Stochastic simulation methods
or: Monte-Carlo methods

Zoltán Néda
UBB-Cluj, Romania
zneda@phys.ubbcluj.ro

Main objective of the course:
To give a PRACTICAL introduction to Monte Carlo methods in physics. To study modern problems in the field. Start research....

What are the Monte Carlo 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)
- Cellular automata (approach to a given phenomena discretized on a lattice, with deterministic or stochastic update rules)
- In general we speak about Monte Carlo simulation methods whenever the use of the random numbers are crucial in the algorithm!
- Monte Carlo techniques are widely used in problems from: statistical physics, soft condensed matter physics, material science, many-body problems, complex systems, fluid mechanics, biophysics, econophysics, nonlinear phenomena, particle physics, heavy-ion physics, surface physics, neuroscience etc....
Some well-known problems where MC methods are useful

2. Percolation problems:

3. The Ising model:

$$H = -J \sum_{<i,j>} S_i S_j - \mu H \sum_i S_i$$

Some well-known problems where MC methods are useful

4. The p-state Potts model

Feeling what Monte Carlo simulation means

- 1. Studying the random walk -

- A basic model in natural sciences
  - Brownian motion, fluctuations, diffusion etc...
  - We consider first the simple 1D case

- Experimental realization: the Galton board

Analytical study

Quantities of interest:

- $\sqrt{\langle x^2 \rangle} = N^\alpha$ for $P(N, k) = ?$
- We study the $N \gg 1$ and $k \ll N$ limit and use the following approximations:

- $\ln(N(k, N)) = \ln(N^2) - \ln((N+k)/2) - \ln((N-k)/2) - N \ln(2)$

- The scaling exponent $\alpha$ is independent of the dimension!

Interesting problems:

- Random walks with restriction or memory

Special case: self-avoiding random walk $\rightarrow$ what is $\alpha$?

Scaling properties

$$\langle x^2 \rangle_N = \frac{N}{N^2} \int_0^N x^2 P(x) \, dx = \frac{N}{2}$$

The 2D and 3D cases:

$$\langle x^2 \rangle_N^{(2)} = \langle x^2 \rangle_N^{(3)} = \frac{N}{2} + \frac{N}{3} = N$$

$$\langle x^2 \rangle_N^{(3)} = \langle x^2 + y^2 + z^2 \rangle_N^{(3)} = \langle x^2 \rangle_N^{(3)} + \langle y^2 \rangle_N^{(3)} + \langle z^2 \rangle_N^{(3)} = N^2 \frac{N}{2} + N^3 \frac{N}{3} = N$$

The scaling exponent $\alpha$ is independent of the dimension!
Solving the problem by MC-type simulations

- the idea: reproducing the random walk by using "random numbers", and realizing the experiment with \( N \) random steps many times, calculating numerically thus \( \Delta x \)

In reality there is no stochastic process in our calculator, so in simulations we use numbers (using for example tunnel diodes, etc.). The speed of these generators are however very low.

In principle one can design interfaces which will be able to generate real random numbers will approach a desired random behavior if their statistics satisfy some constraints. Generating a random integer \( \text{rand}() \) in the \([0, M)\) interval:

\[
R = \text{rand()} \mod M
\]

Uniformly distributed pseudo-random numbers

The core of most of them are the modulo generators.

Primary task is to arrange integers from 1 to \( M-1 \) in "random order":

\[
s_k = (ax_{k-1} + c) \mod M
\]

Generating a random "float" in the \([0,1)\) interval:

\[
x = (\text{rand()} / \text{RAND\_MAX} + 0.0) \times (R_{\text{max}} - R_{\text{min}})
\]

Generating a random integer in \([R_{\text{min}}, R_{\text{max}})\) interval:

\[
x = R_{\text{max}} + (\text{rand()} / \text{RAND\_MAX}) \times (R_{\text{max}} - R_{\text{min}})
\]

Phase transition in a sociological system

In a room with \( L \times L \) sizes, there are \( N \) rats. Each rat can be in two states: either calm (state 0) or nervous (state 1). The system of rats obey the following dynamical rules:

1. The rats randomly run through the whole room. From time to time they stop and look around. Each rat detects only those rats that are within a distance smaller than \( r \).
2. If a nervous rat see no other rat around him, it becomes calm. Otherwise remains nervous.
3. If a calm rat sees a nervous rat around him, it becomes nervous. Otherwise remains calm.
4. With a very small \( p \) probability a calm rat can become nervous accidentally.

Problem: prove, that in the thermodynamic limit \((L \to \infty)\) the system exhibits a phase transition as a function of the rat density \( \rho \). I.e. there is a critical rat density \( \rho_c \) in the system, so that for \( \rho < \rho_c \) the stable dynamic equilibrium is that the rats density decreases and for \( \rho > \rho_c \) it increases.

Analytical solution

- isolated rats: \( \rho = N \) density of random rats, \( \rho_c = N \) density of calm rats (\( \rho_c = 1 \) if \( \rho = 1 \))
- \( P_{\text{eq}}(N,N) \) probability/weight that a nervous rat becomes calm, \( P_{\text{eq}}(N,N) \) probability/weight that a calm rat becomes nervous

The C program can be found on the course home-page.

Random numbers

- the key to MC simulations

- In order to get random numbers we need a real stochastic (random) process like: throwing a dice or tossing a coin
- In reality there is no stochastic process in our calculator, so in simulations we use pseudo-random numbers, generated deterministically by our computer. These numbers will approach a desired random behavior if their statistics satisfy some properties.
- In principle one can design interfaces which will be able to generate real random numbers (using for example tunnel diodes, etc.). The speed of these generators are however very low.

Random numbers

- Uniformly distributed on a given interval (real numbers or integers)
- Distributed according to a given distribution

Managing the analytical solution for eq.2

we consider the \( p \to 0 \) limit

- \( n \) is the order-parameter in the system
- In the vicinity of the critical point \( n = 1 \)

\[
1 - e^{1(n-1)} = 0 \quad \Rightarrow \quad n = \frac{1}{1 - e^{-1}}
\]

\[
N = \left\lfloor \frac{1}{1 - e^{-1}} \right\rfloor
\]

if \( N=N_c \), the stable solution is \( n = n_c \)

- Numerical solution of eq.1 (continuous line)
- Eq.2 (dashed line)

MC simulation of the problem

The 'rats' are placed in new random positions at each simulation step \( \psi \to \) a fast uncorrelated random motion.

