Monte Carlo methods for magnetic systems

Zoltán Néda

Babeş-Bolyai University

Dept of Theoretical and Computational Physics

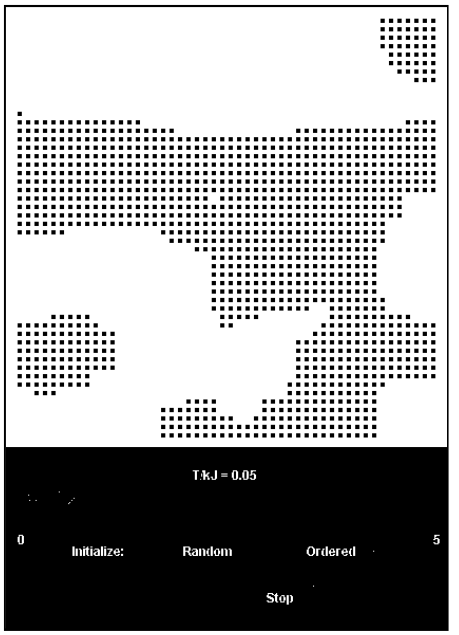
Cluj-Napoca, Romania

Main objective of the lecture:

To give an **introduction** for basic Monte Carlo methods in some simple models of magnetism.

European School on Magnetism, Timişoara, Sept. 1-11, 2009

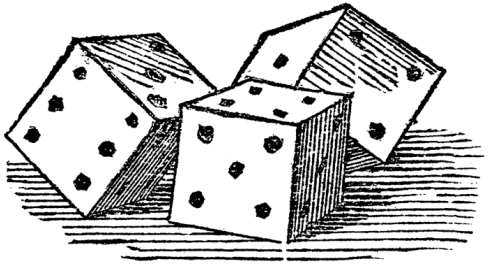
Syllabus

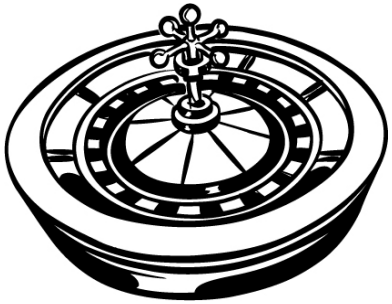


applet made by R. Sumi

- About Monte Carlo methods
- Deterministic versus stochastic simulation methods
- Elements of Stochastic Processes (Markov chains)
- Monte Carlo integration
- Theoretical approach to magnetic models
- What we are interested in?
- The Metropolis MC method for magnetic systems
- Implementing the Metropolis MC method for the 2D Ising model
- Finite-size effects
- Efficient MC techniques

What are Monte Carlo methods?



Computer simulation methods:

- **Molecular dynamics** (deterministic simulations, based on the integration of the equation of motion)
- **Monte Carlo methods** (Stochastic simulation techniques, where the random number generation plays a crucial role)

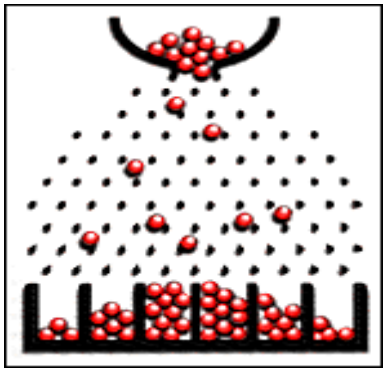
- **In general we speak about Monte Carlo simulation methods whenever the use of the random numbers are crucial in the algorithm!**

MC: the art of using pseudo random numbers

- **Monte Carlo techniques are widely used in studying models of**: statistical physics, soft condensed matter physics, material science, many-body problems, complex systems, fluid mechanics, biophysics, econo-physics, nonlinear phenomena, particle physics, heavy-ion physics, surface physics, neuroscience etc….

