

B 2 Monte Carlo simulations

G.A. Vliegenthart

Institut für Festkörperforschung

Forschungszentrum Jülich GmbH

Contents

1	Introduction	2
1.1	Model systems and interaction potentials	3
1.2	Studying finite system sizes	5
1.3	Random numbers and error estimation	6
2	Statistical Mechanics	8
2.1	Thermodynamic averages and ensembles	8
2.2	Ising model	10
3	Monte Carlo integration	13
3.1	Simple sampling	13
3.2	Importance sampling and the Metropolis rule	14
3.3	The Glauber acceptance rule	17
3.4	Markov chains and the Master equation	17
3.5	Monte Carlo in various ensembles	18
4	'Advanced' Monte Carlo methods	20
4.1	Cluster algorithms and the Ising model	20
4.2	Parallel tempering	23
4.3	Umbrella sampling	26

1 Introduction

As we have seen in chapter B1, classical statistical mechanics deals with the statistical description of many particle systems. In this chapter we briefly discuss the application of the Monte Carlo simulation technique to some selected computational problems of classical statistical mechanics. In chapter A12 the quantum Monte Carlo technique is discussed while chapter B11 deals with some special algorithms that are applied to study protein folding.

The basic idea of the Monte Carlo method is that given a relatively small number of particles (or spins), say $10^2 - 10^6$, the interaction potentials between these particles, the boundary conditions and the thermodynamic ensemble, expectation values of observables like for instance the internal energy can be calculated. These observables are defined as multidimensional integrals (over positions and velocities of all particles) and an estimate of the integrals is obtained by averaging the values of a particular observable for a (finite) number of configurations. The required sequence of configurations has to be generated with a probability density which is defined by the specified thermodynamic ensemble. In Monte Carlo simulations this sequence of configurations is stochastic, rather than the deterministic time evolution of configurations realized in a molecular dynamics simulation. The numerical accuracy of the results depends on the number of configurations involved in the averaging procedure (the more the better) while the finite system size has to be taken into account when making predictions for a system of infinite size, i.e. the thermodynamic limit.

The strength of computer simulation methods lies in the fact that we can make exact (within the model, and within in limited computer time) predictions of bulk, interfacial or single molecule properties once that the interactions are specified. This enables one to test theoretical models by implementing them without further approximations in a computer code. Furthermore, computer simulations can be used to study properties that are difficult or impossible to measure experimentally like for instance the free energy or chemical potential of a system. Finally, simulations allow for a systematic study of the role of different parameters in a model by simply switching them on or off.

The Monte Carlo method has been successfully employed to many different kinds of problems in physics including thermodynamics, structure and dynamics, since the first simulations were employed by Metropolis *et al* in the early 1950's [1]. Due to a constantly increasing capacity and availability of computer power as well as a continuing development of specialized algorithms the Monte Carlo method has become a major tool for physicists and chemists. For those who are interested and want to read more, there are numerous books that deal with Monte Carlo simulations among which [2–6] and references therein. These books discuss the basics as well as specialized techniques for simulating spin systems, polymers, phase equilibria, et cetera.

In this chapter we can by no means cover the whole field of Monte Carlo simulations. We will restrict ourselves to the very basics and give only three examples of slightly more advanced methods. This chapter is organized as follows. In the remainder of this section we will discuss some of the particle interactions that are frequently used as well as some very general concepts in computer simulations that are non-specific to the Monte Carlo method. In section 2 some necessary elementary concepts of statistical mechanics are recapitulated. In section 3 the standard Monte Carlo method is introduced after which in section 4 a few selected more advanced methods are discussed.

1.1 Model systems and interaction potentials

In this section we introduce some commonly used interactions that are employed in coarse grained descriptions of atomic, molecular and colloidal systems. Such coarse grained interactions are useful for studying the generic physical properties of a broad class of materials or for studying general physical trends rather than the properties of specific materials.

Interactions The simplest possible interaction between particles is given by the hard sphere potential, which is zero everywhere, except for interparticle distances r smaller than the diameter σ of a particle,

$$u(r) = \begin{cases} \infty & \text{for } r < \sigma \\ 0 & \text{for } r \geq \sigma. \end{cases} \quad (1)$$

The hard sphere interaction, which is sketched in Fig. 1A, introduces in a simple manner excluded volume interactions between particles. One could think of particles that behave as a group of billiard balls in two or three dimensions. This is a common picture when simulating a liquid at densities and temperatures where further interactions can be neglected. Apart from a model system for the liquid state [7], hard spheres have found an important application in colloid science. It has been demonstrated that the hard sphere interaction gives rise to a fluid-crystal phase transition around a volume fraction of 50% of hard spheres. This phase transition was the subject of a lot of discussion in the early 1950's and was for the first time discovered in computer simulations [8]. In more recent times, this transition was observed experimentally in dense colloidal suspensions [9] of sterically stabilized particles in a good solvent.

In order to reproduce the properties of atomic fluids, a slightly more realistic potential is required. Such a potential should be repulsive for short distances (atoms do not overlap completely), and attractive for larger ones. An example is the Lennard-Jones potential [10] which is given by,

$$u(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] \quad (2)$$

where ϵ is the depth of the attractive well, σ the 'diameter' of the particles and r the center to center distance between the particles. In Fig. 1B the Lennard-Jones potential is shown. The attractive tail ($\sim r^{-6}$) originates from dispersion interactions (van der Waals interactions) between the atoms. The algebraic dependence of the repulsive core ($\sim r^{-12}$) has a historical rather than a physical origin. It is a computationally convenient simplification for the exponential repulsion two atoms experience upon approaching.

A complete overview of possible interactions is beyond the scope of this chapter. Just a glimpse of what can be done is summarized in Fig. 2. Ideal polymers can be modeled by attaching monomers in a string using for instance harmonic springs or a FENE-potential [11]. Excluded volume interactions and/or charges on the chain are then easily incorporated to model more realistic polymers or polyelectrolytes. By making multiple connections between monomers, membranes can be modeled in a polymerized [12] as well as in a fluid state [13]. On a more coarse-grained level one can modify the particle shape to obtain asymmetric particles like rods [14] or discs [15] which can represent liquid crystal molecules, rod-like viruses (tobacco mosaic virus, fd-virus), rod-like inorganic colloids (Boehmite) or plate-like colloids (Gibbsite).