The algorithm:

1. We fix the simulation parameters: \( r \), \( L\times L \). Number of transient steps, \( p \) - probability of getting nervous accidentally, number of steps on which averaging for \( \rho \) is done.
2. We consider simulation with different rat number, output cycle \( \ldots \) For each case we initialize the states (calm or nervous) for each rat
3. Using the dynamical rules 1.-4. we give new random positions for the rats, and update their states. We do this many times, first as many times as many transient steps are, and then as many steps as needed for the average
4. We study the average value of \( n \) as a function of \( \rho \)
Testing the uniform random number generator

1. Determining the repetition period (after how many calls the series will repeat). This must be as big as possible...

2. Testing the uniformity of the distribution – the histogram test.
   (Both for integer and float generators)
   Idea: construct a histogram for the numbers generated in some fixed constant intervals. Denote by \( y_i \) the frequency of generating a number in the "i"-th bin: \( y_i \sim \frac{N}{n} \). The values of \( y_i \) must converge to \( \frac{1}{n} \) for a uniform distribution as \( N \to \infty \).

   \[
   y_i = \frac{1}{n} \sum_{n=1}^{N} \chi_{[x_n,1/n)}(x_n)
   \]

   \( g \to 0 \)

3. The return map test – testing both the uniformity and the absence of correlation in the \([0,1)\) interval. A visual test by plotting on an \(x\)-\(y\) coordinate system \( (x_n,y_n) \) as a function of \( x_n \). If the generator is a proper one, the points must cover uniformly the \([0,1)\times[0,1)\) square.

4. The absence of short-range correlations – the correlation test

   For the total absence of \( k \)th order correlation \( C(k) \)
   we must have \( C(k) \to 0 \)

5. The absence of long-range correlations – the return map test

Pseudo-random numbers distributed according to a desired distribution

We are looking for a Gen2 random number generator, that gives random numbers according to the \( g(x) \) distribution function, on the \([R_{\text{min}}, R_{\text{max}}]\) interval.

Let us suppose that Gen1 gives random numbers distributed uniformly on the \([0,1)\) interval.

\[
\begin{align*}
G_1 &= y \\
G_2 &= G^{-1}(y + G(R_{\text{min}}))
\end{align*}
\]

Properties of a good "random number" generator

- the basic generator should have a long period
- no detectable correlation between the terms
- distribution close to the desired one already for relatively short series
- should be very fast! (should not contain mathematical functions like \( \exp() \), \( \sin() \), etc.)
- should be tested before the use
- should be repeatable for optimal debugging purposes

Elements of statistical physics

- Statistical physics deals with systems of large number of particles or stochastic processes
- The (3D) state-space of one particle
- The (6D) state-space of \( N \) particles (the state of the system is characterized by a characteristic point in this 6D space)
- The allowed region of the 6D state-space (region of the state-space where the characteristic point can move, points permitted by the externally imposed conditions)
- The externally imposed conditions – the ensemble in which the systems is
- The ergodic principle: in a very short time (much shorter than the time needed for a physical measurement) the characteristic point of the system visits all the allowed points of the state-space.

The microcanonical ensemble

All allowed points of the state-space are equally probably realized.

Important microscopic quantity:
\( W \): the number of allowed microstates

\[
S = k \ln(W)
\]

Boltzmann’s equation

Generalization of the Boltzmann equation: Renyi entropy valid for all ensembles:

\[
S = -k \sum_{i=1}^{N} p_i \ln(p_i)
\]

Useful equation for handling analytically or numerically \( W \):

\[
\ln(W) = N \ln(N) - N + \frac{1}{2} \ln(2\pi N)
\]

Stirling’s formula

\[
\ln(N) \approx N \ln(N) - N + \frac{1}{2} \ln(2\pi N)
\]
The order parameter (m)

- Characterizes the degree of the order in the system and the phases
- It is usually a first derivative of the relevant thermodynamic potential
- It is usually adimensional
- Discontinuity at the transition point for first-order phase-transitions
- It’s derivative has a discontinuity for second-order phase-transitions
- Examples of order parameters:
  - It’s derivative has a discontinuity for second-order phase-transitions
  - Defined for second-order phase-transitions
  - It is usually an adimensional
  - It is usually a first derivative of the relevant thermodynamic potential
  - Characterizes the degree of the order in the system and the phases

Critical exponents

- Defined for second-order phase-transitions
- All relevant physical quantities have a power-law behavior in the vicinity of the transition point. The quantities either converge to 0, or diverges.
- Important is the universality of the critical exponents (they do not depend on the microscopic details of the model, just on the symmetry properties of the order parameter and the dimensionality of the system)
- Some important critical exponents:
  - β: critical exponent of the order parameter
  - γ: critical exponent of the susceptibility
  - ν: critical exponent of the correlation length

Elements of Stochastic Processes

Markov processes/ Markov chains

Markov processes (chain) is 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 \( S \) be the state of the stochastic system at step \( t \); a stochastic variable

The time-evolution of the system is described by a sequence of states: \( S_1, S_2, \ldots, S_n \).

The conditional probability that \( S_n \) is realized if previously we had \( S_1, S_2, \ldots, S_{n-1} \):

\[
P(x_n | x_1, x_2, \ldots, x_{n-1}) = P(x_n | x_{n-1})
\]

Definition: A Markov process is a sequence of states that are connected by transition probabilities.

One-step transition probabilities, elements of the stochastic matrix

\[
P(x_n | x_{n-1}) = P_{n-1}(x_n \to x_{n-1})
\]

Definition: A probability distribution over the possible states \( \{x_0\} \) is called invariant or stationary for a given Markov chain if satisfies: \( \{x_0\} \to P(x_0) \).

Elements of critical phenomenon

Critical phenomenon: behavior near a phase-transition point

- First order transitions (the first order derivative of the relevant thermodynamic potential has a discontinuity)
- Examples: boiling, freezing, hardening...
- Second order phase transitions (the second order derivative of the relevant thermodynamic potential has a singularity or discontinuity)
- Examples: para-ferromagnetic transition, percolation...
- Many other type...

Critical exponents are crucial quantities in order to understand and model phase transitions.

\[
A = \frac{1}{x^c}
\]

Ex. for 2D systems \((d=2)\)

\[
\alpha + 2\beta + \gamma = 2
\]

Hyperscaling equation (valid in all dimensions)

\[
d\nu = 2 - \alpha
\]
Brownian Dynamics

- It is a hybrid method, involving both deterministic and stochastic dynamics.
- In molecular dynamics methods all degrees of freedom are explicitly taken into account -> classical equation of motion of particles
- In Brownian dynamics some degrees of freedom are represented only through their stochastic influence

Langevin equation of motion: stochastic equation of motion. The coupling to the heat bath is realized through the $R(t)$ stochastic force.

$$\frac{dv}{dt} = R(t) - \mu v$$

**Question:** What properties should $R(t)$ have, in order to be equivalent with a heat bath at temperature $T$?

- Another way of doing Brownian dynamics:
  - by taking into account the coupling of the system to the heat bath by "statistical" collisions with virtual particles. In this approximation no friction is necessary.
  - each stochastic collision is assumed to be an instantaneous event
  - the colliding virtual particles have a Maxwell-Boltzmann momentum distribution
  - the time intervals at which particles suffer a collision is distributed according to $P(t) = \lambda e^{-\lambda t}$ ($\lambda$ is the mean collision time)

**Algorithm II. for making Brownian Dynamics:**

1. Get initial position and velocity for the particle
2. Choose time intervals according to the above distribution
3. Integrate the equations of motion until the time of a stochastic collision
4. Choose a momentum at random from the Maxwell-Boltzmann distribution at temperature $T$
5. Proceed with step 3

The Monte Carlo method

**Definition:** Monte Carlo methods use random sequence of numbers to calculate statistical estimates on a sample population for a desired parameter.