Deterministic versus stochastic simulations

the Galton table →



- used to exemplify the normal distribution

Molecular dynamics approach: integrating in time the equation of motion of the particles.

advantage → the realistic dynamics

disadvantage → slow even on supercomputers, only short time-scales or small systems can be simulated

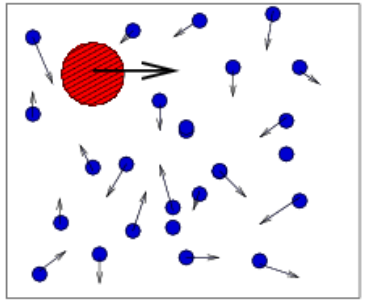
$$x(t) = x(t-1) + \square$$

Random number: 1 with p=1/2 and -1 with p=1/2

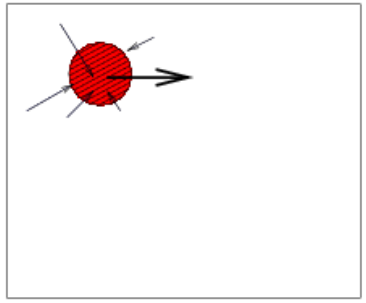
Monte Carlo approach: the result of many deterministic effects is handled as a stochastic (random) force.

advantage → fast, easy to implement

disadvantage → less realistic, many elements of the real phenomena are not in the model



Molecular dynamics

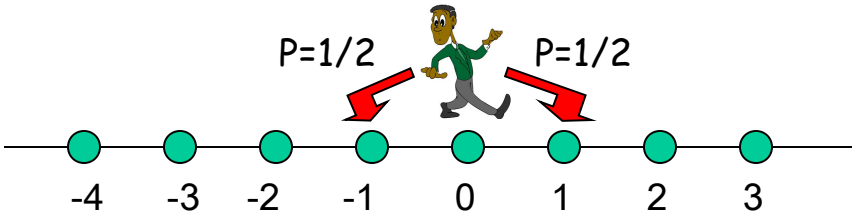


MC

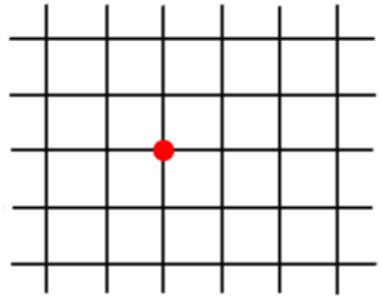
Some necessary elements of Stochastic Processes
Markov processes/ Markov chains

Stochastic process: let $\underline{x}$ label the element of any state-space. A process that randomly visits in time these possible $\underline{x}$ states is a stochastic process:

Example the 1D random walk:



2D random walk:



Markov processes (chain) are characterized by a lack of memory (i.e. the statistical properties of the immediate future are uniquely determined from the present, regardless of the past)

Example: random walk --> Markov process; self-avoiding walk is NOT a Markov process

Let $\underline{x}_i$ be the state of the stochastic system at step "i", a stochastic variable

The time- evolution of the system is described by a sequence of states: $\underline{x}_0$, $\underline{x}_1$, ..., $\underline{x}_n$, ...

The conditional probability that x_n is realized if previously we had: $\underline{x}_0$, $\underline{x}_1$, ..., $\underline{x}_{n-1}$: $P(\underline{x}_n \mid \underline{x}_{n-1},.....\underline{x}_0)$

Definition: For a Markov process we have: $P(\underline{x}_n \mid \underline{x}_{n-1}, \underline{x}_{n-2},..., \underline{x}_0) = P(\underline{x}_n \mid \underline{x}_{n-1})$

$$P(\underline{x}_0,..., \underline{x}_n) = P(\underline{x}_n \mid \underline{x}_{n-1}).P(\underline{x}_{n-1} \mid \underline{x}_{n-2})....P(\underline{x}_1, \underline{x}_0).a_0$$

$$P(\underline{x}_m, \underline{x}_j) = P(\underline{x}_m \to \underline{x}_j) = P_{m,j}$$

one-step transition probabilities, elements of the **stochastic matrix**

Definition: A probability distribution over the possible states (w_k) is called **invariant** or stationary for a given Markov chain if satisfy:

$$\begin{cases} w_m > 0; \\ \sum_m w_m = 1; \\ w_s = \sum_m w_m P_{ms} \end{cases}$$

