Advanced Computational Methods in Condensed Matter Physics

Lecture 10

Classical Monte-Carlo Methods

- Ising Model

Monte-Carlo methods

- Monte Carlo (MC) methods have been used for centuries *(statistical sampling)*.
- However, during World War II, this method was used to simulate the probabilistic issues with neutron diffusion (first real use).
- modern version of the Monte Carlo method invented in the late 1940s by Stanislaw Ulam, while working on nuclear weapons projects at the Los Alamos National Laboratory
- named by Nicholas Metropolis, after the Monte Carlo Casino, where Ulam's uncle often gambled

What is a MC method?

- *Non-Monte Carlo* methods typically involve ODE/PDE equations that describe the system.
- Monte Carlo method refers to any method that makes use of random numbers
  - Simulation of natural phenomena
  - Simulation of experimental apparatus
  - Numerical analysis
- Monte Carlo methods are stochastic techniques.
- It is based on the use of random numbers and probability statistics to simulate problems.

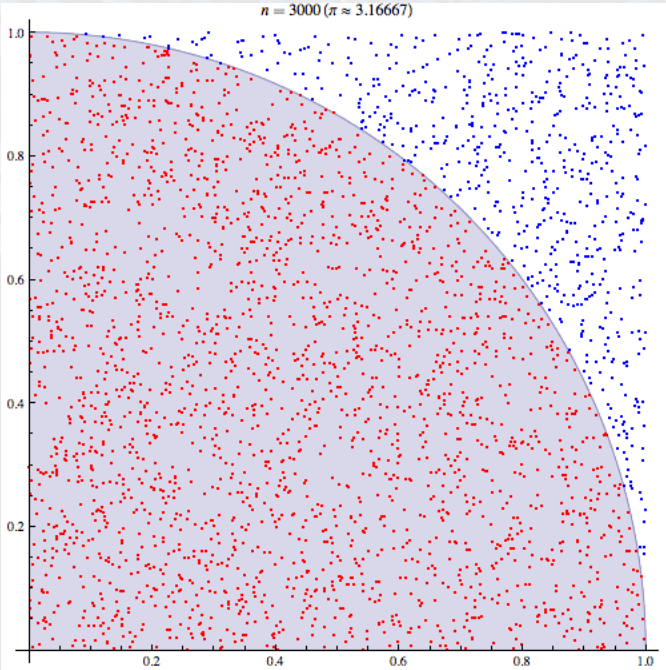
Why is MC used?

- It allows us to examine complex systems. And is usually easy to formulate (independent of the problem).
- For example, solving equations which describe two atom interactions. This would be doable without using Monte Carlo method. But solving the interactions for thousands of atoms using the same equations is impossible.
- However, the solutions are imprecise, and it can be very slow if higher precision is desired.

Simple Example 1: π

Consider a circle inscribed in a unit square: the circle and the square have a ratio of areas that is π/4 → the value of π can be approximated using a Monte Carlo method:

- Draw a square, then inscribe a circle within it
- Uniformly scatter some objects of uniform size (grains of rice or sand) over the square.
- Count the number of objects inside the circle and the total number of objects.
- The ratio of the two counts is an estimate of the ratio of the two areas, which is π/4. Multiply the result by 4 to estimate π.