- known examples: calculating PI, calculating recession thresholds...
- other examples: calculating average magnetization and energy for the Ising model
- in general: applications are enormous and fascinating ....

The outline of MC methods:

1. Description of the system in terms of a Hamiltonian
2. Selecting an appropriate ensemble for the problem
3. Observables are computed using an associated distribution function. Ultimately the goal is to compute quantities appearing as results of high-dimensional integrations on the state-space. The idea is to sample the main contributions to get an estimate for the observable.

Starting point: one dimensional Monte Carlo integration

One dimensional Monte Carlo Integration

**Problem:** given a function $f(x)$, compute the integral:

$$\int f(x) dx$$

**Straightforward sampling**

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

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

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

(1) The above affirmation is also true if the random numbers are correlated, or the intervals are finite

**How rapidly the method converge?**

For $\mu(x)$ close to very bad!!

**Central limit theorem:** If:

$$\sigma = \int f(x) \mu(x) dx$$

then:

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} f(x_i) - I \leq \frac{\sigma \sqrt{2\pi}}{\sqrt{n}}$$

**Algorithm for simulating the Brownian dynamics:**

1. Assign initial position and velocity for the particle
2. Draw a random number from a Gaussian distribution with mean zero and variance as described above. This will give us $R(t)$
3. Integrate the equation of motion with the obtained value of $R$, and get the new positions and velocities.
4. Proceed with step 2.

**Exercises**

1. Prove by computer simulations that the given recipe for $R(t)$ leads to a Maxwell-Boltzmann distribution of the particles velocities (in 1D)
2. Study the motion of a particle in a harmonic potential and subject to a heat-bath at temperature $T$. (in 1D)
3. Study the motion of a particle in a W potential valley, in contact with a heat-bath at temperature $T$. Both parts of the W potential valley are harmonic (in 1D)
4. Study problem nr. 3 when the two minimum of the W potential valley is modulated in anti-phase by a time-like harmonic component. Calculate the correlation function between the particle’s position and the external modulating field (1D case) (the phenomenon of stochastic resonance)

**Proving:**

$$P(V) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{V^2}{2\sigma^2}}$$

$$V = \int f(x) \mu(x) dx$$

$$\sigma = \int f(x)^2 \mu(x) dx$$

$$I = \int f(x) \mu(x) dx$$

$$J = \int f(x) \mu(x) dx$$

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

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} f(x_i) - I \leq \frac{\sigma \sqrt{2\pi}}{\sqrt{n}}$$

**Exercises**

1. Prove by computer simulations that the given recipe for $R(t)$ leads to a Maxwell-Boltzmann distribution of the particles velocities (in 1D)
2. Study the motion of a particle in a harmonic potential and subject to a heat-bath at temperature $T$. (in 1D)
3. Study the motion of a particle in a W potential valley, in contact with a heat-bath at temperature $T$. Both parts of the W potential valley are harmonic (in 1D)
4. Study problem nr. 3 when the two minimum of the W potential valley is modulated in anti-phase by a time-like harmonic component. Calculate the correlation function between the particle’s position and the external modulating field (1D case) (the phenomenon of stochastic resonance)

**Proving:**

$$P(V) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{V^2}{2\sigma^2}}$$

$$V = \int f(x) \mu(x) dx$$

$$\sigma = \int f(x)^2 \mu(x) dx$$

$$I = \int f(x) \mu(x) dx$$

**Algorithm for simulating the Brownian dynamics:**

1. Assign initial position and velocity for the particle
2. Draw a random number from a Gaussian distribution with mean zero and variance as described above. This will give us $R(t)$
3. Integrate the equation of motion with the obtained value of $R$, and get the new positions and velocities.
4. Proceed with step 2.

**Exercises**

1. Prove by computer simulations that the given recipe for $R(t)$ leads to a Maxwell-Boltzmann distribution of the particles velocities (in 1D)
2. Study the motion of a particle in a harmonic potential and subject to a heat-bath at temperature $T$. (in 1D)
3. Study the motion of a particle in a W potential valley, in contact with a heat-bath at temperature $T$. Both parts of the W potential valley are harmonic (in 1D)
4. Study problem nr. 3 when the two minimum of the W potential valley is modulated in anti-phase by a time-like harmonic component. Calculate the correlation function between the particle’s position and the external modulating field (1D case) (the phenomenon of stochastic resonance)
For straightforward sampling (p(x)) = const) the error ~ 1/n^2!!!

- The error is dependent on the choice of f(x) and p(x)! = influencing σ

A better method for calculating

\[ \int f(x) dx = \int \frac{f(x)}{p(x)} p(x) dx \]

We generate random x, x', x points according to the p(x) distribution

The basic idea: if we choose p(x) as close as possible to f(x), we get σ = 0 and the method converges rapidly for small values of n!!!

**Problem:** The methods needs advance knowledge of f(x)!

One way to overcome the problem is by guessing some p(x) functions, that mimics well the behavior of f(x). The error is also considerably reduced!

**Importance sampling:** Sampling in the neighborhood where f(x) is large!

---

### Monte Carlo for statistical physics problems

We want to compute integrals like:

\[ \sum \frac{A(x)}{f(H(x))} f(H(x)) \]

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

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

If in the MC integration we choose the states with probability P(x) ->

\[ A \geq \frac{1}{n} \sum \frac{A(x)}{f(H(x))} f(H(x)) \]

By choosing \( P(x) = \frac{f(H(x))}{Z} \) \( \sigma \to 0 \) and thus the error \( \to 0 \)

**Problem:** we still don’t know Z!

---

### Example: the canonical ensemble

\[ \frac{1}{Z} \exp \left( -\frac{H(x)}{T} \right) \]

**Metropolis dynamics:**

\[ W(x) \propto \exp(\Delta E(x, x')/k_B T) \] if \( \Delta E(x, x') > 0 \);

\[ W(x) = 1 \] if \( \Delta E(x, x') = 0 \)

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

**Glauber dynamics:**

\[ W(x) = \exp(\Delta E(x, x')/k_B T) \] if \( \exp(\Delta E(x, x')/k_B T) \)

\[ \Delta E(x, x') = H(x') - H(x) \]

**Algorithm for Monte Carlo simulations:**

1. Specify an initial point \( x \) in the phase space
2. Generate a new state \( x' \)
3. Compute the \( W(x) \) Transition probability
4. Generate a uniform random number \( r \) between [0,1]
5. If \( r = \frac{W(x)}{W(x')} \) jump to the new state, and return to 2.
6. Average the desired A quantity on all states after the initial \( n=1 \) “transient” states

---

### The Ising model

\[ H = -J \sum_{\langle i,j \rangle} S_i S_j - \mu B \sum S_i \]

Interaction with nearest neighbors only!