$w_k \rightarrow$ The probability that $\underline{x}$=k during an infinitely long process

- A Markov chain is **irreducible** if and only if every state can be reached from every state! (the stochastic matrix is irreducible)

-A Markov chain is **aperiodic**, if all states are **aperiodic**. A state $\underline{x}$ has a period T>1 if $P_{ii}^{(n)}$=0 unless n=zT (z: integer), and T is the smallest integer with this property. A state is aperiodic if no such T>1 exist.

(Here we denoted by $P_{ik}^{(n)}$ the probability to get from state i to state k through n steps)

Definition: An irreducible and aperiodic Markov chain is **ergodic**

The **basic theorem for Markov processes**:

An ergodic Markov chain posses an invariant distribution w_k over the possible states

One dimensional Monte Carlo integration

Problem: given a function f(x), compute the integral: $I = \int_a^b f(x)dx$

The integral can be computed by choosing n points (x_i) randomly on the [a,b] interval, and with a uniform distribution : $\rho(x) = 1/(b-a) = Const$

$\rho(x)$ ⟶ probability density: $P(x, x+dx) = \rho(x)dx$ normalization: $\int_a^b \rho(x)dx = 1$

Straightforward sampling ⟶ $I = \langle f(x) \rangle \cdot (b-a) = \frac{b-a}{n} \sum_{i=1}^{n} f(x_i)$

The **strong law of large numbers** guarantees us that for a sufficiently large sample one can come arbitrary close to the desired integral!

Let $x_1, x_2, \ldots, x_n$ be random numbers selected according to a normalized probability density $\rho(x)$, then :

$$I = \int_{-\infty}^{\infty} f(x)\rho(x)dx;$$

$$P\left(\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} f(x_i) = I \right) = 1$$

(!) *the above affirmation is also true if the random numbers are correlated, or the interval is finite*

How rapidly the sum converge? --> for $\rho(x) = Const$ very badly!!!

Central limit theorem → the convergence improve if the shape of $\rho(x)$ approximates f(x) → we are sampling in the neighborhood where f(x) is big

Important sampling

The important sampling MC method will calculate the I integral by sampling on random points on the [a,b] interval according to a $\rho(x)$ distribution which approximates the shape of |f(x)|

If one generates n points, x_i ,according to an arbitrary $\rho(x)$

$$I = \int_a^b f(x)dx = \int_a^b \frac{f(x)}{\rho(x)}\rho(x)dx = \frac{1}{n}\sum_{i=1}^{N}\frac{f(x_i)}{\rho(x_i)}$$

the convergence is infinitely fast if $\rho(x) = |f(x)|$

Before getting to excited….→ one cannot simply choose $\rho(x) = |f(x)|$, since in this case one cannot normalize $\rho(x)$ (normalization of $\rho(x)$ is equivalent with the initial problem → one cannot generate thus random numbers simply according to the desired $\rho(x) = |f(x)|$ distribution

Theoretical approach to a magnetic ordering

Usually <u>canonic ensemble</u> is used → T, N, h is fixed

(T → temperature, h → external magnetic field, N→ particle number)

relevant energies

[internal interactions + interaction with external magnetic field, h, +kinetic terms]

↓

a deterministic effect → could favor ordering

↓

Hamiltonian, $H(\underline{x}_i,H)=E_i$

($\underline{x}_i$ → labels the microstates)

heat bath

↓

is a stochastic effect → favor randomness

↓

T or $\beta = 1/(kT)$

↓

Statistical thermodynamics approach

$$Z = \sum_i \exp\left(-\frac{E_i}{kT}\right)$$

$$F = -kT\ln(Z)$$

$$Z = \int_\Omega \exp\left(-\frac{H(x)}{kT}\right)dx$$

assuming that the density of state-space points is constant

F→ the free energy; k→ the Boltzmann constant

What are we interested in ?