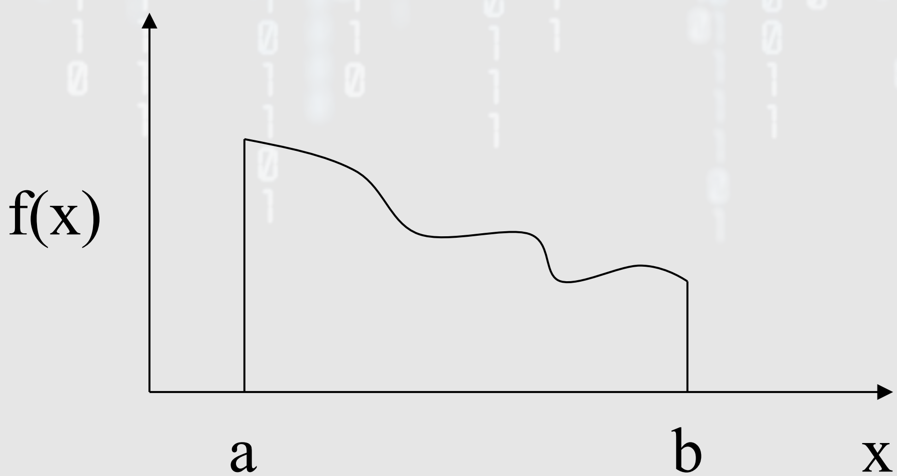
Simple Example 2: dice

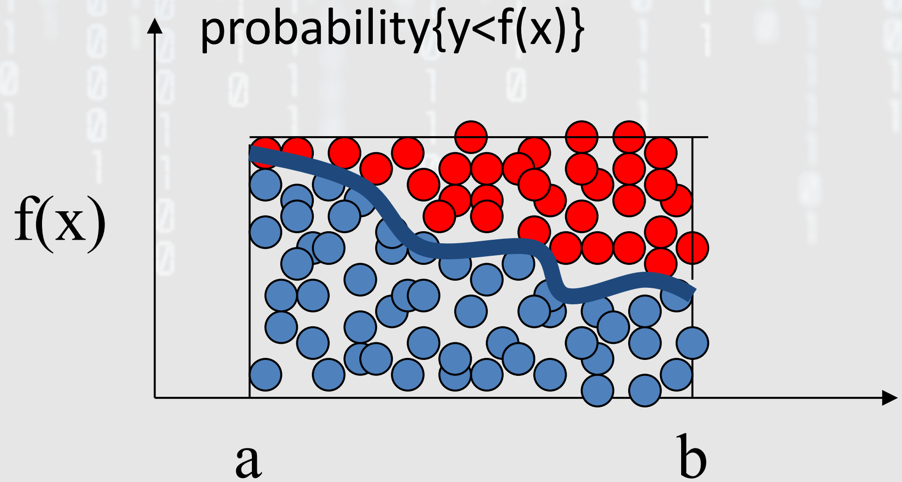
- **Problem**: What is the probability that 10 dice throws add up exactly to 32?
- **Exact Way.** Calculate this exactly by counting all possible ways of making 32 from 10 dice.
- **Approximate (Lazy) Way.** Simulate throwing the dice (say 500 times), count the number of times the results add up to 32, and divide this by 500.
- **Lazy Way can get quite close to the correct answer quite quickly.**

Simple Example 3: integration

- Method 1: Analytical Integration
- Method 2: Numerical, e.g., Quadratures
- Method 3: MC – random sampling the area enclosed by a<x<b and 0<y<max (f(x)): *hit and miss integration*

$$\int_a^b f(x)\,dx \approx \max(f(x))(b-a)\left(\frac{\#\textcolor{blue}{\bullet}}{\#\textcolor{blue}{\bullet}+\#\textcolor{red}{\bullet}}\right)$$





General case: integrate function over complicated region G

- Pick a simple (e.g. rectangular) region G'
- Sample N' random points over G'
- Count points in G: N

$$\frac{\mathrm{vol}(G)}{\mathrm{vol}(G')} = \frac{N}{N'}$$

Mean-value integration

Other way of MC integration is based on the mean-value theorem for a continuous function f(x) in x ∈[a,b], which states that a z∈[a,b] exists that:

$$\int_a^b \mathrm{d}x f(x) = f(z)(b-a)$$

where $f(z) \equiv \langle f \rangle$ is the mean or expectation value of $f(x)$.

Therefore:

$$\boxed{\frac{1}{b-a}\int_a^b \mathrm{d}x' f(x') \simeq \bar{f} \pm \sqrt{\frac{\overline{f^2} - \bar{f}^2}{N}}}$$

error, vanishes for N→∞

where we calculate from uniform random x_i ∈[a,b] :

$$\bar{f} = \frac{1}{N}\sum_{i=1}^{N} f(x_i) \qquad \overline{f^2} = \frac{1}{N}\sum_{i=1}^{N} f^2(x_i)$$

… more general

The expectation value is in general defined as

$$\langle f \rangle = \int \mathrm{d}x f(x) p(x)$$

where $x \in \mathbb{R}^d$ and p(x) is a pdf

A typical example is the calculation of the thermal expectation value in **statistical physics** where the pdf p(x) is given by the normalized BOLTZMANN distribution:

$$p(x) = \frac{1}{Z} \exp\left[-\frac{E(x)}{k_\mathrm{B} T}\right]$$

where E(x) is the energy, k_B the Boltzmann constant, T the temperature, and Z the canonical partition function (normalization)

Metropolis algorithm

The Metropolis algorithm is a more sophisticated method to produce random numbers from given distributions and is particularly useful to treat problems in statistical physics where thermodynamic expectation values of some observable O are of interest

$$\boxed{\langle O\rangle = \int \mathrm{d}x O(x) q(x)}$$

x is a set of parameters (e.g. coordinates and momenta of particle) and q(x) the Boltzmann distribution:

$$q(x) = \frac{p(x)}{Z} \qquad Z = \int \mathrm{d}x\, p(x)$$

The Metropolis algorithm now aims at generating sampling points x_i according to the pdf q(x), such that the expectation value can be evaluated.

rejection method

Suppose we already have n parameter sets: x_0, x_1,…, x_n={x_n} which follow q(x), and we need to decide if a new value x_t should be added:

The Metropolis method introduces the following acceptance probability for x_t:

$$\Pr(A|x_t, x_n) = \begin{cases} 1 & \text{if} \quad \dfrac{q(x_t)}{q(x_n)} \geq 1 \\ \dfrac{q(x_t)}{q(x_n)} & \text{otherwise.} \end{cases}$$

Hence, if $\Pr(A|x_t,x_n)=1$, we set $x_{n+1}=x_t$, and if $\Pr(A|x_t,x_n) < 1$, we draw a random number $r\in[0, 1]$ and accept x_t if $r<\Pr(A|x_t,x_n)$ and reject x_t otherwise.

or more compact: $$\Pr(A|x_t, x_n) = \min\left(\frac{p(x_t)}{p(x_n)}, 1\right) = p(x_t|x_n)$$

The underlying reason for this to work requires knowledge of stochastics in general and of MARKOV-chains in particular.