- spontaneous magnetization is possible (N=4 for B=0)
- first model for understanding ferromagnetic and antiferromagnetism for localized spins
- for \( J > 0 \) \( \rightarrow \) ferromagnetic order
- for \( J < 0 \) \( \rightarrow \) anti-ferromagnetic order
- no phase transition in 1D
- ferro-paramagnetic phase transition for \( D > 1 \)
- second order phase transition (order parameter)
- Approximation methods: mean-field theory, renormalization, high and low temperature expansion

---

**Important quantities:**

- \( m(T) \) curve
- \( T_c \)
- the critical exponent of susceptibility \( \gamma \)
- order parameter \( m \)
- Specific heat(c) and correlation length(h)

**Exact results:**

- (0D) \( T_c = 0 \)
- (2D) \( T_c = 2.269118 J/k_B \) (square lattice); \( \beta = 1/8; \gamma = 0 \)
- (3D) \( T_c = 4.44 J/k_B \) (square lattice); \( \beta = 1/8; \gamma = 5/4 \)
- (4D) no exact results (believed that: \( T_c = 1.09 \))
The transfer matrix solution to the 1D Ising chain

\[ H = -J \sum_{i=1}^{N} \sigma(i) \sigma(i+1) - h \sum_{i=1}^{N} \sigma(i) = -\frac{h}{2} \left( \sum_{i=1}^{N} \sigma(i) + \sigma(i+1) \right) \]

\[ P = \sum_{\sigma} \left[ \prod_{i=1}^{N} \exp \left( \beta \sum_{i} \sigma(i) \sigma(i+1) + h \sigma(i) \right) \right] \]

\[ Z = \text{Tr}(\hat{P}^N) = \lambda_0^N + \lambda_2^N \]

\[ F = -NkT \ln \cosh(\beta) + \sqrt{e^{2\beta} \sinh^2(\beta) + e^{-2\beta}} \]

Exact results in 1D:

\[ m = <\sigma(1)> = -\frac{1}{N} \frac{\partial F}{\partial h} = \frac{\sinh(\beta)}{\sqrt{\sinh^2(\beta) + e^{-2\beta}}} \]

For h=0:

\[ <\sigma(i)\sigma(i+j)> = \tanh(\beta) \]

\[ \zeta = -\ln(\tanh(\beta)) \]

no phase transition at T>0 !!! \((T_c=0)\)

For \( T\to 0 \) we have that \( \zeta \to \infty \)

A graphical solution:

\[ t=1 \text{ the critical point} \quad \rightarrow \text{paramagnetic behavior} \]

\[ t>1 \text{ two solutions}; \quad \begin{cases} <S>=0 \text{ (unstable solution)} & \text{for } t=1 \text{ the critical point} \\ <S> \neq 0 \text{ (stable solution)} & \text{for } t>1 \end{cases} \]

\[ t=0 \text{ the only possible solution}; \quad <S> = 0 \rightarrow \text{paramagnetic behavior} \]

\[ T_c = \frac{Jq}{k} \]

Implementing the Metropolis and Glauber Monte Carlo for the 2D Ising model

**Problem:** Study \(<m(T)>, <E(T)>, <C(T)> \text{ and } T_c \) for 2D Ising models by using the Metropolis or Glauber algorithm.

We consider \( B=0 \) and fix \( J=1 \). The units are considered that \( k_B=1 \).

\[ H = -\sum_{\langle i,j \rangle} S_i S_j \]

- Let us assume a lattice \( N \times N \) with free boundary conditions
- We consider a canonical ensemble and fix thus \( N \) and \( T \)
- We would like to calculate:

\[ <m(T)>, <E(T)>, <C(T)> \]

\[ T_c \] will be determined from the maxima of \(<C(T)>\) and \(<\chi(T)>\)
In order to get the desired quantities we have to calculate the following \( N \times N \) dimensional sums (integrals):

\[
\begin{align*}
\langle M(T) \rangle &= \sum_{i,j} \sum_{S_i} \sum_{S_j} \frac{1}{Z} \exp \left( - \frac{H(S_i)}{k_B T} \right) \\
\langle M^2(T) \rangle &= \left( \sum_{i,j} \sum_{S_i} \sum_{S_j} \frac{1}{Z} \exp \left( - \frac{H(S_i)}{k_B T} \right) \right)^2 \\
\langle E(T) \rangle &= - \sum_{i,j} \sum_{S_i} \sum_{S_j} \frac{1}{Z} \exp \left( - \frac{H(S_i)}{k_B T} \right) \\
\langle E^2(T) \rangle &= \left( \sum_{i,j} \sum_{S_i} \sum_{S_j} \frac{1}{Z} \exp \left( - \frac{H(S_i)}{k_B T} \right) \right)^2
\end{align*}
\]

We will use the Metropolis MC method to calculate these sums (integrals):

1. Fix a given temperature
2. Fix an initial spin configuration \( \{q\} \)
3. Calculate the initial value of \( E \) and \( M \)
4. Consider a new spin configuration by virtually “flipping” one randomly selected spin \( \{q'\} \)
5. Calculate the energy \( E' \) of the new configuration, and the energy change due to this spin-flip
6. Calculate the Metropolis (Glauber) \( W(q \rightarrow q') \) probabilities for this move
7. Generate a random number \( r \) between 0 and 1
8. If \( r \leq W(q \rightarrow q') \) accept the flip and update the value of the energy to \( E' \) and magnetization to \( M' \)
9. Repeat the steps 4-8 by collecting the values of \( E, E', M, M' \), and calculate their average
10. Compute this average for a large number of steps
11. Calculate the value of \( \langle m(T) \rangle, \langle E(T) \rangle, \langle C_v(T) \rangle, \langle \chi(T) \rangle \) by the given formulas
12. Change the temperature and repeat the algorithm for the new temperatures as well.
13. Construct the desired \( \langle m(T) \rangle, \langle E(T) \rangle, \langle C_v(T) \rangle, \langle \chi(T) \rangle \) curves

writing the code—see the computer code

A simple 2D Glauber dynamics code

Variation of \( \langle m(T) \rangle \) as a function of \( T \) for different system sizes

Variation of \( \langle C_v(T) \rangle \) as a function of \( T \) for various system sizes

Variation of \( \langle \chi(T) \rangle \) as a function of \( T \) for various system sizes

Estimates for \( T_c \) for various system sizes (step in \( T \) is 0.1)
In order to have time steps satisfying this probability generating uniformly distributed numbers according to the distribution, we need to generate thus q random numbers. 