The primary goal of the MC type simulations in magnetic systems is to estimate some averages at various T, h and N values

$\langle M \rangle$ ⟶ average magnetization

$\langle E \rangle$ ⟶ average energy

$\langle E^2 \rangle$ ⟶ average square energy

$\langle M^2 \rangle$ ⟶ average square magnetization

in canonical ensemble

$$\langle X \rangle = \frac{1}{Z} \sum_i X_i \exp(-\beta E_i)$$

a sum with huge number of terms (number of terms increasing exponentially with system size…ex: 2^N)…or very high dimensional integrals

$$\langle X \rangle = \frac{1}{Z} \int_\Omega X(x) \cdot \exp[-\beta H(x)] dx \qquad X \to M; M^2; E; E^2$$

the problem is that these sums cannot be usually analytically calculated → MC methods !!

exact enumeration is possible for small N (not of thermodynamic interest)

we are also interested in measurable quantities like: C_V ; χ

$$C_V = \left(\frac{\partial \langle E \rangle}{\partial T} \right)_{V,N} = \frac{1}{kT^2 N} \left(< E^2 > - < E >^2 \right)$$ ⟶ heat capacity at constant volume

$$\chi = \left(\frac{\partial \langle M \rangle}{\partial h} \right)_{H \to 0} = \frac{1}{NkT} (< M^2 > - < M >^2)$$ ⟶ susceptibility

The Metropolis MC method for magnetic systems

We want to compute integrals (sums) like:

$$< A >= \frac{1}{Z}\int_{\Omega} A(x)u[H(x)]dx$$

x_-->elements of the state-space

Ω--> the entire state-space

H(x)--> the Hamiltonian of the system

$$Z = \int_{\Omega} u[H(x)]dx$$

$$u[H(x)] = \exp[-\beta H(x)]$$

Very high dimensional integral which is exactly computable only for a limited number of problems!!!

Basic idea: to use the **important sampling** for calculating these integrals

IF in the MC integration we choose the states with probability $\rho(x)$:

$$\langle A \rangle = \frac{\int_{\Omega} A(x)\cdot\rho^{-1}(x)\cdot u[H(x)]\cdot\rho(x)\cdot dx}{\int_{\Omega}\rho^{-1}(x)\cdot u[H(x)]\cdot\rho(x)\cdot dx} = \frac{\sum_{i=1}^{n} A(x_i)\cdot\rho^{-1}(x_i)\cdot u[H(x_i)]}{\sum_{i=1}^{n}\rho^{-1}(x_i)\cdot u[H(x_i)]}$$

by choosing $\rho(x) = \frac{u[H(x)]}{Z}$ the sum converges rapidly and: $< A >= \frac{1}{n}\sum_{i=1}^{n} A(x_i)$

Problem: we still don't know Z!

The Metropolis et al. idea...

N. Metropolis et. al; J. Chem. Phys., vol. 21, 1087 (1953)

an algorithm has to be derived that generates states according to the desired $\rho(x)$!