Interaction cut-off Long-ranged interactions like the Lennard-Jones potential are computationally expensive since the state of one particle involves interactions with all the other particles

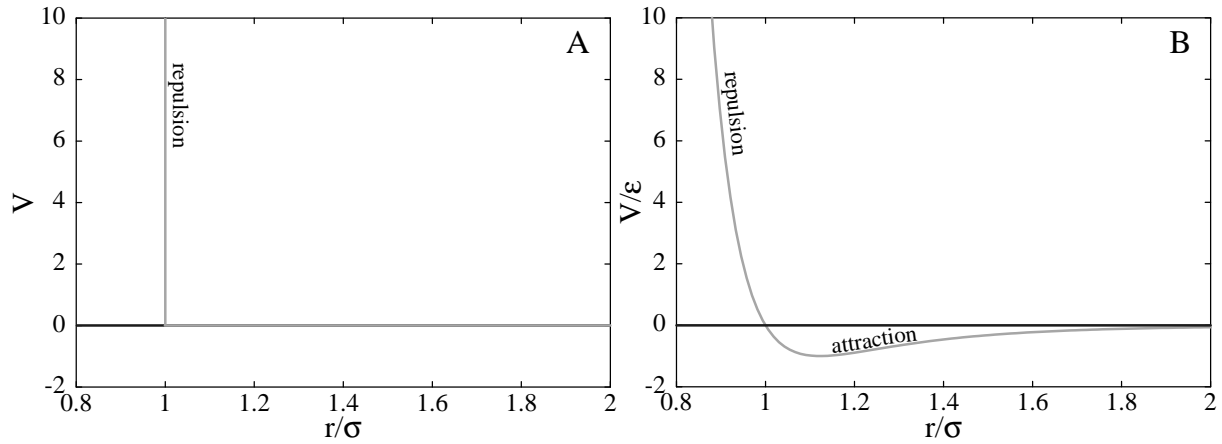


Fig. 1: The hard sphere interaction (A) contains a hard (infinitely repulsive) core while attractions between the particles are absent. The Lennard-Jones potential (B) models long-ranged van der Waals attractions between atoms, on top of a 'soft' repulsive core.

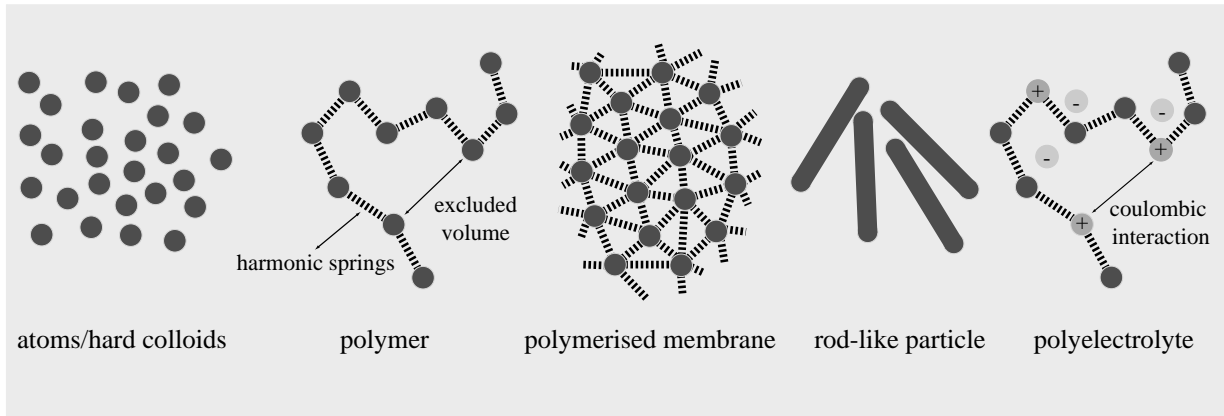


Fig. 2: Sketch of some possible structures combining different interactions and particle shapes.

in the system. Therefore, the interactions are in practice cut off at some distance r_c where the potential has decayed to a small value. In other words, the potential is set to zero for $r > r_c$. In addition, it is convenient to shift the potential such that $u(r_c) = 0$. It is important to realise that thermodynamic properties of a system with a cut-off and shifted potential are in general different from the original system. For instance, the critical temperature of a Lennard-Jones fluid goes down from $k_B T/\epsilon \approx 1.32$ for the full potential to $k_B T/\epsilon \approx 1.09$ for a system with $r_c = 2.5 \sigma$ and which is shifted [16]. This is a decrease of almost 20%.

Units Most of the parameters used in simulations have in principle arbitrary values. However, these numbers should be related with values in real systems such that comparisons can be made. The first step is to choose three independent parameters as a consistent set of units. In this chapter, we have chosen one commonly used option which is to express length in units of σ , mass in units of m and energy in units of ϵ . In this way, time is expressed as $t/\sigma\sqrt{m/\epsilon}$, temperature as $k_B T/\epsilon$, pressure as $P\sigma^3/\epsilon$, force as $F\sigma/\epsilon$ and so on.

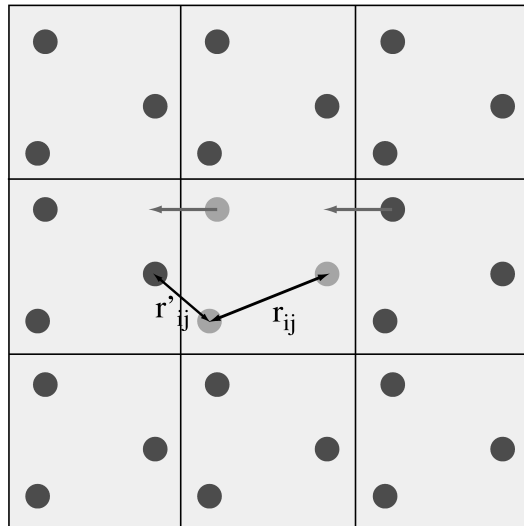


Fig. 3: *Periodic boundary conditions in two dimensions. Indicated are two particles i and j that interact via their minimum distance.*

1.2 Studying finite system sizes

Periodic boundary conditions and the minimum image convention An important aspect of computer simulations are the boundary conditions. In general, one is interested in the bulk behavior of a system and for these cases it is convenient to implement periodic boundary conditions. This means that the central simulation cell has a finite size but that this cell is periodically replicated in all directions. Particles that leave the box at one side, enter the box at the opposite side. The number of particles in a simulation box is thus conserved. This is shown in Fig. 3. In principle each particle in the central cell interacts with all other particles in the central cell as well as with all the periodic images. In practice, for most systems, the interactions are sufficiently short ranged that particles are interacting only with other particles within some cut-off distance. For that case the minimum image convention is employed which selects for each pair of particles the closest distance. Special care has to be taken with intrinsically long-ranged interactions like electrostatic or dipolar interactions. These interactions, which decay slower than r^{-d} with d the dimensionality of the system, lead to divergence of the energy and have to be handled by special methods such as the Ewald summation technique [3, 5, 6, 17] (see also Chapter B3 by R. G. Winkler).

Finite size effects The systems used in computer simulations always contain a finite number of degrees of freedom. Typically, the number of particles or spins is between 10^2 and 10^6 in contrast to 'real' systems where we have order 10^{23} particles. The consequence of this is that the thermodynamic limit is never reached and that we may expect strong dependence of properties on the system size. This is intuitively easily to understand if we consider a system close to a critical point. If a critical point is approached the correlation length increases and diverges right at the critical point. This means that long ranged fluctuations become more and more important. However, in a finite system of box length L the smallest wave vector that fits is $2\pi/L$ and thus the spectrum is cut off. Therefore the expected critical exponents are not observed very close to the critical point but instead mean field behavior is found [18]. Since we are in general

interested in the properties of the 'infinite' system it is important to have a way to extract these properties from finite (and in most cases relatively small) systems. A thorough discussion of finite size scaling methods is given in [6].