We apply this to the Ising model, which illustrates this at a concrete example

Example of a simulation process
(e.g. Ising model)

- Initialize the system
  – Put the system in a random state
- Make a trial move
  – Randomly make a trial move
- Calculate the energy change
  – Reevaluate the interactions of the moved particles with its neighbors and calculate the energy change
- Accept the trial move with the Metropolis scheme

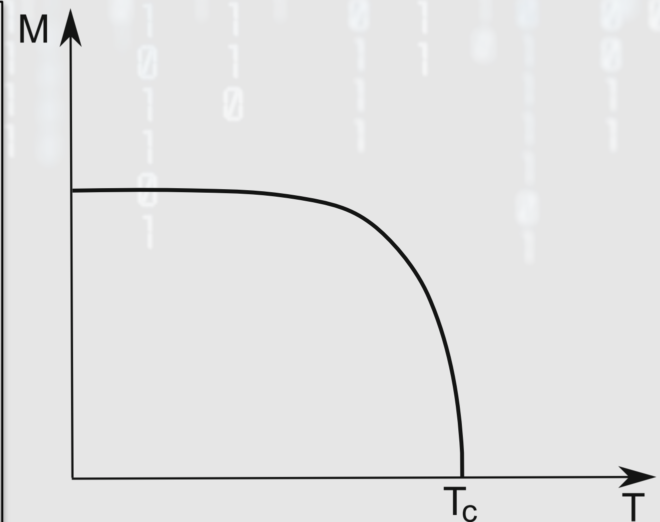
$$P = \begin{cases} \exp\left(-\frac{\Delta E}{k_B T}\right) & \Delta E > 0 \\ 1 & \Delta E < 0 \end{cases}$$

- Keep trying the moves until system approach equilibrium
  – Either monitor the total energy change, or monitor the structure formed in the simulation box
- Sampling
  – Sample a certain property over a certain number of configurations

Ising Model - Intro

- The Ising model describes a ferromagnetic (or antiferromagnetic) material.
- A ferromagnet (FM) has a finite magnetization M without applied magnetic field below the Curie temperature T_c.
- At T_c the FM has a second order phase transition to the paramagnetic state at $T>T_c$. M goes to zero at T_c and serves as an order parameter.

The microscopic origin of this macroscopic phenomenon is based on the exchange interaction between identical particles, the atoms or molecules forming the material. The exchange interaction is a ***purely quantum-mechanical effect*** which is a consequence of the *COULOMB interaction* in combination with the *PAULI exclusion principle.*

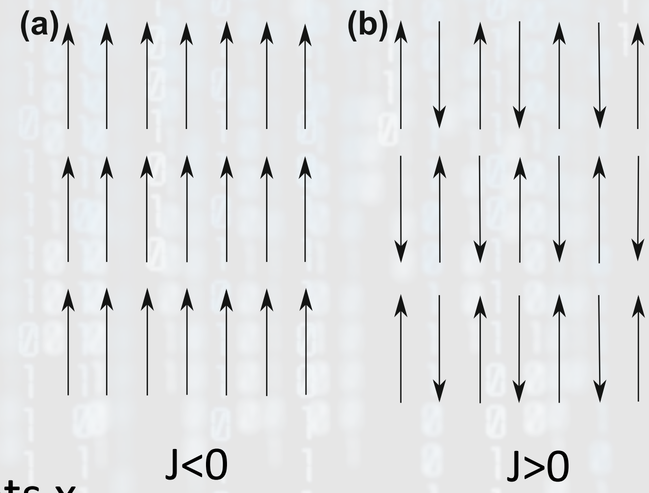


Spins in a solid – Heisenberg model

Given two atoms or molecules with spins S_1 and S_2, where $S_1, S_2 \in \mathbb{R}^3$, the exchange interaction energy is of the form:

$$E = JS_1 \cdot S_2$$

with exchange constant J. The ground state for a system of spins depends on the sign of J →
For the antiferromagnetic case (J>0) the transition temperature T_c is called Neel temperature.



For atoms (spins) on a cubic lattice with lattice points x_ℓ, we can write the Hamilton function

$$H = \frac{1}{2}\sum_{\ell\ell'} J_{\ell\ell'} S_\ell \cdot S_{\ell'} = \frac{1}{2}\sum_{\ell\ell'} J_{\ell-\ell'} S_\ell \cdot S_{\ell'}$$

Heisenberg model

where $J_{\ell\ell}=0$ and $J_{\ell\ell'} \rightarrow J_{\ell-\ell'} = J_{\ell'-\ell}$ to account for translation invariance.

Ising model

In the Heisenberg model the spin directions are arbitrary. In the Ising model one restricts the direction to one direction, typically the z-direction. Thus the quantum mechanical description (spin quantization) of the (general) Ising model is given by the Hamiltonian

$$H = \frac{1}{2}\sum_{\ell\ell'} J_{\ell-\ell'} S^z_\ell S^z_{\ell'}$$

where S^z_ℓ are the spin operators in z-direction.
For spin ½ particles, the eigenvalues of the spin operator are $\sigma_\ell=\pm1$ (in units of $\hbar/2$) and we write

$$H = -\frac{1}{2}\sum_{\ell\ell'} J_{\ell-\ell'} \sigma_\ell \sigma_{\ell'} - h\sum_\ell \sigma_\ell$$

"classical" Ising model

where the factor $\hbar^2/4$ is absorbed in the coupling constants and h is an external magnetic field in z-direction.

