and y is the scaling variable. Near the phase transition the correlation length $\xi \sim x^{-\nu}$ is the only length scale with which to describe the physics of the system, therefore the only adimensional quantity we can construct from it is:

$$y = L/\xi \sim L x^\nu = (L^{1/\nu} x)^\nu = L x \quad (35)$$

We now can remove any dependence on the scale of the system by rescaling the observables by the appropriate power of L given by the relative critical exponent. This way we expect each observable, expressed as functions of the scaling variable y , to collapse on the same curve $f_i(y)$ regardless of the value of L (at least in the vicinity of β_c).

ν	1
α	0
β	1/8
γ	7/4

Table 3. Exact critical exponents for the Ising model in 2d.

With a procedure completely analogous to that of (Sect.1.5), one could also obtain estimates for the critical exponents α, β, γ and ν by extrapolating the infinite size limit.

Remark: we observe that since the exact value of exponent α of the model is zero, the quantity C should, in principle, not scale with the lattice size. In reality C has a logarithmic divergence at the critical point, hence we adopt the scaling law:

$$C \sim \log(L^{1/\nu}) \quad (36)$$

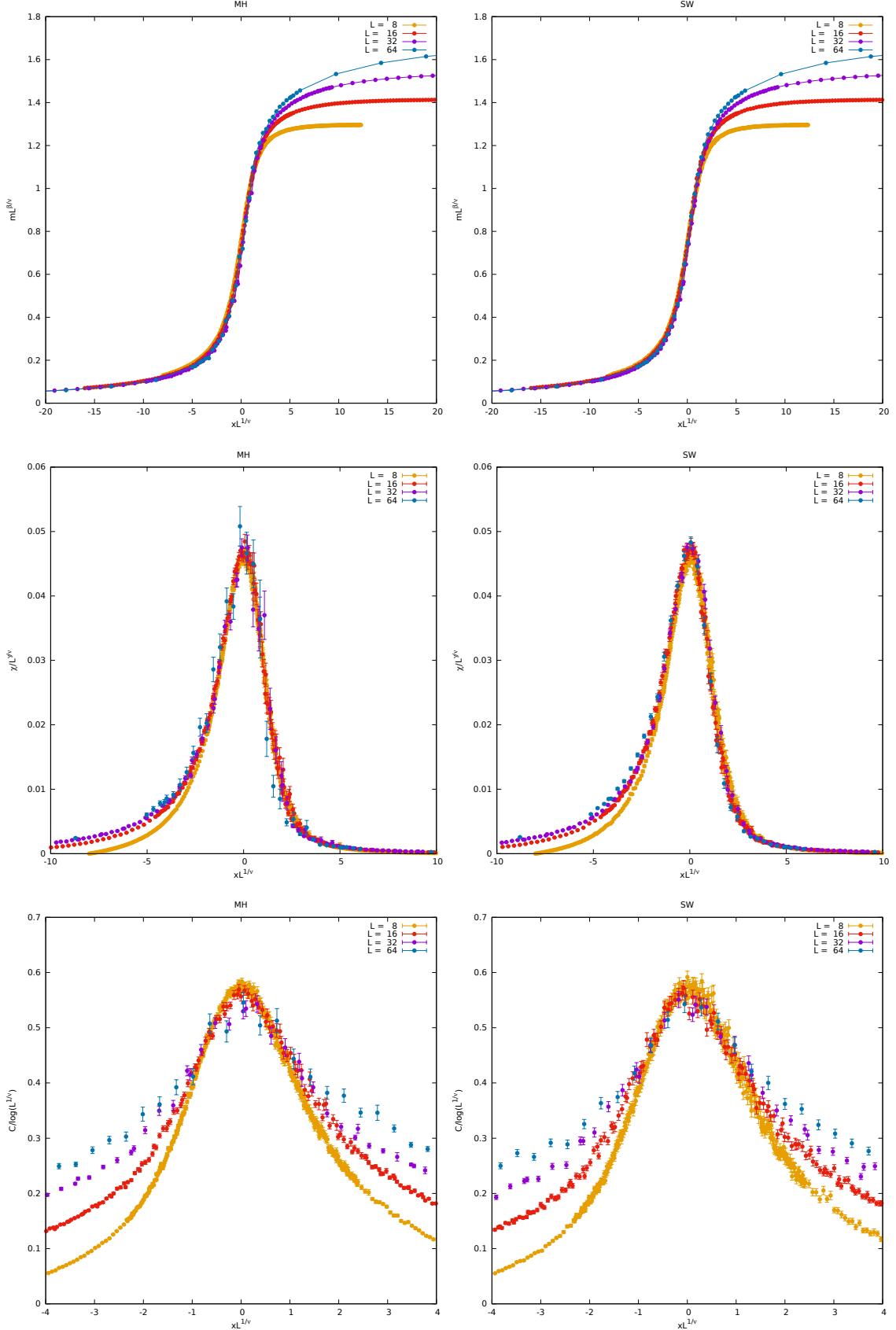


Figure 23. Finite Size Scaling study for the lattice sizes $8^2, 16^2, 32^2, 64^2$. Left MH, right SW. The top plot represents the scaling of the magnetization, the center one the magnetic susceptibility and the bottom one the heat capacity.