1.3 Random numbers and error estimation

Random numbers The Monte Carlo method strongly depends on the ability to generate a large sequence of random numbers. Software generated random numbers are not truly random since they are always produced by a deterministic procedure. One of such algorithms is of the multiplicative type. A multiplier c and initial 'seed' X_0 are chosen after which numbers are generated by,

$$X_n = (c \times X_{n-1} + a_0) \text{MOD} N_{\max} \quad (3)$$

Random number generators of this type are not very good for Monte Carlo simulations since the subsequent numbers are correlated and the period is rather small. An extensive discussion of various types of random number generators as well methods to test random number generators is given in [6, 19]. Popular generators for Monte Carlo are R250 [20] and Mersenne twister [21] for which implementations in C, C++ and Fortran (including parallel versions) are available on the web.

Statistical errors In a computer simulation we cannot (in most cases at least) determine averages exactly since we can only perform the simulation for a finite time. Apart from the average value of a quantity, it is useful to have some estimate of the variance or error in that quantity. Suppose we have a large number n of samples of some fluctuating quantity A that we have measured in a simulation. Assume that the samples were taken after equilibration of the simulation and with some interval. The expectation value $\langle A \rangle$ can be estimated using,

$$\langle A \rangle = \frac{1}{n} \sum_{i=1}^n A_i. \quad (4)$$

An estimate of the variance of A (for independent A_i 's) is given by,

$$\sigma^2(A) = \frac{1}{n} \sum_{i=1}^n (A_i - \langle A \rangle)^2 = \langle A^2 \rangle - \langle A \rangle^2. \quad (5)$$

Whereas if the measurements A_i are uncorrelated the variance of the mean is just,

$$\sigma^2(\langle A \rangle) = \sigma^2(A)/n \quad (6)$$

Equation (6) can be more carefully written by inserting Eq. (4) in Eq. (5),

$$\sigma^2(\langle A \rangle) = \frac{\gamma_0}{n} \left[1 + 2 \sum_{t=1}^{n-1} \left(1 - \frac{t}{n} \right) \frac{\gamma_t}{\gamma_0} \right] \quad (7)$$

with,

$$\gamma_t = \gamma_{i,j}, \quad t = |i - j| \quad (8)$$

and

$$\gamma_{i,j} = \langle A_i A_j \rangle - \langle A_i \rangle \langle A_j \rangle. \quad (9)$$

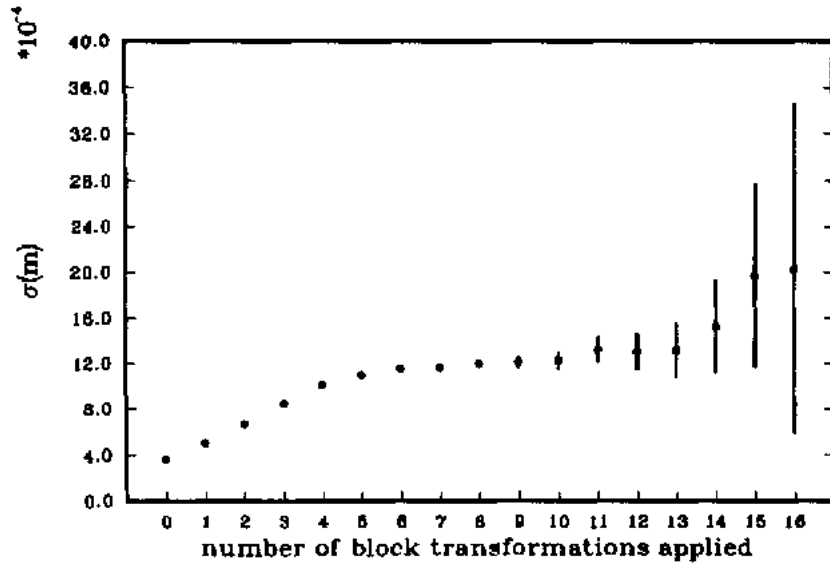


Fig. 4: Block transformation of magnetization data of a 2 dimensional Ising model [22]. After about 6 blocking operations the variance remains constant.

and

$$\gamma_0 = \langle A^2 \rangle - \langle A \rangle^2 \quad (10)$$

The term γ_t/γ_0 is just an autocorrelation function of the data set A_i . If the samples A_i are uncorrelated all the terms γ_t with $t > 0$ in Eq. (7) are zero and the variance is only determined by γ_0 . However, in most cases, the samples A_i are correlated because it is impossible to know the correlation time (and thus the appropriate sampling frequency) in advance. A detailed error analysis would involve the calculation of γ_t but fortunately there is a much more simple method available based on so-called block averaging which was developed by Flyvbjerg and Petersen [22]. Lets start with the original data set A_1, \dots, A_n of correlated data. This data set is now transformed into a new one by the following transformation rule,

$$A'_i = (A_{2i-1} + A_{2i})/2 \quad (11)$$

which generates a data set half as large as the original one but with the same average. Its variance is given by,

$$\sigma^2(A') = \frac{1}{n'} \sum_{i=1}^{n'} \left(A'_i - \langle A \rangle \right)^2 \quad (12)$$

Interestingly, a repeated block transformation of the data leads to a decreasing correlation between the new data A'_i so that the block transformation drives γ_t in Eq. (7) to zero. From the moment that the data are truly decorrelated, the variance $\sigma^2(A')$ reaches a constant value which is exactly the variance that we were looking for. In Fig. 4 we show an example of the standard deviation of the magnetization measured in a 2 dimensional Ising model using Monte Carlo simulation. After approximately 6 block transformations the standard deviation becomes constant at 0.0012. This number of block transformation depends of course on the details of the simulation.

2 Statistical Mechanics

In this section we recapitulate some important concepts of statistical mechanics that we need in Monte Carlo simulation. For a more detailed discussion we refer to Chapter B1 by J.K.G. Dhont as well as to standard textbooks [23, 24].

2.1 Thermodynamic averages and ensembles

Suppose we have a system of N spherical particles in 3 dimensions where each of the particles is characterized by its position \mathbf{r} and momentum \mathbf{p} . The Hamiltonian of the system is given by,

$$\mathcal{H} = \mathcal{K}(\mathbf{p}^N) + \mathcal{U}(\mathbf{r}^N) \quad (13)$$

with \mathcal{K} the kinetic energy and \mathcal{U} the potential energy. The state of the system is described by a single $6N$ -dimensional vector $\Gamma(\mathbf{p}^N, \mathbf{r}^N)$, the phase-space vector. Since all particles are moving according to Newton's equations of motion, Γ will be a function of time and therefore we write $\Gamma(t)$. The observables in the system will have some dependence on Γ so that we can write for the time average of such an observable $A(\Gamma)$,

$$\langle A \rangle = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t A(\Gamma(t')) dt' \quad (14)$$

The average $\langle A \rangle$ can be differently formulated if we consider the complete set of state points Γ which is also called the ensemble of state points. These state points are distributed in phase space according to a probability distribution function that is determined by the thermodynamic ensemble. If the time evolution of $\Gamma(t)$ is such that all states are visited eventually irrespective of the initial conditions, the system is ergodic and we can replace the time average from Eq. (14) by an ensemble average,

$$\langle A \rangle = \sum_{\Gamma} A(\Gamma) \rho_{ens}(\Gamma) \quad (15)$$

where the sum runs over all states Γ and $\rho_{ens}(\Gamma)$ is the probability density function for the ensemble. This probability density function acts as a weight function in the averaging procedure. Below we discuss briefly the probability density functions for some commonly used ensembles.