Motivation for the equation of time

- The C=1 constant results from the \( f(0)=1 \) condition
- If we generate a random number \( u \in [0,1) \) and multiply it by \( R_3=1.1 \), this will correspond to one visit of a point on the line. The probability to get transition \( w(2) \) will be proportional with the distance between \( R_2=0.6 \) and the system)
- Problem: how to update the "time" in order to get the good values of the averages

Simulating dynamic processes

1. Simulating grain-growth by using a T=0 temperature Potts- model (a trivial application)

Grain-growth in metals: growth of the size of the mono-crystalline domains. Small domains are "eaten up" by larger ones, the grain-boundary moves. Experimentally observed that the <d> mean-grain size increases as:

\[
< d(t) > \sim (< d(0)^2 + \alpha t >)^{1/2}
\]

The Potts model:

\[
H = -q \sum_{<i,j>} \delta (\sigma_i, \sigma_j)
\]

\[
\sigma_i(t) = 1, 2, ..., q
\]

- we use a square or triangular lattice
- we start with a large \( q \) number of initial states
- each lattice site will have a randomly chosen Potts variable

- Make a Metropolis or Glauber dynamics at T=0 (accept only steps that do NOT increase the energy of the system)
In the classical Metropolis or Glauber algorithm, one would randomly select between all the lattice sites, and accept those moves that do not increase the energy --> assuming the used Metropolis or Glauber dynamics real, the elapsed time here would be proportional with the number of attempts...one MC step (trial equal with the number of spins) would be considered as a 1-time.

- In the BKL algorithm, we visit ONLY the GRAIN BOUNDARIES, and flip one of this sites in...state equal to it's neighbor! The probability to flip any site on the GB is the same. The time, however, has to be updated according to the BKL scheme.

- If the number of possible GB transitions is $q$, we get that $R=q$.

The increase of $<d(t)>$ follows the expected form:

$$d(t) = \frac{1}{R} \ln(t)$$

2. Simulating dynamics of atoms on crystal surfaces, or/and the deposition of atoms on substrates

- Atoms deposited on surfaces can jump from one site to a nearby site, after escaping from a potential barrier (binding energy to nearby atoms).

The escape probability

$$P = \exp(-\frac{E_b}{kT})$$

The transition rate

$$r = w \cdot \exp(-\frac{E_b}{kT})$$

$w$: is the attempt frequency per unit time, or the frequency of the vibration (of order $10^{12-10^{15}}$Hz)

- The binding energy is proportional with the number of its nearest neighbors.

- Usually there is also a deposition process coexisting with the diffusion one, with a rate $r_{\text{dep}}$ (number of atoms deposited per unit time per unit area).

- Many times there is a co-deposition or co-diffusion of more types of atoms, let’s say A and B.

- Problem: to simulate the dynamics of the deposited atoms, and the formed structures.

**Main difficulty:** the largely different time-scales (transition rates) for the possible stochastic processes → makes the simulation very lengthy in time.

Cluster Algorithms (Swendsen and Wang MC method and the Wolf algorithm)

**Cluster Algorithms:**

- consider a lattice for the substrate

- fix the interaction constants (fix the binding energies to nearby atoms), fix the temperature and deposition rates

- calculate the transition rates for different situations (different species and different number of nearest neighbors)

- update the time in units $t$: smaller than the smallest period for the transitions or deposition

- at each time-step randomly choose as many sites as many atoms are on the substrate, and attempt their transition to a nearby site with the probability $P_r = \exp(-\frac{E_b}{kT})$

- at each time-step deposit atom on a randomly chosen site with probability $P_{\text{dep}} = \exp(-\frac{E_b}{kT})$

- continue the simulation until the desired time

**Problem and solution:** due to the largely different transition rates (effect of the exponential in the transition rate formula) the simulation is very slow, and ineffective. Most of the time nothing happens --> can be much improved by the BKL method!

**Spin-spin correlation and the correlation length**

The correlation between spin $S_i$ and spin $S_j$ is measured by:

$$\langle S_i \cdot S_j \rangle$$

(average is an ensemble average). If two spins are uncorrelated $\langle S_i \cdot S_j \rangle \approx 0$.

- At $T=0$ the spins are lined up in one direction $\rightarrow \langle S_i \cdot S_j \rangle \approx 1$ (this is a trivial correlation, because $S_i$ will hardly affect $S_j$ unless it is not a neighbor of $S_j$)

- Near $T_c$ the situation is very different, the spins are constantly changing, but not independently $\rightarrow$ there are large domains of parallel spins which persist for long periods of time. Thus, spins far apart from one another are strongly correlated

- At high temperatures, the spins fluctuate rapidly but almost independently of each other.

- To describe the real spin correlation (and eliminate the trivial correlation) we define the pair correlation function $\xi(r)$.

- In general $\xi_2$ depends only on the distance between the two spins $g(r)$ (for large systems, r can be considered a continuous variable. For $r >>a$ (lattice spacing) we have:

$$g(r) = \exp\left(-\frac{r^2}{\xi^2}\right)$$

(2/3) diverges at $T_c$ as $\rightarrow \xi(2/3) \rightarrow (r-1)^2$

At $T= T_c$ we obtain thus $\eta$ is another critical exponent ($\eta=1/4$) for the 2D Ising model.

**The critical slowing down (first picture)**

- The Ising model does not have real dynamics built into it (there is no kinetic energy term associated with the spins, so there is no real time evolution)

- In a real dynamical system, the dynamical variables are functions of time. The relaxation time characterizes the time scale over which the system approaches equilibrium. If $\langle A(t) \rangle$ is a quantity that relaxes towards its equilibrium value $A_0$, the relaxation time can be characterized as:

$$\tau = \frac{1}{\langle A(t) \rangle - A_0}$$

- The Monte Carlo method generates successive spin configurations

- Although there is no real time evolution the MC simulation process might be viewed as kind of dynamics...

- We can follow how a configuration of spins relaxes to another configuration as the simulation goes by $\rightarrow$ this defines a relaxation time

- The relaxation time is linked to the correlation length by the dynamical critical exponent, $z$:

$$\tau = \frac{\xi^z}{2} \rightarrow T \rightarrow T_c$$

- If $z=2$ for the Metropolis or the Glauber dynamics $\rightarrow$ the relaxation time diverges more quickly than the correlation length $\rightarrow$ critical slowing down

- There is necessary many MC steps to generate independent (uncorrelated configurations) $\rightarrow$ the sampling is restricted only to a small portion of the state-space

- The system has a long memory....
The relaxation time in Metropolis MC simulations \( \rightarrow \) the autocorrelation time