…

usually the interaction is limited to nearest neighbors (n.n.) such that for $J_{\ell\text{-}\ell'}$:

$$J_{\ell-\ell'} = \begin{cases} J & \text{if} \quad \ell, \ell' \quad \text{n. n.} \\ 0 & \text{otherwise.} \end{cases}$$

→ $$\boxed{H = -\frac{J}{2} \sum_{\langle \ell\ell' \rangle} \sigma_\ell \sigma_{\ell'} - h \sum_\ell \sigma_\ell}$$

summation over n.n.

For an infinite system, this model can be solved analytically in 1D (Ising's solution) and 2D (Onsager solution)

The magnetization of the system in a certain configuration is given by $\mathcal{M}(\mathcal{C}) = \langle \sigma_\ell \rangle$.

Statistical Physics and definitions

The Boltzmann distribution is given by: $$p(\mathscr{C}) = \frac{1}{Z_N} \exp\left[-\frac{E(\mathscr{C})}{k_\mathrm{B}T}\right]$$

where E($\mathcal{C}$) is the energy of a spin configuration $\mathcal{C}$

The partition function is defined as

$$Z_N = \sum_{\mathscr{C}} \exp\left[-\frac{E(\mathscr{C})}{k_\mathrm{B}T}\right]$$

from which we can derive the average energy

$$\langle E \rangle = \sum_{\mathscr{C}} p(\mathscr{C}) E(\mathscr{C}) = k_\mathrm{B}T^2 \frac{\partial}{\partial T} \ln Z_N$$

and magnetization

$$\langle M \rangle = \sum_{\mathscr{C}} p(\mathscr{C}) \mathscr{M}(\mathscr{C}) = k_\mathrm{B}T \frac{\partial}{\partial h} \ln Z_N$$

using these we can define

$$\chi = \frac{\partial}{\partial h} \langle M \rangle$$ **susceptibility**

$$c_h = \frac{\partial}{\partial T} \langle E \rangle$$ **specific heat**

...

Final expressions are:

$$
\begin{aligned}
c_h &= \frac{1}{k_\mathrm{B}} T^{\overset{-2}{\cancel{2}}} \sum_{\mathscr{C}} p(\mathscr{C}) \left[E^2(\mathscr{C}) - E(\mathscr{C}) \langle E \rangle \right] \\
&= \frac{1}{k_\mathrm{B} T^2} \left(\langle E^2 \rangle - \langle E \rangle^2 \right) \\
&= \frac{1}{k_\mathrm{B} T^2} \operatorname{var}(E) \ .
\end{aligned}
$$

$$
\begin{aligned}
\chi &= \frac{1}{k_\mathrm{B} T} \sum_{\mathscr{C}} p(\mathscr{C}) \left[\mathscr{M}^2(\mathscr{C}) - \mathscr{M}(\mathscr{C}) \langle M \rangle \right] \\
&= \frac{1}{k_\mathrm{B} T} \left(\langle M^2 \rangle - \langle M \rangle^2 \right) \\
&= \frac{1}{k_\mathrm{B} T} \operatorname{var}(M) \ .
\end{aligned}
$$

1D solution

The partition function in 1D for N spins can be calculated as

$$Z_N = \lambda_1^N + \lambda_2^N$$

with

$$\lambda_{1,2} = \exp\left(\frac{J}{k_B T}\right)\cosh\left(\frac{h}{k_B T}\right) \pm \sqrt{\exp\left(\frac{2J}{k_B T}\right)\sinh^2\left(\frac{h}{k_B T}\right) + \exp\left(-\frac{2J}{k_B T}\right)}$$

The expectation value for the energy per particle is given by

$$\langle \varepsilon \rangle = \frac{k_B T^2}{N}\frac{\partial}{\partial T}\ln Z_N$$

In the thermodynamic limit N→∞: $\lim_{N\to\infty}\frac{1}{N}\cancel{Z_N} = \lim_{N\to\infty}\frac{1}{N}\ln\left(\lambda_1^N + \lambda_2^N\right) = \ln\lambda_1$

$\ln(Z_N)$

and for h=0: $\boxed{\lim_{N\to\infty}\frac{1}{N}\cancel{Z_N} = \ln\left[2\cosh\left(\frac{J}{k_B T}\right)\right]}$

smooth function of T for T > 0
→ no phase transition in the one dimensional Ising model