The canonical ensemble The canonical ensemble corresponds to a system with a constant number of particles N in a fixed volume V at temperature T that can exchange energy with a heat bath. The probability density function for the canonical ensemble equals,

$$\rho_{NVT} = \frac{e^{-\beta \mathcal{H}(\Gamma)}}{\sum_{\Gamma} e^{-\beta \mathcal{H}(\Gamma)}} \quad (16)$$

In the classical limit of a continuous distribution of energies the denominator of Eq. (16) transforms into,

$$Q(N, V, T) = \frac{1}{N! h^{3N}} \int d\mathbf{p}^N d\mathbf{r}^N e^{-\beta \mathcal{H}(\mathbf{p}^N, \mathbf{r}^N)} \quad (17)$$

which is called the canonical partition sum. It is related to the Helmholtz free energy via,

$$\mathcal{F} = -k_B T \ln Q(N, V, T) \quad (18)$$

from which all thermodynamic properties can be calculated by taking the appropriate derivatives. The integrals over the momentum and position coordinates in Eq. (17) can be separated after which the integral over the momentum coordinates can be carried out analytically giving,

$$Q(N, V, T) = \frac{1}{N! \Lambda^{3N}} \int d\mathbf{r}^N e^{-\beta \mathcal{U}(\mathbf{r}^N)} \quad (19)$$

with $\Lambda = (h^2/2\pi m k_B T)^{1/2}$ the thermal wavelength. The remaining integral over the positions is called the configuration integral $Z(N, V, T)$.

In the continuum limit, averages in the canonical ensemble are defined as,

$$\langle A \rangle = \frac{\int d\mathbf{r}^N A(\mathbf{r}^N) \exp[-\beta \mathcal{U}(\mathbf{r}^N)]}{Q(N, V, T)} \quad (20)$$

By inspection of Eq. (20) we see that only states with a large Boltzmann factor $\exp[-\beta \mathcal{U}(\mathbf{r}^N)]$ contribute significantly to the average $\langle A \rangle$.

The isothermal-isobaric ensemble The isothermal-isobaric ensemble is specified at constant number of particles, pressure and temperature. In this ensemble the energy and volume of the system are free to fluctuate. In the isothermal-isobaric ensemble the probability density function is equal to,

$$\rho_{NPT} = \frac{e^{-\beta (\mathcal{H} + PV)}}{Q(N, P, T)} \quad (21)$$

with

$$Q(N, P, T) = \frac{1}{N! h^{3N}} \int dV \int d\mathbf{r}^N d\mathbf{p}^N e^{-\beta (\mathcal{H} + PV)} \quad (22)$$

which is related to the Gibbs free energy via,

$$\mathcal{G} = -k_B T \ln Q(N, P, T) \quad (23)$$

Averages in the NPT -ensemble are defined as,

$$\langle A \rangle = \frac{\int dV \int d\mathbf{r}^N A(\mathbf{r}^N) \exp\{-\beta [\mathcal{U}(\mathbf{r}^N) + PV]\}}{Q(N, P, T)} \quad (24)$$

The grand canonical ensemble The grand canonical ensemble fixes the chemical potential of the particles and in this ensemble the number of particles and the energy are allowed to fluctuate. The probability density function is in this case,

$$\rho_{\mu VT} = \frac{e^{-\beta (\mathcal{H} - \mu N)}}{Q(\mu, V, T)} \quad (25)$$

with partition sum,

$$Q(\mu, V, T) = \sum_{N=0}^{\infty} \frac{e^{\beta \mu N}}{N! h^{3N}} \int d\mathbf{r}^N d\mathbf{p}^N e^{-\beta \mathcal{H}} \quad (26)$$

where μ is the imposed chemical potential. The thermodynamic potential corresponding to the grand canonical ensemble is the so-called q -potential (after Kramers),

$$-PV = -k_B T \ln Q(\mu, V, T) \quad (27)$$

Averages in the μVT ensemble are defined as,

$$\langle A \rangle = \frac{\sum_{N=0}^{\infty} \frac{\exp(\beta \mu N)}{N! \Lambda^{3N}} \int d\mathbf{r}^N A(\mathbf{r}^N) \exp[-\beta \mathcal{U}(\mathbf{r}^N)]}{Q(\mu, V, T)} \quad (28)$$

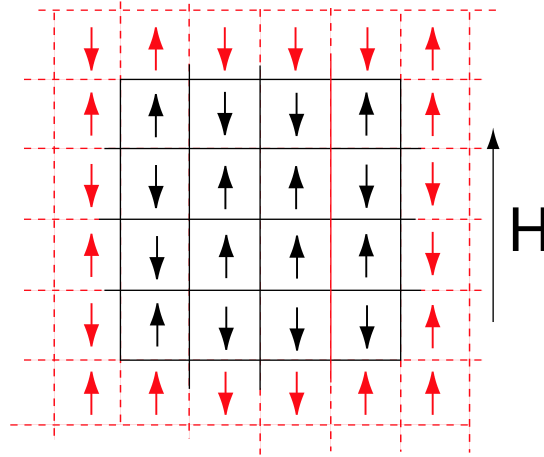


Fig. 5: Spins on a two-dimensional periodic cubic lattice. The central cell is shown in black, parts of the periodic images are indicated in red.

2.2 Ising model

The Ising model is a lattice model for ferromagnets that was presented by Lenz in 1920 [25]. In the past 85 years, a humongous amount of research resources have been invested in the Ising model and with that it has become really one of the *Drosophila melanogasters* of theoretical physics. The model, and the fact that there are exact solutions for closely related models and sub-classes of the model (changing lattice symmetry, dimensionality, external fields, interaction range et cetera) has been proven invaluable to check different kinds of aspects of phase transitions and critical phenomena. In this section we briefly discuss some general features of the Ising model. In a later stage we will use the Ising model as an example to show how Monte Carlo methods can be applied to real problems.

Consider a two-dimensional cubic lattice of N sites. At each lattice site a spin is defined that can either point up or down, $s_i = \pm 1$. The spins can interact with neighbouring spins with a coupling constant J and, with an external magnetic field H . This expressed in the Hamiltonian,

$$\mathcal{U}_\nu = - \sum_{i=1}^N H s_i - J \sum'_{ij} s_i s_j \quad (29)$$

where \mathcal{U}_ν is the energy of a particular configuration ν and s_i the value of a spin i , J is the coupling constant between neighbouring spins. The primed sum in Eq. (29) runs over nearest neighbours only. Periodic boundary conditions are applied which means that the system is semi-infinite, i.e. surface effects are absent but the maximum correlation length is bounded by the size of the central cell. The model is schematically illustrated in Fig. 5.