- The relaxation time in the Metropolis MC characterizes how many MC steps to skip in order to generate statistically independent configurations.
- If the relaxation time is of the order of a single MC step, every configuration can be used in measuring averages.
- If the relaxation time is longer, approximately \( T \) MC steps should be discarded between every point.
- To define a relaxation time we first define the autocorrelation function for a quantity \( A \):
  \[
  c_{\text{aut}}(t) = \langle A_{\text{MC-step}} \rangle - \langle A \rangle^2
  \]
  which defines the exponential correlation (relaxation) time \( (\tau_{\text{rel}}) \)
- Starting from the equations from below one can suggest another correlation time:
  - If the relaxation time is of the order of a single MC step, every configuration can be used in computer simulations.
- The Wolff single cluster algorithm:
  - even more efficient than the S-W algorithm
  - the way of constructing the clusters is different, and the clusters are always correlated spins, they are correlated (realization of a giant spin, flipping it according to a random criterion).
  - The Swendsen and Wang cluster algorithm for \( J > 0 \) Ising model
    - the basic idea is to identify the clusters of like and correlated spins and treat the clusters as a giant spin, flipping it according to a random criterion.
    - It is necessary that the algorithm should lead to an ergodic Markov process and the detailed balance condition is satisfied.
    - The algorithm can be generalized for arbitrary \( J > 0 \) or \( J < 0 \) Potts models
- Construction of the clusters of correlated spins:
  - The simple clusters of like nearest neighbor spins are NOT the clusters of correlated spins, these are too large…
  - \( \{x_{i+1} = x_{i} \} \) are still spins with like orientation, although the correlation between them in this case should vanish.
  - The way of constructing the clusters of correlated spins is to put a link between nearest neighbors and like spins, with a probability \( p = 1 - \exp(-2J/kT) \).
- Critical slowing down (a second look)
  - For finite lattices when \( T \to T_c \), we get \( \tau_{\text{rel}} \to L \) (size of the lattice) \( \rightarrow \) no real divergence.
  - We get thus:
    \[
    \tau \approx \frac{\Lambda}{\epsilon(T)}
    \]
  - For the Metropolis and Glauber dynamics \( z = 2 \), and we get:
    \[
    \tau = L^d
    \]
  - Good news: the value of the dynamical critical exponent is not universal; it depends on the MC algorithm (paper of Swendsen and Wang).
- The problem: elaborate a MC method for which the value of the dynamical critical exponent is smaller \( \rightarrow \) these must be non-local algorithms.
- Cluster algorithms:
  - We flip together all correlated spins.
  - In each flip we generate statistically independent configurations.
  - The value of the dynamic exponent becomes as low as \( z = 0.15 \)!

Flipping the clusters:
- The clusters are flipped with probability \( 1/2 \) (we assign a new common value, +/-1 to all spins in the cluster).
- The spins in the whole lattice are in this manner updated.
- The algorithm satisfies detailed balance \( \rightarrow \) appropriate for important sampling.

The Swendsen and Wang algorithm for the 2D Ising model:
1. Consider a lattice of spins with size \( N \times N \)
2. Fix the parameters \( (T, J = 1, k_B T) \)
3. Consider an initial configuration of the spins
4. Put "virtual bonds" with probability \( 1 - \exp(-2J/k_B T) \) between nearest neighbor and like spins
5. Construct the clusters of correlated spins
6. "Flip" the clusters with probability \( 1/2 \) (this is one MC step)
7. Get the new configuration of the spin system, and count it in the calculation of the desired averages
8. Identify like nearest neighbor spins and repeat the algorithm starting from 4.

Main difficulty: \( \rightarrow \) the construction of the clusters of correlated spins

An example code \( \rightarrow \) program nr. ????? (identification of clusters with recursion).

The Wolff single cluster algorithm:
- Even more efficient than the S-W algorithm.
- Difference: constructing and flipping only one cluster at a time!
- The way of constructing the correlated spins cluster is the same as in the S-W algorithm.

The basic of the Wolff algorithm:
1. Choose a spin randomly in the lattice.
2. Construct the cluster of correlated spins starting from this spin as a "seed" by connecting nearest neighbor and like spins with a probability \( p = 1 - \exp(-2J/k_B T) \).
3. Do this process recursively until the cluster cannot grow more.
4. Flip this cluster of correlated spins (this will be one MC step).
5. Update the time proportionally with the number of flipped spins.
6. Count the new configuration in the average of the desired quantities.

- The Wolff algorithm is more efficient, because we construct only one cluster an always flip it, the probability to choose a cluster is proportional with the size of the cluster \( \rightarrow \) we will usually flip bigger clusters \( \rightarrow \) we generate statistically independent configurations.
- An example program is given as code nr. ?????

A simple visual program to compare the effectiveness of the Metropolis, BKL, and Wolff algorithms: \( \text{MC.exe} \).
Now, we can calculate $f(E, T)$ at any desired temperature

$$f(E, T) = \frac{1}{\Omega(T)} D(E) \exp\left(-\frac{E}{k_B T}\right)$$

The $\Omega(T)$ constant can be obtained from the normalization condition

$$\sum_i \int \Omega(E) \exp\left(-\frac{E}{k_B T}\right) dE = 1$$

- we get thus for any $T$ value:

$$f(E, T) = \frac{\Omega(T)}{\Omega(T)^{\frac{1}{2}}} \exp\left(-\frac{E}{k_B T}\right)$$

Calculating other averages

- we obtain from here two equations:

$$E_i(T) = \int f(E, T) E \, dE = \int \frac{E \Omega(T)}{\Omega(T)^{\frac{1}{2}}} \exp\left(-\frac{E}{k_B T}\right) dE$$

- we can construct thus the $E_i(T)$ for any $T$ temperature and calculate the desired averages

$$<E(T)> = \frac{1}{\Omega(T)} \int f(E, T) E \, dE = \frac{1}{\Omega(T)} \int \frac{E \Omega(T)}{\Omega(T)^{\frac{1}{2}}} \exp\left(-\frac{E}{k_B T}\right) dE$$

Calculating $M(T)$:

$$f(M, T) = \frac{1}{\Omega(T)} D(M) \exp\left(-\frac{M}{k_B T}\right)$$

After substituting $g(M)$ and $E(M)$

$$f(M, T) = \frac{1}{\Omega(T)} \frac{1}{\Omega(T)^{\frac{1}{2}}} \exp\left(-\frac{M}{k_B T}\right) \int \frac{M \Omega(T)}{\Omega(T)^{\frac{1}{2}}} \exp\left(-\frac{M}{k_B T}\right) dM$$

The value of $M$ can be calculated by the normalized condition for $M(T)$

$$\sum_i f(M, T) \Delta M = 1$$

$$f(M, T) = \alpha(T, T_i) \frac{1}{\Omega(T)} \frac{1}{\Omega(T)^{\frac{1}{2}}} \exp\left(-\frac{M}{k_B T}\right) \int \frac{M \Omega(T)}{\Omega(T)^{\frac{1}{2}}} \exp\left(-\frac{M}{k_B T}\right) dM$$

The Quantum Monte Carlo Method

Quantum-statistical systems $\rightarrow$ quantum mechanical systems in contact with a heatbath