Basic idea: using a Markov chain, such that starting from an initial state x_0 further states are generated which are ultimately distributed according to $\rho(x)$

$\rho(x)$

→ is an **invariant distribution** over the possible states

for this Markov chain need to specify the $P(x \to x')$ transition probabilities from state x to state x'. In order that the invariant limiting distribution be $\rho(x)$ we need:

- 1. The Markov chain should be ergodic (any state point should be reachable from any other state-point through the Markov chain)
- 2. For all possible x microstates: $\sum_{x' \in \Omega} P(x \to x') = 1$
- 3. For all possible x microstates: $\sum_{x' \in \Omega} P(x \to x')\rho(x') = \rho(x)$ (condition for the existence of the limiting distribution)

Instead of 2. and 3. a stronger but simpler condition can be used, the so called **detailed balance**:

$$P(x \to x')\rho(x) = P(x' \to x)\rho(x')$$

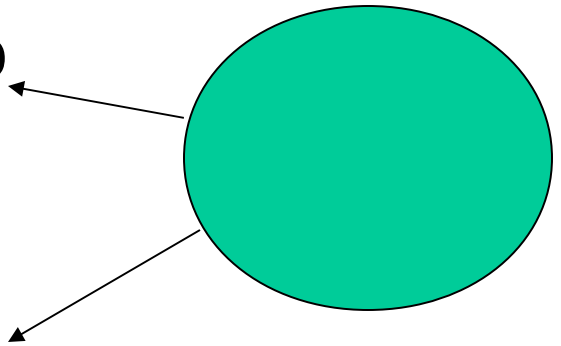
Result: We can construct Markov chains leading to the desired $\rho(x)$ distribution, without the prior knowledge of Z !!!

The Metropolis algorithm:

In the canonical ensemble: $u[H(x)] \propto \exp[-\frac{H(x)}{kT}]$

$$P(x \to x') = \begin{cases} \exp[-\beta \Delta E(x,x')] \cdots for \cdots \Delta E(x,x') > 0 \\ 1 \cdots for \cdots \Delta E(x,x') \le 0 \end{cases}$$

$$\Delta E(x,x') = H(x') - H(x)$$



another possibility (Glauber algorithm):

$$P(x \to x') = \exp\left\{-\frac{\Delta E(x,x')}{kT}\right\} \Big/ \left\{1 + \exp\left[-\frac{\Delta E(x,x')}{kT}\right]\right\}$$

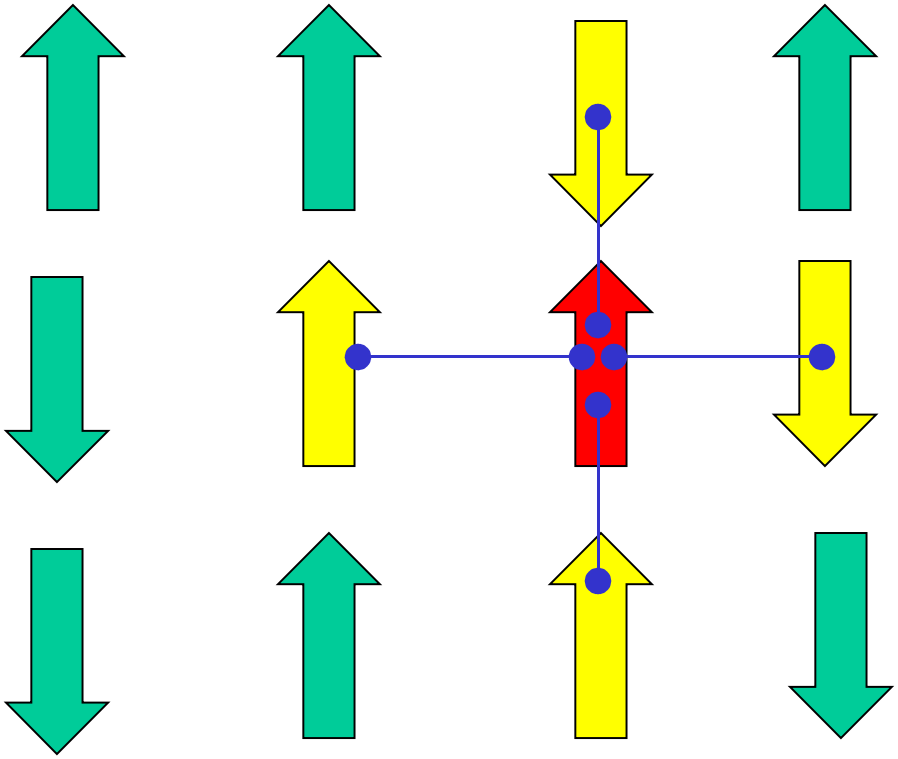
Algorithm for MC simulations:

1. Design an ergodic Markov process on the possible microstates (each state should be reachable from each other)
2. Specify an initial x microstate for starting
3. Choose randomly a new x' microstate (preferably so that $P(x \to x') > 0$)
4. Compute the value of $P = P(x \to x')$
5. Generate a uniformly distributed random number r between [0,1].
6. If $r \le P$ --> jump to the new state, and return to 3.
   If $r > P$ --> count the old state as new and return 3.
7. Average the quantity A for the generated states. Repeat steps 1-6 until the average converge

The Ising spin system

$$H = -J \sum_{<i,j>} \sigma_i \sigma_j - \mu h \sum_{\{i\}} \sigma_i$$

$$\sigma_i = \pm 1$$



- spontaneous magnetization is possible ($M \neq 0$ for $h=0$)
- first model for understanding ferro- and anti-ferromagnetism for localized spins
- for $J>0$ --> **ferromagnetic** order
- for $J<0$ --> **anti-ferromagnetic** order
- no phase transition in 1D
- ferro-paramagnetic phase transition for $D>1$
- second order phase transition (order-disorder)

- In 1D and 2D exactly solvable!
- Due to the local interactions calculating Z is difficult.
- exact solution very difficult in 2D
- no exact solution in 3D
- **Approximation methods**: mean-field theory, renormalization, high and low temperature expansion

Implementing the Metropolis MC for the 2D Ising model

Problem: Study m(T), <E(T)>, $C_v(T)$, $\chi(T)$ and T_c for 2D Ising model

We consider **h=0**, and fix **J=1**.

The temperature units are considered so that **k=1**.

Square lattice topology is considered

$$H = -\sum_{<i,j>} \sigma_i \sigma_j$$

$$\sigma_k = \pm 1$$

-Let us assume a lattice L x L with **free** boundary conditions

-We consider a canonical ensemble and fix thus **N** and **T**

We plan to calculate:

$$\square = \left\langle \frac{|M|}{N} \right\rangle = \left\langle \frac{\left|\sum_i \sigma_i\right|}{N} \right\rangle = \frac{\left\langle \left|\sum_i \sigma_i\right| \right\rangle}{N}$$

$$\square = \frac{1}{NkT}\left[< M^2(T) > - < M(T) >^2 \right]$$

$$\square = \langle H(\{\sigma_i\}) \rangle = \left\langle -\sum_{<i,j>} \sigma_i \sigma_j \right\rangle$$

$$\square = \frac{1}{Nk_B T^2}\left[< E^2(T) > - < E(T) >^2 \right]$$

→ from the maxima of $C_v(T)$ and $\chi(T)$

The Metropolis MC algorithm for the problem:

1. Fix a temperature (T)

2. Consider an initial spin configuration ($\{\sigma_i\}$). For example for all $i = 1, N \rightarrow \sigma_i = \pm 1$

3. Calculate the initial value of E and M

4. Consider a new spin configuration by virtually "flipping" one randomly selected spin

5. Calculate the energy E' of the new configuration, and the energy change ΔE due to this spin-flip

6. Calculate the Metropolis P=P(x-->x') probabilities for this change

7. Generate a random number "r" between 0 and 1

8. If $r \leq P$ accept the flip and update the value of the energy to E' and magnetization to M'

If $r > P$ reject the spin flip and take again the initial E and M values in the needed averages