2D Onsager solution

For h=0 one observes a second order phase transition with transition temperature defined by:

c_h and χ diverge at T_c

$$2\tanh^2\left(\frac{2J}{k_\mathrm{B}T_C}\right) = 1 \qquad k_B T_c = \frac{2J}{\log(1+\sqrt{2})} \approx 2.269J$$

and for the energy per particle

$$\langle\varepsilon\rangle = -J\coth\left(\frac{2J}{k_\mathrm{B}T}\right)\left\{1 + \frac{2}{\pi}K_1(\xi)\left[2\tanh^2\left(\frac{2J}{k_\mathrm{B}T}\right) - 1\right]\right\}$$

where $K_1(\xi)$ is the complete elliptic integral with

$$\xi = \frac{2\sinh\left(\frac{2J}{k_\mathrm{B}T}\right)}{\cosh^2\left(\frac{2J}{k_\mathrm{B}T}\right)}$$

and magnetization per particle

$$\langle m\rangle = \begin{cases} \dfrac{(1+z^2)^{\frac{1}{4}}(1-6z^2+z^4)^{\frac{1}{8}}}{\sqrt{1-z^2}} & \text{for } T < T_C \\ 0 & \text{for } T > T_C \end{cases}$$

with $z = \exp\left(-\frac{2J}{k_\mathrm{B}T}\right)$

$$= \left(1 - \left[\sinh\left(\log(1+\sqrt{2})\frac{T_c}{T}\right)\right]^{-4}\right)^{\frac{1}{8}} \qquad T<T_c$$

Phase transitions

Ehrenfest classification of Phase Transition:

- **First-order phase transitions** exhibit a discontinuity in the first derivative of the chemical potential with a thermodynamic variable. Such as solid/liquid/gas transitions.
- **Second-order phase transitions** (also called continuous phase transition) have a discontinuity or divergence in a second derivative of the chemical potential with thermodynamic variables.

c_h and χ are second derivatives

Critical exponents

Reduced temperature: $\tau \equiv \frac{T - T_C}{T_C}$

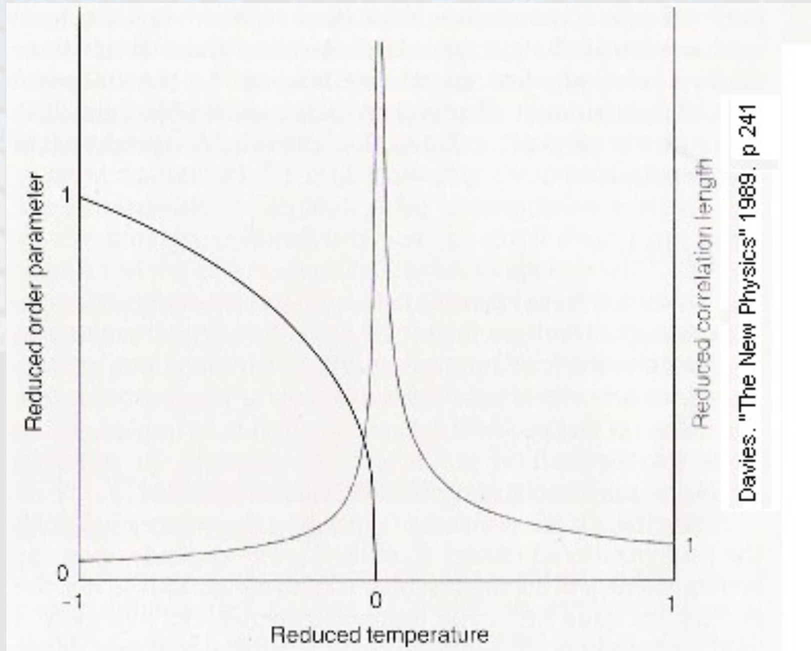
Critical exponent: $k \stackrel{\text{def}}{=} \lim_{\tau \to 0} \frac{\log|f(\tau)|}{\log|\tau|}$

Specific heat $C \propto |\tau|^{-\alpha}$ ($T<T_c$)

Magnetization $M \propto |\tau|^{\beta}$

Magnetic susceptibility $\chi \propto |\tau|^{-\gamma}$

Correlation length $\xi \propto |\tau|^{-\nu}$



Critical behavior of the order parameter an the correlation length. The order parameter vanishes with the power β of the reduced temperature t as the critical point is approached along the line of phase coexistance. The correlation length diverges with the power ν of the reduced temperature.

The exponents display critical point universality (don't depend on details of the model). This explains the success of the Ising model in providing a quantitative description of real magnets.

Ising values

d	2	3	4
α	0 (log div)	0.110(1)	0
β	1/8	0.3265(3)	1/2
γ	7/4	1.2372(5)	1
δ	15	4.789(2)	3
η	1/4	0.0364(5)	
ν	1	0.6301(4)	1/2
ω	2	0.84(4)	

At $T=T_c$: $H \propto M^{\delta}$, $\langle\sigma(0)\sigma(r)\rangle \propto r^{2-d-\eta}$

Numerics

Here we concentrate on the 2D case of the Ising model and place the spins on a square lattice Ω with coordinates (x_i, y_j), i,j,=1,…,N.

$$H = -J \sum_{\left\langle \substack{ij \\ i'j'} \right\rangle} \sigma_{i,j}\sigma_{i',j'} - h \sum_{ij} \sigma_{i,j}$$

with $\sigma_{i,j}=\pm 1$ – i.e. we treat these as “classical” spins. We consider only n.n. interaction.

Task: calculate numerically observables like the expectation value of the energy or of the magnetization, which will then be compared with analytic results.

Problem: We cannot sample all possible configurations of the system which grow exponentially with grid points, e.g. for N=100 we have $2^{N*N}=2^{10000}\approx 10^{3000}$ possible arrangements of spins.