The physical picture is that of an array of magnets that are obliged to align parallel to one another and parallel to the field H – the model is a ferromagnet. If we set $J < 0$, the neighbouring spins try to align anti-parallel and we have an anti-ferromagnet. The phase behavior of the Ising model is as usual the result of a competition between energy that tries to align spins and entropy that tries to disorder the spins.

Although the Ising model is simple in its structure, it was only solved by Ising in 1925 for the one-dimensional case [26]. Ising did not find no phase transition in one dimension and then (wrongly) concluded that there would be also no phase transition in higher dimensions. It

then it took 20 years until Onsager solved the two-dimensional Ising model with zero external field [27] and only in 1989, Zamalodchikov solved the two-dimensional case with external field [28]. The exact solution in three dimensions is still unknown. For the general case we can write the partition function,

$$Z = \sum_{\nu} \exp(-\beta \mathcal{U}_{\nu}) \quad (30)$$

which is in fact a configurational integral since momenta play no role. The number of states in the Ising model is finite, and grows exponentially ($\sim 2^N$) with the number of lattice sites. The free energy is given by,

$$\mathcal{F} = -k_B T \ln Z. \quad (31)$$

From the free energy all the relevant thermodynamic quantities can be calculated being, the average magnetization,

$$M = -\frac{\partial \mathcal{F}}{\partial H} \quad (32)$$

the susceptibility,

$$\chi = \frac{\partial M}{\partial H} = -\frac{\partial^2 \mathcal{F}}{\partial H^2} \quad (33)$$

the energy,

$$\mathcal{U} = \mathcal{F} - T \frac{\partial \mathcal{F}}{\partial T} = -T^2 \frac{\partial (\mathcal{F}/T)}{\partial T} \quad (34)$$

and finally the specific heat,

$$C = \frac{\partial \mathcal{U}}{\partial T} \quad (35)$$

Below, we discuss some of the easy examples that can be worked out without great difficulty.

$H = 0, J \neq 0$ Suppose we have N spins arranged on a line which only interact with their direct neighbours. In the case that the external field is zero, the partition function is,

$$Z = \sum_{\nu} \exp \left(\beta J \sum_i^N s_i s_{i+1} \right) \quad (36)$$

which can be reformulated slightly since the spin products of neighbours is limited to $s_i s_{i+1} = a_i = \pm 1$ to give,

$$Z = \sum_{\nu} \exp \left(\beta J \sum_i^N a_i \right) \quad (37)$$

where we have neglected end effects. This can be written as a sum of products,

$$Z = \sum_{\nu} \prod_{i=1}^N \exp(\beta J a_i) \quad (38)$$

and yields the simple result,

$$Z = (e^{\beta J} + e^{-\beta J})^N \quad (39)$$

In search for a phase transition we seek a temperature for which there is a spontaneous magnetization, i.e. a finite magnetization at zero field. This can be done by calculating the partition function in presence of a small external field, then calculate the free energy and from the free

energy the magnetization and then take the limit of zero field. This calculation shows that there is no spontaneous magnetization at any temperature for the one-dimensional Ising model (for the details see Refs. [23, 29]) and thus there is no phase transition in one dimension. This can be understood intuitively as well. Suppose we have a one-dimensional array of spins. If all spins are aligned the energy is at a minimum but the entropy is zero (perfect order). If we now choose one spin and flip all spins at the right of this spin (we make a wall), the energy increases by $2J$ but the entropy increases by $k_B \ln(N-1)$ since there are $(N-1)$ places to put the wall. One can now introduce a second wall which can be put at $(N-2)$ positions giving an energy increase of $2J$ and an entropy gain of $k_B \ln(N-2)$ and so on. In the limit of long chains, the entropy gain associated with the insertion of an extra wall will always out-balance the increase in energy and thus the system will fully randomize due to entropic effects. Therefore there is no phase transition in one dimension.

$H \neq 0, J = 0$ In the case that $J = 0$ the partition sum can be evaluated easily and equals,

$$Z = (e^{-\beta H} + e^{\beta H})^N = \cosh(\beta H)^N \quad (40)$$

The free energy (compare Eq. (18)) then equals $\mathcal{F} = -k_B T \ln Z$ and from that the energy,

$$\mathcal{U} = -T^2 \frac{\partial(\mathcal{F}/T)}{\partial T} = -NH \tanh(\beta H) \quad (41)$$

and magnetization

$$M = -\frac{\partial \mathcal{F}}{\partial H} = N \tanh(\beta H). \quad (42)$$

can be calculated. We will encounter this example again in the next section.

Phase transitions and the Ising model As mentioned before, the phase behavior of the Ising model is determined by competition between entropy and energy. At high temperatures, entropy wins and a disordered state with zero magnetization is found. At low enough temperature, the energy contribution to the free energy dominates and the system shows a spontaneous magnetization M . At the critical point, a transition from a disordered to ordered state is found and this phase transition is second order. This means that the order parameter (in this case the magnetization which is the first derivative of the free energy with respect to the external field) changes continuously when we cross the critical point (see Fig. 6). It is the second derivative of the free energy with respect to the field, the magnetic susceptibility, that changes discontinuously with the field. This in contrast first order transitions where one finds a discontinuous change of the order parameter (the first derivative of the free energy with respect to the chemical potential). The thermodynamic quantities given above diverge at the critical point with exponents that depend only on the dimensionality and internal symmetries of the Hamiltonian and not on microscopic details of the model. The divergent behavior at the critical point of for instance the correlation length results in what is called critical slowing down. Relaxation processes that require the break-up of large domains get slower and slower as the size of these domains increases.

A much more complete discussion of the Ising model and related models can be found in standard statistical mechanics text books [23, 24, 30, 31] or in one of the many papers.

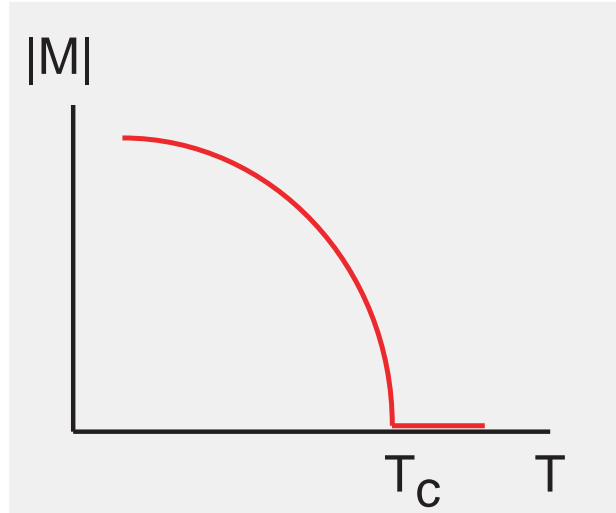


Fig. 6: Artist impression of the spontaneous magnetization as a function of temperature in the Ising model. For $T = T_C$ there is a second order phase transition from the disordered to the ordered state. Spontaneous symmetry breaking results in a finite magnetization that is either positive or negative.