9. Repeat the steps 4 - 8 many times (drive the system to the desired canonical distribution of the states)

10. Repeat the steps 4 -8 by collecting the values of E, E^2, M, M^2, for the needed averages

11. Compute this average for a large number of microstates

12. Calculate the value of m(T), <E(T)>, $C_v(T)$ and $\chi(T)$ using the given formulas

13. Change the temperature and repeat the algorithm for the new temperatures as well.

14. Construct the desired m(T), <E(T)>, $C_v(T)$, $\chi(T)$ curves

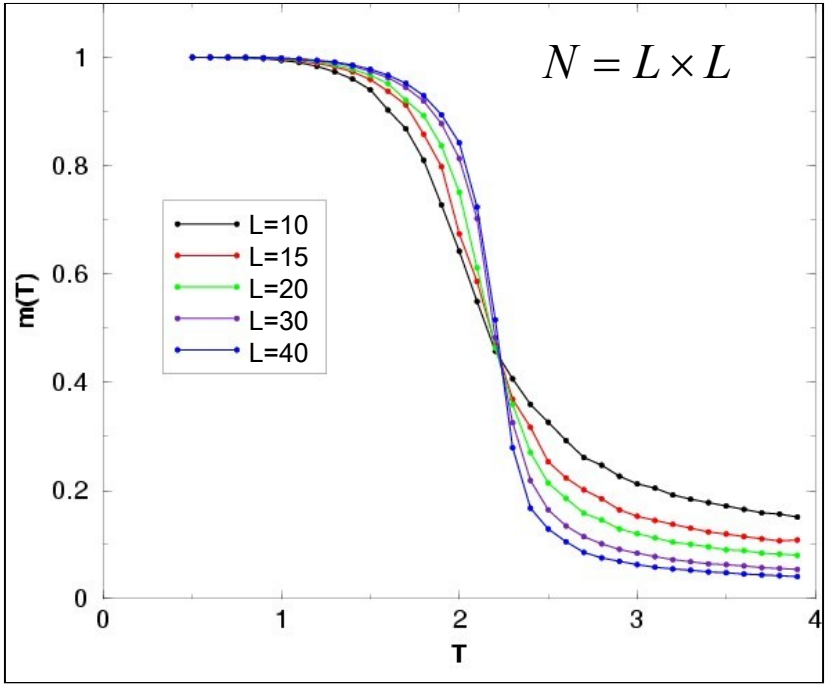
Finite-size effects

-The biggest problem with computer simulations is that it can be performed for relatively small systems (far from the ones needed in thermodynamics) $N << N_A$

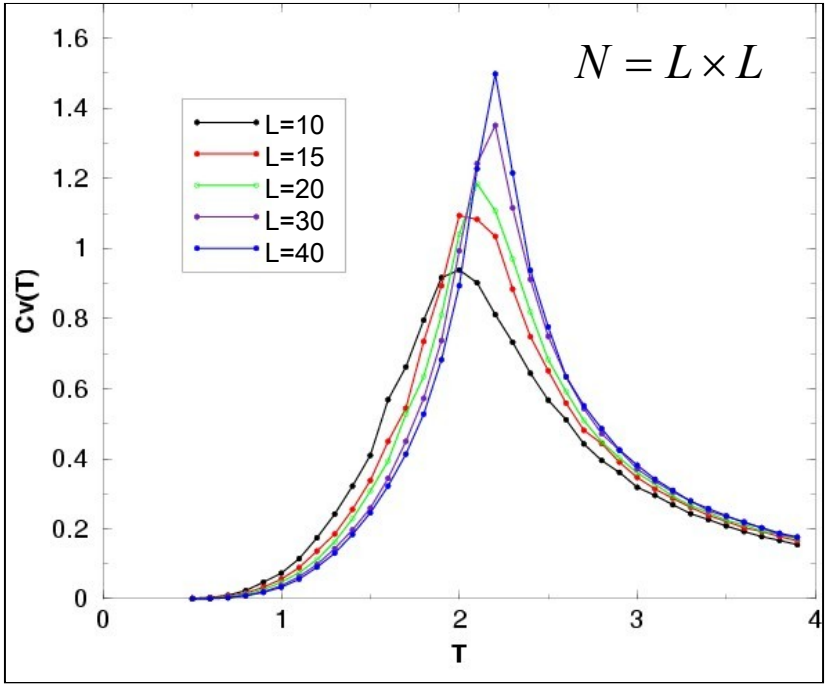
- Real phase-transition (real divergences in the thermodynamic quantities – derivatives of the thermodynamic potential) is possible only in infinite systems! In a finite-size system the correlation length cannot diverge and it is cut by the size of the system → instead of divergences rounded maximum or continuous behavior is obtained.