Solution: Use Monte-Carlo methods – here the Metropolis algorithm

For example, using M (not magnetization here) configurations, we can get an approximation for the energy expectation value

$$\langle E\rangle = \frac{1}{M}\sum_{i=1}^{M} E(\mathscr{C}_i) \pm \sqrt{\frac{\mathrm{var}\,(E)}{M}}$$

$\mathrm{var}(E)=\langle E^2\rangle-\langle E\rangle^2$

$$\langle E^2\rangle = \frac{1}{M}\sum_{i=1}^{M} E_i^2$$

if these M configurations follow the Boltzmann distribution (this basically means that we neglect most configurations which have exponentially small probability to be physical.)

Metropolis rejection for Ising model

After generation a new trial configuration $\mathcal{C}^t$ from the current, valid configuration $\mathcal{C}$, we accept the new one with the following probability

$$\Pr(A|\mathscr{C}^t,\mathscr{C}) = \min\left(\frac{p(\mathscr{C}^t)}{p(\mathscr{C})},1\right) = \min\left\{\exp\left[-\frac{E(\mathscr{C}^t)-E(\mathscr{C})}{k_\mathrm{B}T}\right],1\right\}$$
$$= \min\left[\exp\left(-\frac{\Delta E_{ij}}{k_\mathrm{B}T}\right),1\right].$$

$\mathcal{C}^t$ is generated by just flipping one spin on the grid. Since we only consider n.n. interaction, the energy difference of the two configurations is given by the simple expression:

$$\Delta E_{ij} = 2J\sigma_{i,j}\left(\sigma_{i+1,j}+\sigma_{i-1,j}+\sigma_{i,j-1}+\sigma_{i,j+1}\right)+2h\sigma_{i,j}$$

Practical considerations

- We consider a two-dimensional NxN square lattice with periodic boundary conditions in order to reduce finite volume effects

$$\sigma_{N+1,j} = \sigma_{1,j} \qquad \text{and} \qquad \sigma_{i,N+1} = \sigma_{i,1}$$

- **Do not use** the n.n. matrix suggested in the book – this is usually very inefficient (requires memory access). N.n. coordinates are easy to calculate on-the-fly
- We need a good PRNG to choose random sites
- Initial configuration: the Metropolis algorithm produces configurations which become independent of the initial state and follow the Boltzmann distribution → start with random spins
- Run the algorithm for "a while" to ***thermalize*** the system – disregard these initial trial movements. How to determine "a while"? One should check if thermal equilibrium has been reached by analyzing the observable under consideration as function of time, then determine when initial biases are gone. In this case the observable reaches some saturation value as a function of the number of measurements.

…

- To check if saturation is reached, one should start with two different initial conditions and check when the observables will converge
- typically we study temperature dependencies and change the temperature once the equilibrium has been reached: either hot (random intial spins) or cold (ordered initial spins) start

Running the code

The Metropolis algorithm for the Ising model is executed in the following steps:

1. Choose an initial configuration $\mathcal{C}_0$
2. we go through all lattice sites (either systematically/sequentially, by random permutation, or completely random) and flip the spin: $\sigma_\ell \rightarrow -\sigma_\ell$. Calculate $\Delta E_{i,j}$ – one complete loop through all site is called a sweep. MC simulations typically need many sweeps
3. Accept the new configuration according to

$$\Pr(A|\mathscr{C}^t, \mathscr{C}_k) = \min\left[\exp\left(-\frac{\Delta E_{ij}}{k_\mathrm{B} T}\right), 1\right]$$

4. go to next lattice site, repeat 2

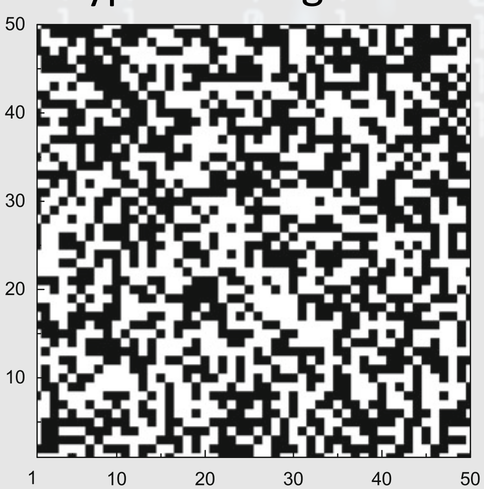
Measurement:

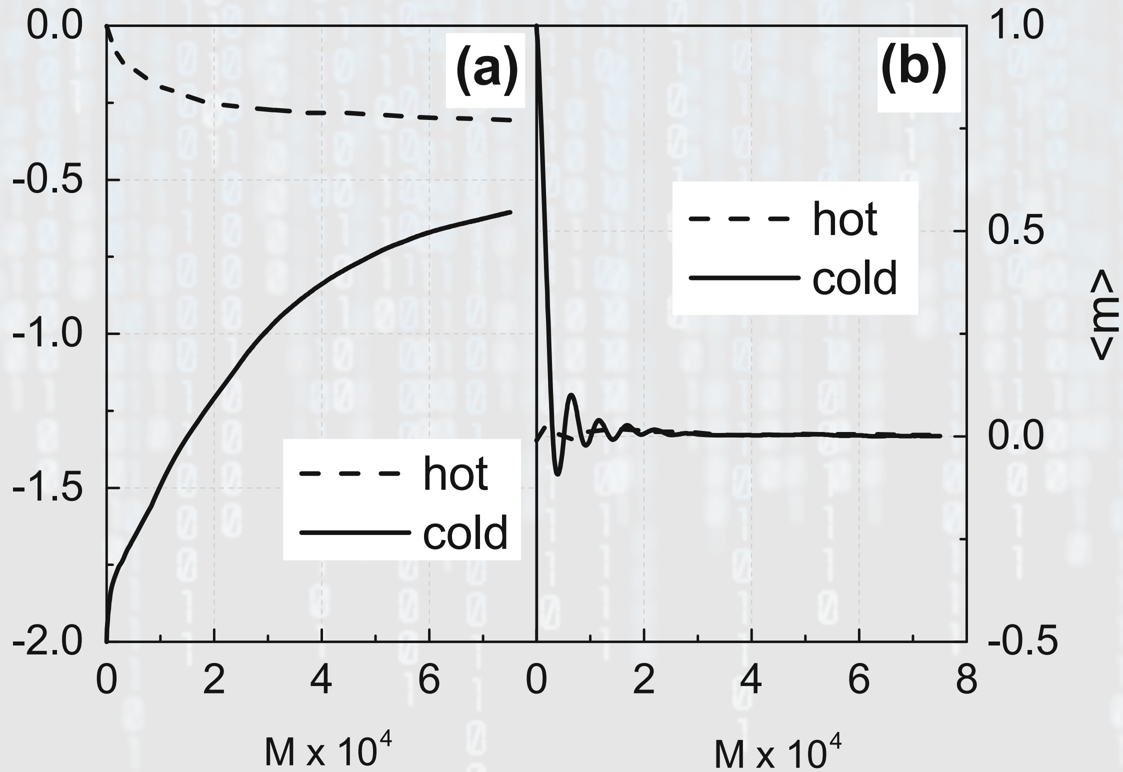
1. do the above till thermal equilibrium is reached
2. start calculation of observables and average over N trial configurations, when accepted
3. (optional, repeat) change external parameter (T, h), re-equilibrate (typically shorter than the initial equilibration), and then average observable

Some results

h=0, J=0.5
N=50,
~10^5 measurements
(32 sweeps)
$k_B T=3>k_B T_c$ →
paramagnetic state,
$\langle m\rangle=0$