3 Monte Carlo integration

In this section the fundamental concepts of the Monte Carlo method are introduced and illustrated by the Ising model and the hard sphere model.

3.1 Simple sampling

Using simple or random sampling Monte Carlo, integrals like,

$$I = \int_a^b f(x) dx \quad (43)$$

can be approximated as,

$$\frac{b-a}{M} \sum_{i=1}^M f(x_i) \quad (44)$$

by evaluating $f(x)$ at a large number M of randomly chosen points x_i in the interval $[a, b]$. This approach works in general quite well for low dimensional integrals and if the integrand is a smooth and not a too sharply peaked function. Let us now see what happens if we apply this random sampling to some physically interesting models.

Application to the Ising model Consider the Ising model for a very small number of spins. If we want to calculate the partition function in absence of interactions between the spins Eq. (40) we should be able to generate all possible configurations. For very small numbers of spins, all configurations and therefore all properties of the system can be determined exactly. For a moderate number of spins, say $N = 25$ – there are then $\approx 4 \cdot 10^7$ configurations – one can attempt to obtain results from a brute force Carlo simulation. This can be done as follows.

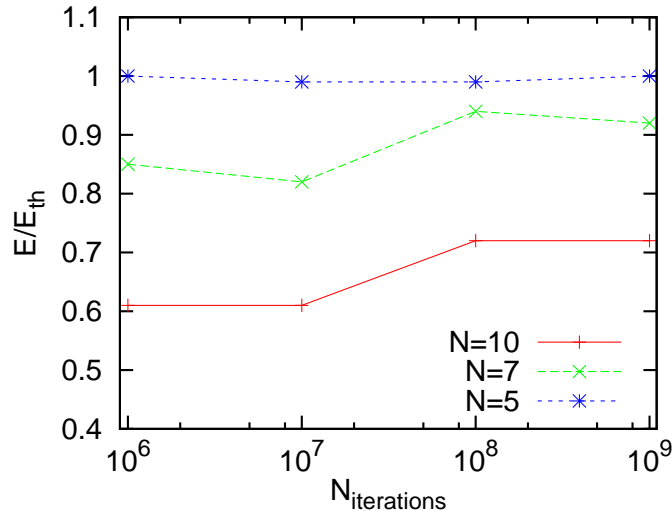


Fig. 7: Energy as a function of the number of Monte Carlo steps.

Generate a random configuration ν by assigning random values to all the spins and calculate the total magnetization M . Repeat this procedure at least 2^{25} times to ensure that all possible states have been visited. The average energy is calculated using Eq. (17). In Fig. 7 some results are shown for the brute force MC simulation with $J = 0$ and $\beta H = 1$ for $N = 5, 7$ and 10 . Only for the smallest system we see that the energy converges to the theoretical value $\mathcal{U}_{th} = -NH \tanh(\beta H)$. Clearly, the above described Monte Carlo procedure is doomed to fail for any serious calculation, simply because the number of configurations that contributes significantly to the average is small and cannot be adequately sampled in this way.

Application to a hard sphere fluid For a system of hard spheres it is even more obvious that simple sampling cannot work. Suppose we have a (not too dilute) fluid with interactions as described in Eq. (1). For such a system the partition function is zero whenever two particles overlap. By generating configurations randomly, the vast majority of configurations will contain one or more pairs of particles that overlap and therefore by far most of the Boltzmann factors generated in this way are zero.

3.2 Importance sampling and the Metropolis rule

In order to improve on random sampling Metropolis *et al* [1] introduced what is called importance sampling, configurations are generated proportional to their Boltzmann weight. This guarantees that the phase space is sampled often where the Boltzmann factor is large and less frequent when the Boltzmann factor is small. Let us now derive the Metropolis scheme for determining the transition probabilities from a state $\mathbf{r}^N = o$ ($o = \text{old}$) to another state $\mathbf{r}^N = n$ ($n = \text{new}$) and thus for generating a sequence of states obeying the Boltzmann distribution. State o has a Boltzmann factor given by $\exp[-\beta \mathcal{U}(o)] / Q(N, V, T)$, where $\mathcal{U}(o)$ is the potential energy of this configuration and $\beta = 1/k_B T$ the prescribed thermal energy. In equilibrium there is no net flow between states o and states n . This means that, *on average*, the number of accepted trial moves resulting in leaving states o must be exactly balanced by the number of accepted trial moves arriving in states o from *any other state* n (see Fig. 8). In practice, a

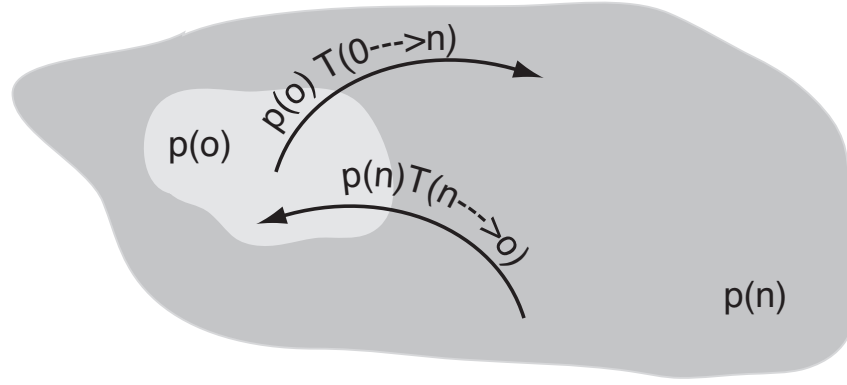


Fig. 8: In equilibrium, the flow from an old state (o) is balanced by the flow from all new states (n).

stronger condition is imposed, namely that in equilibrium the number of accepted trial moves from a state o to a particular state n should be exactly balanced by the number of accepted trial moves from a particular state n to o . This condition is called detailed balance and can be written as,

$$p(o)T(o \Rightarrow n) = p(n)T(n \Rightarrow o) \quad (45)$$

where $p(\alpha)$ is the probability to be in state α and $T(\alpha \Rightarrow \beta)$ denotes the transition probability to go from a state α to a state β . The subtle difference is thus that in equilibrium the net flow from each state is zero (which means that there can be still flow from this state to a specific other state under the condition that it is balanced by flow backwards from other states) while detailed balance implies that the net flow from each state to each other state is exactly zero.