Quantum spin models (taking in consideration the quantum nature of the spins)

$$H = -J \sum_{i} \hat{S}_i \cdot \hat{S}_{i+1} - J \sum_{i} \hat{S}_i \cdot \hat{S}_j$$

$$\hat{S}_i = \sin(\theta) \hat{S}_i + \cos(\theta) \hat{S}_i$$

$$\hat{S}_i \cdot \hat{S}_j = \frac{1}{2} \left( \hat{S}_i^2 + \hat{S}_j^2 - \hat{S}_i \hat{S}_j \right)$$

$$\hat{S}_i \cdot \hat{S}_j = \frac{1}{8} \left( \hat{S}_i + \hat{S}_j \right) \cdot \left( \hat{S}_i - \hat{S}_j \right)$$

$$\hat{S}_i \cdot \hat{S}_j = \frac{1}{2} \left( \hat{S}_i^2 + \hat{S}_j^2 \right) - \frac{1}{2} \hat{S}_i \cdot \hat{S}_j$$

$$\hat{H} = -J \sum_{i} \hat{S}_i \cdot \hat{S}_{i+1} - J \sum_{i} \hat{S}_i \cdot \hat{S}_j$$

$$\hat{H} = -J \sum_{i} \left( \hat{S}_i \cdot \hat{S}_{i+1} + \hat{S}_i \cdot \hat{S}_j \right) - J \sum_{i} \hat{S}_i \cdot \hat{S}_j$$
The Hubbard model
- many body hamiltonian in a second quantized form
- describes interacting electrons in the periodic potential of a lattice
- second quantization in Wannier states (quasi-localized) electron states on a given orbital and with a given spin orientation at a given ion at position \( \mathbf{r}_i \)
- Hamiltonian written with the creation \( \hat{c}_i^\dagger \) and destruction \( \hat{c}_i \) operators, acting on states described with occupational number
- interaction is strongly screened by the electron gas, thus it is restricted at the same site

In 1D, and considering jumps only at nearest neighbor lattice sites, the hamiltonian becomes

\[
\hat{H} = -t \sum_{\alpha = \sigma} (\hat{c}_{i+1,\alpha}^\dagger \hat{c}_{i,\alpha} + \hat{c}_{i,\alpha} \hat{c}_{i+1,\alpha}^\dagger) + U \sum_{\alpha = \sigma} \hat{n}_{i,\alpha} \hat{n}_{i,\alpha}^\dagger
\]

\( \hat{n}_{i,\alpha} \) are the particle number operators in state \((\alpha, \sigma)\)

Considering more than one interacting electrons on the same lattice is possible by considering extra terms. \( H_0 \) is an on-site repulsion and \( H_{\alpha} \) is interaction between particles in neighboring cells.

\[
\hat{H}_0 = V \sum_{i \alpha} \hat{n}_{i,\alpha} \hat{n}_{i,\alpha}^\dagger
\]

\[
\hat{H}_\alpha = V \sum_{i \alpha} \hat{n}_{i,\alpha} \hat{n}_{i,\alpha}^\dagger
\]

Taking into account the spin of the particles we can add an additional discrete parameter characterizing the projection of the spin.

Basic idea for the QMC method
- transferring the \( d \)-dimensional quantum-statistical problem in a \( d+1 \) dimensional classical statistical problem
- we consider the system in canonical ensemble

\[
Z = \text{Tr}(\text{exp}(-\hat{H} \beta)) = \sum_{|\{n_i\}|} \text{exp}(-\hat{H} \beta |\{n_i\}|)
\]

First possibility to tackle the problem: diagonalize \( \hat{H} \) and consider only that the Metropolis or other MC schemes for the resulted states

Problem: if there are only 2 possible particles in each of the \( L \) possible cells we have to diagonalize a \( 2^L \times 2^L \) matrix which is practically impossible for \( L > 1 \).

A lattice model for itinerant electrons in 1D
- for 1 electron: \( \hat{H}_0 = -\frac{\hbar^2}{2m} \nabla^2 + V(x) = E \phi \)
- discretizing it on a lattice with \( \Delta x \rightarrow L \) coupled linear equations (L system I the length of the considered space)
- \( \psi_i \) is the medium of \( \psi(x) \) in box \( y \)
- \( V \) is the medium of \( V(x) \) in box \( y \)

\[
\sum_{\alpha = \sigma} \exp(ik(\hat{R}_{i} - \hat{R}_{j}))
\]

\( \alpha \) is the medium of \( \alpha \)

Second possibility: make a classical MC algorithm without diagonalization
- consider an \( \{|n\}\} \) initial configuration
- consider a possible new \( \{|n\}\} \) configuration
- accept the change \( \{|n\} \rightarrow \{|n\}\} \) with probability:

\[
P = \frac{\langle |n\| \exp(\beta \hat{H}) |n\rangle}{\langle |n\| \exp(\beta \hat{H}) |n\rangle} = \frac{\langle |n\| \exp(\beta \hat{H}) |n\rangle}{\langle |n\| \exp(\beta \hat{H}) |n\rangle}
\]

- continue the algorithm until the thermodynamic equilibrium is reached
- collect periodically the relevant data

Third possibility (also solution): to rewrite \( Z \) in a form in which the calculation of the \( P \) probabilities are easy when only a few \( n \) numbers are changed:

we can write thus:

\[
\frac{\langle |n\| \exp(-\beta \hat{H}) |n\rangle}{\langle |n\| \exp(-\beta \hat{H}) |n\rangle} = \frac{\langle |n\| \exp(-\beta \hat{H}) |n\rangle}{\langle |n\| \exp(-\beta \hat{H}) |n\rangle}
\]

\[
\langle |n\| \exp(-\beta \hat{H}) |n\rangle = \langle |n\| \exp(-\beta \hat{H}) |n\rangle
\]

\[
\langle |n\| \exp(-\beta \hat{H}) |n\rangle = \langle |n\| \exp(-\beta \hat{H}) |n\rangle
\]

Ideal would be to do this for the whole \( H_0 \hat{H}_\alpha \). This is not possible however, because \( H_\alpha \) do not commute and have terms acting on the same \( n \) numbers.

- SOLUTION: the Trotter-Suzuki (TS) approximation: if \( A \) and \( B \) are sufficiently SMALL operators:

\[
e^{\hat{A}} \approx e^{\hat{A}} = e^{\hat{A}} e^{\hat{B}} = e^{\hat{A} \hat{B}}
\]

Applying the TS approximation if \( M = 1 \) (integer), so that \( [\hat{H}_0, \hat{H}_\alpha] = [\hat{H}_\alpha, \hat{H}_\beta] = 0 \) is small enough we can write

\[
Z = \sum_{\{n\}} \langle \{n\| \exp(-\frac{\beta}{M} \hat{H}_0 + \hat{H}_\alpha)^{M} |\{n\}\rangle = \sum_{\{n\}} \langle \{n\| \exp(-\frac{\beta}{M} \hat{H}_0) \exp(-\frac{\beta}{M} \hat{H}_\alpha)^{M} |\{n\}\rangle
\]

To write \( Z \) as product of simple terms we insert \( 2M \) complete sets of state vectors

\[
\sum_{\{n\}} \langle \{n\| \exp(-\frac{\beta}{M} \hat{H}_0 |\{n\}\rangle \langle \{n\} | \exp(-\frac{\beta}{M} \hat{H}_\alpha)^{M} |\{n\}\rangle
\]

between the exponential terms
Taking into account that $H_\alpha$ acts only on the occupation numbers $n_i$ and $n_{i+1}$, we can write