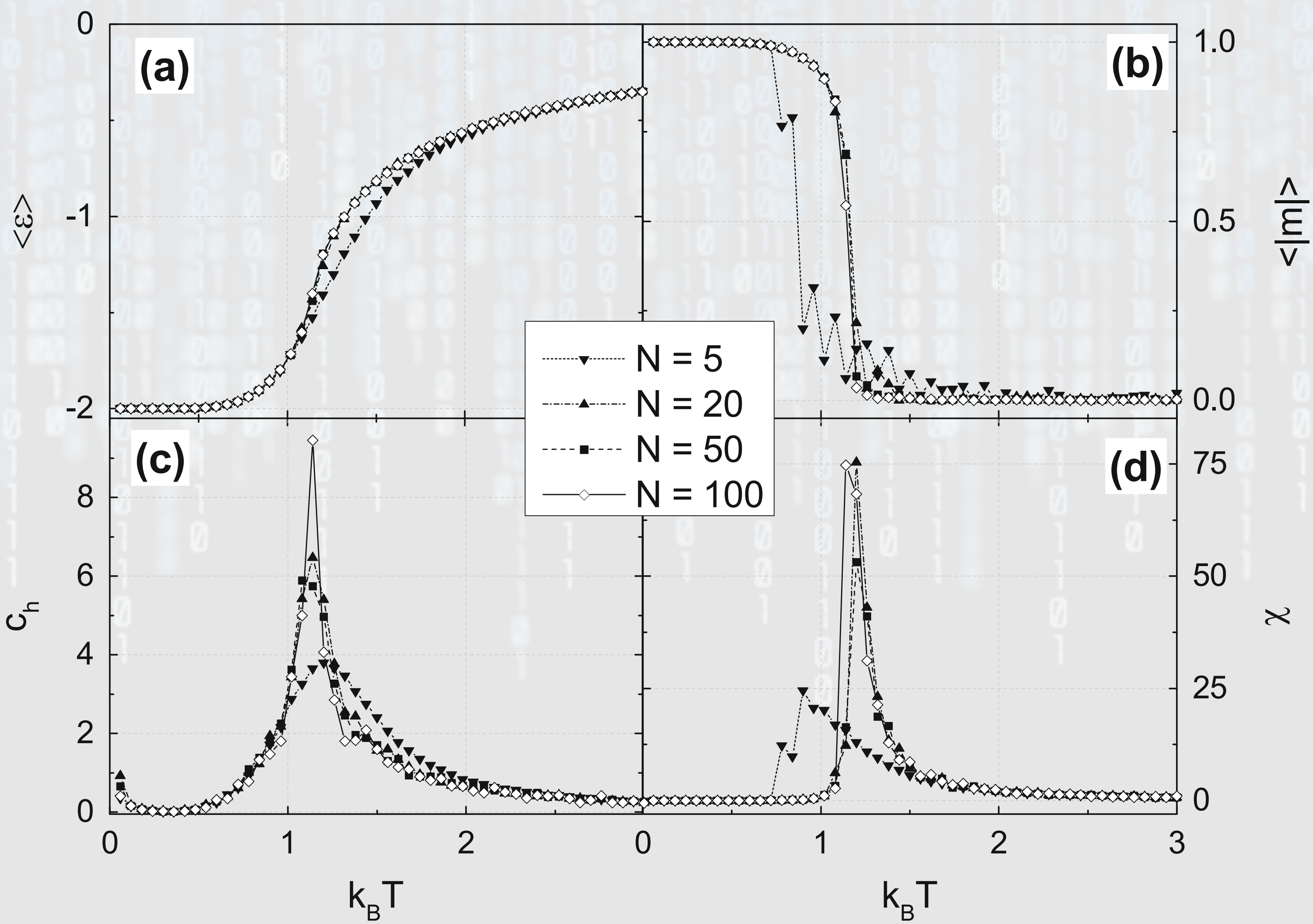
typical configuration



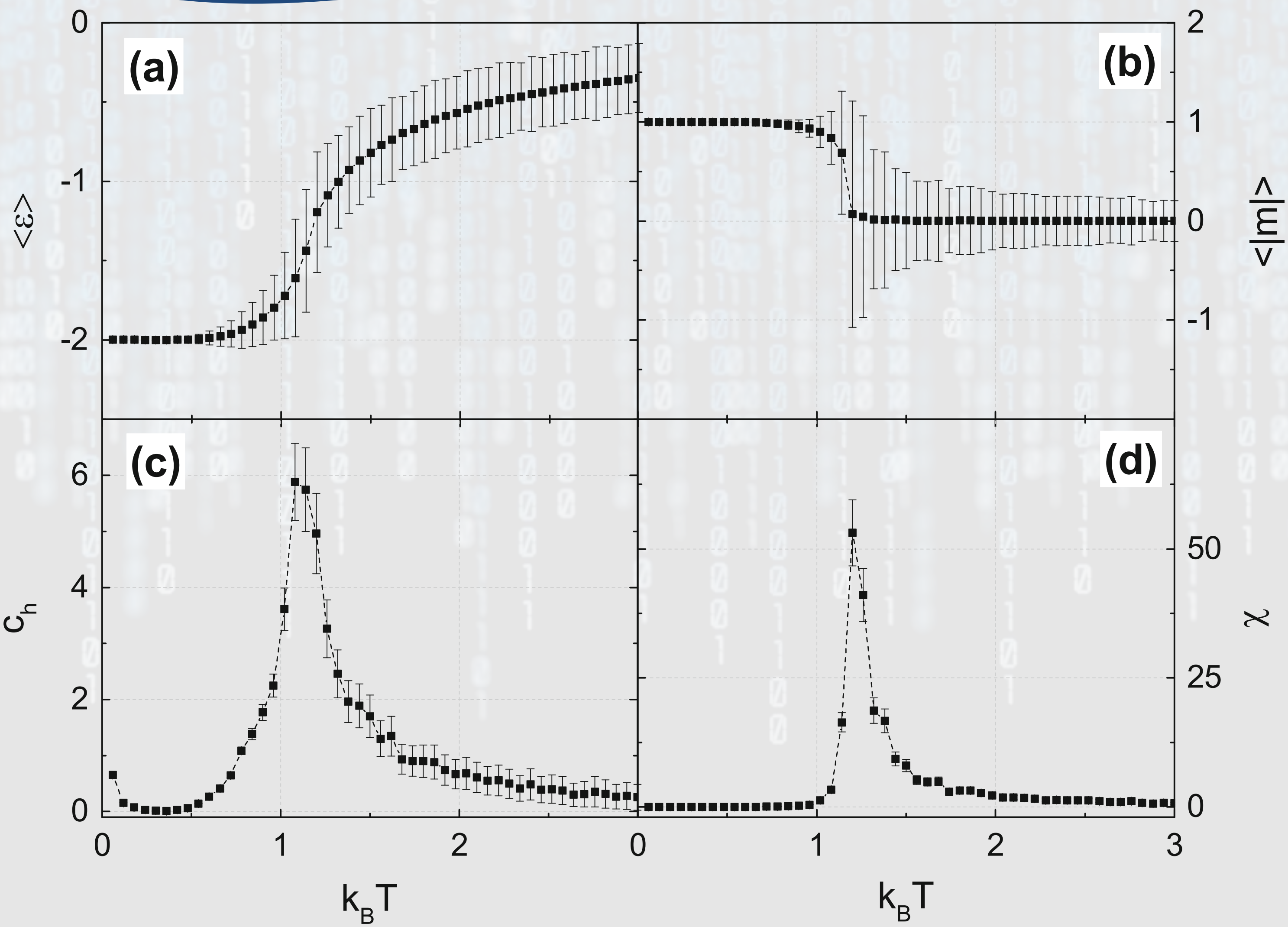


→ 32 sweeps are not long enough to equilibrate the energy

temperature dependence



with error bars



domains near T_c