The transition probability is in itself the product of two processes, the creation of a trial move $C(o \Rightarrow n)$ and the acceptance of this trial move $A(o \Rightarrow n)$,

$$T(\alpha \Rightarrow \beta) = C(\alpha \Rightarrow \beta)A(\alpha \Rightarrow \beta) \quad (46)$$

In many Monte Carlo applications the creation of a trial move is a symmetric process, i.e. the creation of the forward and backward moves have the same probability and thus $C(\alpha \Rightarrow \beta) = C(\beta \Rightarrow \alpha)$. Using this form, inserting Eq. (46) in Eq. (45) and using the Boltzmann weight $p(\alpha) = \exp[-\beta \mathcal{U}(\alpha)]$ gives,

$$\frac{p(n)}{p(o)} = \frac{A(o \Rightarrow n)}{A(n \Rightarrow o)} = \exp\{-\beta [\mathcal{U}(n) - \mathcal{U}(o)]\} \quad (47)$$

There are many possible choices for $A(o \Rightarrow n)$ that fulfill Eq. (47) but one of the most efficient and most commonly used is the Metropolis rule,

$$\begin{aligned} A(o \Rightarrow n) &= \exp\{-\beta [\mathcal{U}(n) - \mathcal{U}(o)]\} & \text{if } \mathcal{U}(n) > \mathcal{U}(o) \\ &= 1 & \text{if } \mathcal{U}(n) \leq \mathcal{U}(o) \end{aligned} \quad (48)$$

This can be written in short as,

$$A(o \Rightarrow n) = \min(1, \exp\{-\beta [\mathcal{U}(n) - \mathcal{U}(o)]\}) \quad (49)$$

In practice, a Monte Carlo translational move is performed as follows.

1. select a particle i at random
2. calculate the present energy $\mathcal{U}_i(o)$ of particle i
3. move particle i randomly

$$\mathbf{r}_i(n) = \mathbf{r}_i(o) + \text{RAN}\Delta R$$

4. calculate the new energy $\mathcal{U}_i(n)$ of particle i
5. accept/reject the move according to Eq. (49)

Here RAN is a random number in the interval $\{-1 \dots 1\}$ and ΔR the maximum displacement (which is a tunable parameter of the program). The acceptance is implemented by generating a random number RAN in the interval $\{0 \dots 1\}$, if $\text{RAN} < p(o)/p(n)$ the move is accepted. If $\text{RAN} \geq p(o)/p(n)$ the move is rejected. The magnitude of the parameter ΔR determines the efficiency of the Monte Carlo procedure. If ΔR is large, many of the trial moves are rejected, on the other hand, if ΔR is too small we will only sample phase space very slowly. A good choice for ΔR is a value such that half of the trial moves are accepted.

Now that the translational move of the particles is defined we can set up a general scheme for a *NVT* Monte Carlo simulation,

1. initialise
 - generate N particle positions in a volume V
 - set temperature $k_B T$
2. equilibrate
 - do $i = 1, N_{\text{eq}}$
 - move particles
 - end do
3. main loop
 - do $i = 1, N_{\text{pr}}$
 - move particles
 - collect data
 - end do
4. finish
 - calculate averages
 - write output

The simulation starts an initialization step in which the initial particle positions, and all other relevant parameters (temperature, density ...) are set. Subsequently, the system is equilibrated

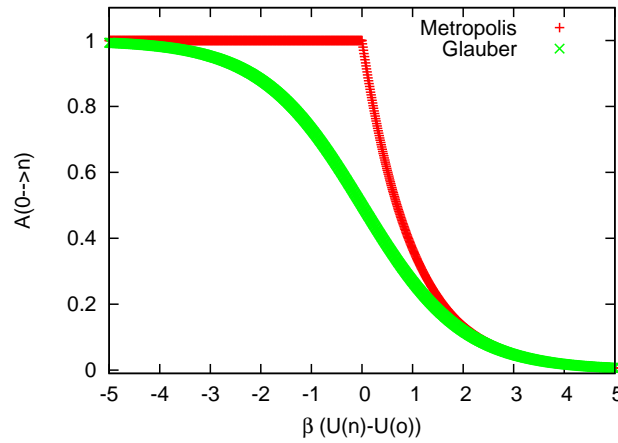


Fig. 9: Metropolis and Glauber acceptance probability.

for a number of integration steps so that the initial configuration (which could be ordered for instance) relaxes towards some state representative for the equilibrium state. This is typically monitored by looking at the internal energy, pressure or paircorrelation function. Then, the real simulation starts and a large number of configurations is created. Every n_{samp} iterations the relevant observables are sampled and stored. Finally, the simulation terminates by calculating averages, errors and writing output and the final configuration. If necessary the simulation can be restarted from this final state to improve statistics.

3.3 The Glauber acceptance rule

As mentioned in the previous section, the Metropolis acceptance rule is not unique since there are many possible choices that would fulfill Eq. 47. An alternative to Metropolis rule is the so-called Glauber rule [32],

$$A(o \Rightarrow n) = \frac{1}{2} \left(1 - \tanh \left\{ \frac{\beta}{2} [\mathcal{U}(n) - \mathcal{U}(o)] \right\} \right) \quad (50)$$

Figure 9 shows both the Metropolis and Glauber acceptance probability as a function of the energy difference between the new and the old state. Whereas the Metropolis rule accepts all trial moves for which the energy of the new configuration is lower than the energy of the old configuration, the Glauber rule is symmetric around $\Delta\mathcal{U} = 0$. In the case of spin systems, Metropolis sampling can give rise to non-ergodic behavior at high temperatures. As at high temperatures all spin flips are accepted, the system jumps back and forward between only two states.

3.4 Markov chains and the Master equation

The sequence of states generated by the Monte Carlo method is a so-called Markov chain. This is a sequence of states that result from a stochastic process and in which each state depends only on the immediate predecessor. The 'dynamics' of a Markov chain where we identify each configuration by a Monte Carlo time can be described by a master equation,

$$\frac{\partial p_n(t)}{\partial t} = - \sum [p_n(t)T(n \Rightarrow m) - p_m(t)T(m \Rightarrow n)] \quad (51)$$

where $p_n(t)$ is the probability to be in state n at time t and $T(n \Rightarrow m)$ is the transition rate to go from state n to m . In equilibrium the net flow from each state is zero (see Fig. 8) and so

$$p_n(t)T(n \Rightarrow m) = p_m(t)T(m \Rightarrow n). \quad (52)$$

which is known as the condition of detailed balance.

If we now use $p_n(t) = \exp(-\beta \mathcal{U}_n)/Q(N, V, T)$ we arrive exactly at Eq. (47).

3.5 Monte Carlo in various ensembles

Isothermal-isobaric Monte Carlo The first isothermal-isobaric Monte Carlo simulations were performed by Wood [33] on hard disks and McDonald [34] on Lennard Jones particles. The important difference with a canonical simulation is that in an isothermal-isobaric simulation, the volume and thus the density of the system fluctuates.