- $i$ and $\alpha$ label occupation numbers
- from a 1D lattice $\rightarrow$ 2D lattice
- all $n_i$ numbers are independent!
- due to the original trace $\sum_{n_i=1}^{n_i=\infty} n_i$,

\[ Z = \sum_{n_i=1}^{n_i=\infty} \prod_{i=1}^{M} P_{i,a} \]

where: $\beta = \frac{1}{M}$

\[ P_{i,a} = (n_i, n_{i+1}) = [n_{i+1}; n_{i+1}^{-1}) \]

We transform now the quantum-statistical problem to a classical one!

\[ Z = \sum_{n_i=1}^{n_i=\infty} \exp\left(-\beta \sum_{i=1}^{M} E_i\right) \]

\[ \beta E_i = \ln P_i \]

\[ \Delta E = \sum_{i=1}^{M} E_i \]

- $E_i$ (effective energies) can be calculated from the occupation numbers at sites: $(i,\alpha), (i+1,\alpha), (i+1,\alpha+1)$.
- $\Delta E$ is the value of the physical quantity $\Delta$.

Performing the MC simulation:

- we have four-site interaction between neighboring sites $(i,\alpha), (i+1,\alpha), (i+1,\alpha+1), (i+1,\alpha-1)$ with the values of $i$ and $\alpha$ both even or odd.
- in the $i$ and $\alpha$ space this interaction is represented by a check-board pattern where the interaction is around the dark plaquettes.
- the condition for $P_i,a$ to be satisfy:

\[ n_{i,a} + n_{i,a+1} = n_{i,a+1} + n_{i+1,a+1} \]

acceptable changes: (leading to non-infinite final "energies")
- the occupation numbers for the left side of a white plaquet are increased by unity, and the values for the right side are decreased by unity.
- the occupation numbers for the left side of a white plaquet re decrease by unity and the values from the right side are increased by unity.
- the occupation numbers cannot become negative!

Calculation of the $e_{i,a}$ factors

- in general these "energies" are only numerically calculable, after Taylor expansion of the formula for $P_i$...
- in some cases however, it is analytically possible.
- example: spinless fermions with only nearest neighbor interactions. $P_i,a^{0,1}$

\[
\begin{align*}
E(1,1,1,0) &= \frac{\Delta \tau}{\beta} \\
E(0,0,0,0) &= 0 \\
E(0,1,0,1) &= \frac{1}{\beta} \ln \left[ \cosh \left( \frac{1}{\beta} \right) \right] \\
E(1,0,1,0) &= \frac{1}{\beta} \ln \left[ \sinh \left( \frac{1}{\beta} \right) \right] \\
p &= \frac{\hbar^2}{8 \pi \mu \kappa}
\end{align*}
\]

The Microcanonical Monte Carlo method


- Studies the thermodynamic system not in a canonical but in microcanonical ensemble $\rightarrow E, V, N$ are fixed!
- $S(E,V,N) = k \ln(W)$.
- $W$: the number of microstates corresponding to a macro-state with fixed $E, V, N$.
- All microstates have the same probabilities $\rightarrow$ needing no extra random number generation to compute averages for the given ensemble.

\[
A_i := \sum_{W=1}^{W} A_i \]

- The main problem is how to sample states with fixed $E$ values!
- An immediate possibility would be to randomly try to change the state of the system (using an ergodic Markov chain), and accepting only the changes that keeps the total energy of the system constant $\rightarrow$ the big problem is that many trials are in vain, and are disregarded $\rightarrow$ loosing a lot of computation time!
- The idea for improving this is the "demon algorithm"
The demon algorithm

- We add an extra degree of freedom to the original macroscopic system of interest to model the demon.
- The total energy of the demon plus thermodynamic system (TS) is conserved.
- There is an energy transfer between the demon and the thermodynamic system, the demon can receive energy from the TS and can give energy to the TS. The only constraint is that the demon energy cannot become negative.
- We realize the following ergodic dynamics to map the micro-canonical ensemble of the TS (specific case of the Ising model):

  - Choose an initial spin configuration with the desired energy
  - Choose another spin configuration by flipping one spin
  - Compute $\Delta E$, the change in the energy of the system due to the considered change
  - If $\Delta E < 0$, the system gives the amount $|\Delta E|$ to the demon and the trial configuration is accepted.
  - If $\Delta E > 0$ and the demon has sufficient energy for this change then the demon gives the necessary energy for the system and new configuration is accepted. If the demon has not enough energy the new trial configuration is rejected.
  - The above steps are continued until a representative sample is obtained, and averages on this are calculated.

Since the demon is only one degree of freedom in comparison to the many degrees of freedom of the system, the energy fluctuations of the system will be of order $1/N$, which is very small for $N \gg 1$, so the calculated average will approximate well the conditions for the micro-canonical ensemble at the fixed energy

The demon energy distribution, the demon as a thermometer

- The demon in many respects acts as an ideal thermometer since it is only one degree of freedom in comparison to the many degrees of freedom system (TS) with which it changes energy.
- From classical statistical physics we know, that in the considered setup, the probability that the energy of the demon at a given step is $E_d$ is:
  
  $$ P(E_d) = \frac{1}{Z} \exp\left(-\frac{E_d}{kT}\right) $$

  where $Z = \sum_{E_d} \exp\left(-\frac{E_d}{kT}\right)$ is the partition function for the demon.

- One way to characterize this "temperature" that the demon (thermometer) experiences is by plotting $\ln[P(E_d)]$ as a function of $E_d$. The slope will give indication for $1/kT$.

- A simpler way to determine the temperature experienced by the demon is by calculating $\langle E_d \rangle$. For the case of the Ising model it can be shown, that:

  $$ \frac{kT}{2} = -\frac{4}{\ln(1 + 4/\langle E_d \rangle)} \quad \text{for } |\langle E_d \rangle| \ll 1 \Rightarrow kT \approx \langle E_d \rangle $$

Problems to solve by microcanonical Monte Carlo

1. Compute $T$ for $N=100$ and $E=-20, -40, -60$ and -80 from the inverse slope of $P(E_d)$ and $\langle E_d \rangle$. Compare your results to the exact results of the infinite one-dimensional lattice $E/N = -\tanh(J/kT)$. Consider 500 MC steps to heat up the system, and than 1000 Monte Carlo steps to study the statistics for $E_d$.

2. Simulate the two-dimensional Ising model on a square lattice by using the demon algorithm. Compute $\langle E_d \rangle$, $\langle M^2 \rangle$ and $\langle |M| \rangle$ as a function of $E$. From here study the susceptibility as a function of the temperature of the system. Consider a lattice with sizes $100 \times 100$, 500 MCS to heat up the system and 1000 additional MC steps to study the necessary averages.