-The results obtained by MC simulations for finite systems has to be carefully evaluated and extrapolated for infinite systems! → finite size scaling is needed!

- Important quantities that have to be scaled: m(T), Cv(T), χ(T) curves and the value of T_c



The order parameter m(T) as a function of T for different system sizes



Specific heat C_v(T) as a function of T for various system sizes

Observations and technical points:

- the considered P($\underline{x}$-->$\underline{x}'$) transitions leads to an **ergodic Markov process**
- **one MC step** is defined as N spin flip trials !
- By applying the above algorithm for $T<T_c$ one can also follow how the order arises in the system. This dynamics might not necessarily be the "real one". **The Metropolis MC method is intended to yield equilibrium properties and not dynamical simulation of the system!**
- It is believed that the Glauber probabilities gives a realistic picture for the dynamics as well!

- One way of making the system quasi-infinite is to impose **periodic boundary conditions** (see the exercise in the computer codes!) → however this cuts also the correlation length
- The simple Metropolis and Glauber algorithm can be further improved, designing more clever and faster methods

Efficient MC techniques

I. At low temperatures $T << T_c$ the Metropolis and Glauber algorithm is inefficient. After equilibrium is reached (spins are ordered) most of the spin-flips are rejected, and computer time is wasted → very long simulations are needed to get a reasonable estimate for the averages. This drawback is eliminated by the BKL MC algorithm, see

A. B. Bortz, M.H. Kalos and J.L. Lebowitz, J. Comp. Phys. Vol. 17, 10 (1975)

II. In the neighborhood of T_c the Metropolis and Glauber algorithm is inefficient due to the critical slowing down →the relaxation time is linked to the correlation length by the **dynamical critical exponent**, z.

$\tau \sim \xi^z$

as T-->Tc we have $\xi \dashrightarrow \infty$ and get that $\tau \dashrightarrow \infty$

The big problem: for the Metropolis or the Glauber algorithm **z=2** !!! ---> There are many MC steps necessary to generate independent (uncorrelated configurations) --> the sampling is restricted only to a small part of the state-space **(**The system has a long memory). This problem is partially solved by flipping together clusters of correlated spins (cluster algorithms) see: U. Wolff, PRL vol. 62, 361 (1989); R.H. Swendsen and J-S. Wang, PRL vol. 58, 86 (1987)

III. Quantum-statistical models (Hubbard, Stoner, T-J, etc…) can be studied by Quantum MC methods, see: J. Tobochnik, G. Batrouni and H. Gould, Computers in Physics, vol. 6, 673 (1992)

IV. Frustrated, spin-glass type models (Edward-Anderson, Potts glass, etc…) can be studied also by MC methods. One of these is the simulated annealing method, see:
S. Kirckpatrick, G.D. Gelatt and M.P. Vecchi, Science vol. 220, 671 (1983)

Conclusions

-MC methods are powerful tools for numerically studying various models of magnetism.

-MC methods can be implemented on normal PC type computers, no supercomputers are needed.

- MC methods are easy to learn … however some basic programming experience is needed

-Mastering the MC method opens possibilities for studying many other models in solid-state physics, biophysics, ecology, economics, sociology, nuclear and medical physics, etc….

-The most cited paper in statistical physics is the paper of Metropolis et. al!