In the NPT ensemble the configurational average is given by Eq. (24) which we can rewrite slightly different as [3, 5],

$$\langle A \rangle = \frac{\int dV V^N \exp(-\beta PV) \int ds^N \exp[-\beta \mathcal{U}(s^N)]}{Z(N, P, T)} \quad (53)$$

with $Z(N, P, T)$ the isothermal-isobaric configuration integral. Here V is the volume of the fluid and s indicates the set of scaled coordinates $s = \mathbf{r}/L$ with $L = V^{1/3}$ the linear dimension of the (cubic) simulation box. The factor V^N stems from the scaling of the coordinates $s = \mathbf{r}/L$. The Metropolis scheme is implemented by generating a sequence of states with limiting distribution,

$$\exp\{-\beta [PV + \mathcal{U}(s^N)] + N \ln V\} \quad (54)$$

This sequence of states is obtained by performing trial moves, for which there are two kinds the first being translational trial moves of particles

$$s(n) = s(o) + \text{RAN} \Delta \mathbf{S} \quad (55)$$

which are accepted with probability,

$$A(o \Rightarrow n) = \min(1, \exp\{-\beta [\mathcal{U}(n) - \mathcal{U}(o)]\}). \quad (56)$$

Here $s(o)$, $s(n)$ indicate the old and new position of a particle and $\mathcal{U}(o)$ and $\mathcal{U}(n)$ the energy of the old the new configuration.

Secondly there are trial volume changes,

$$L_n = L_o + \text{RAN} \Delta L. \quad (57)$$

which are accepted with probability,

$$A(L_o \Rightarrow L_n) = \min[1, \exp(-\beta \Delta H_{on})] \quad (58)$$

where

$$\Delta H_{on} = \mathcal{U}(n) - \mathcal{U}(o) + P(V_n - V_o) - N\beta^{-1} \ln V_n/V_o \quad (59)$$

where L_n , L_o the new old and new length of the simulation cell and V_n and V_o indicate the new and the old volume. RAN is again a random number between $\{-1 \dots 1\}$ and $\Delta \mathbf{S}$, ΔL are the

maximal displacement and change in box length. In order to establish volume changes one can alternatively make a random walk in $\ln V$ instead of in V itself which changes the factor N in Eq. (59) to $N + 1$ [35]. This kind of move is computationally more convenient.

The extra computational expense of an NPT simulation compared to a NVT simulation comes from the volume change which requires an evaluation of the energy of all particles. For potentials of the form $1/r^n$ or of combinations thereof, the energy of the system with a new box size L_n can be obtained using a simple exact scaling relation. For the Lennard Jones potential (Eq. 2) this scaling relation can be obtained using the following expressions for the total energy,

$$\mathcal{U}(L) = 4\epsilon \sum_i \sum_{j>i} \left(\frac{\sigma}{L s_{ij}} \right)^{12} - 4\epsilon \sum_i \sum_{j>i} \left(\frac{\sigma}{L s_{ij}} \right)^6 \quad (60)$$

$$= \mathcal{U}_{12}(L) - \mathcal{U}_6(L) \quad (61)$$

where $\mathcal{U}_{12}(L)$ and $\mathcal{U}_6(L)$ denote the contributions of $1/r^{12}$ and $1/r^6$ terms to the total energy in a box with length L . The 'new' energy for a box that is scaled to have length L_n , the energy simply is,

$$\mathcal{U}(L_n) = \mathcal{U}_{12}(L_o) \left(\frac{L_o}{L_n} \right)^{12} + \mathcal{U}_6(L_o) \left(\frac{L_o}{L_n} \right)^6 \quad (62)$$

If we store the two contributions to the total energy separately we can always calculate the new energy after scaling the box in one step instead of having to perform a double loop over all particles.

Grand canonical Monte Carlo In the grand canonical ensemble the volume, temperature and chemical potential of the particles are fixed. The last property is implemented by letting the system being in contact with a reservoir of particles that is at fixed chemical potential μ . The method was first implemented by Norman and Filinov [36] and later extended by a number of other groups (see for an overview Ref. [5] and references therein).

For the grand canonical ensemble, the averages are defined as,

$$\langle A \rangle = \frac{\sum_{N=0}^{\infty} (N!)^{-1} V^N z^N \int d\mathbf{s}^N A(\mathbf{s}^N) \exp[-\beta \mathcal{U}(\mathbf{s}^N)]}{Q(\mu VT)} \quad (63)$$

with $z = \exp(\beta\mu N)/\Lambda^{3N}$ the fugacity and $\Lambda = (h^2/2\pi m k_B T)^{1/2}$. The factor V^N stems from the scaling of the coordinates $\mathbf{s} = \mathbf{r}/L$. The Metropolis scheme is implemented by generating a sequence of states with limiting distribution,

$$\exp\{\beta[\mu N - \mathcal{U}(\mathbf{s}^N)] + N \ln V - 3N \ln \Lambda - \ln N!\} \quad (64)$$

In this simulation there are two types of trial moves, the first being a translational moves of particles,

$$\mathbf{s}(n) = \mathbf{s}(o) + \text{RAN} \Delta \mathbf{S} \quad (65)$$

which is accepted with probability,

$$A(o \Rightarrow n) = \min(1, \exp\{-\beta[\mathcal{U}(\mathbf{s}^N) - \mathcal{U}(\mathbf{s}^N)]\}) \quad (66)$$

The second type of move involves the creation or removal of a particle. A particle is created by randomly generating a position in the simulation box which is accepted with probability,

$$A(N \Rightarrow N + 1) \min \left(1, \frac{V}{\Lambda^3(N + 1)} \exp\{\beta[\mu - \mathcal{U}(\mathbf{s}^{N+1}) + \mathcal{U}(\mathbf{s}^N)]\} \right) \quad (67)$$

Similarly, a randomly selected particle can be removed with acceptance probability,

$$A(N \Rightarrow N - 1) \min \left(1, \frac{N\Lambda^3}{V} \exp\{-\beta[\mu + \mathcal{U}(\mathbf{s}^{N-1}) - \mathcal{U}(\mathbf{s}^N)]\} \right) \quad (68)$$

In order to satisfy detailed balance for particle creation/removal the number of attempts to create or destroy a particle needs to be the same. The insertion of particles is in general effective only for not too large densities so that for concentrated systems other methods have to be used [37].

4 'Advanced' Monte Carlo methods

4.1 Cluster algorithms and the Ising model

After we introduced the Ising model in section 2.2 we have seen in section 3.1 how the Monte Carlo method can be implemented in a straightforward way to calculate the properties of the Ising model far from the critical point. When approaching the critical point, correlations in space build up and finally diverge at the critical point. In order to break up these diverging spatial correlations increasingly more time is required. This phenomenon is known as critical slowing down. In general the time constant for a relaxation process (at the critical point) scales with the system size like $\tau(L) \sim L^z$ where z is the dynamical critical exponents which depends on how the simulation is implemented. For the two-dimensional Ising model and using a local update scheme (Metropolis Monte Carlo) $z \approx 2$. Consequently, the standard Metropolis Monte Carlo technique breaks down as it becomes very inefficient to de-correlate a configuration via single spin flips. In order to tackle the problem of critical slowing down a special class of methods has been developed, the so-called cluster methods of which we will discuss only one, the Swendsen-Wang method [38]. The basic idea of all cluster algorithms is that spins are flipped in groups at the same time which increases the efficiency. The groups of spins are constructed by constructing clusters of connected spins that are in the same state.

Swendsen-Wang method The Swendsen-Wang method is based on the fact that a spin model can be mapped on a percolation model on the same lattice [39,40]. This is conveniently illustrated by a two state Potts model which is apart from constants equal to the Ising model. Let us now derive the algorithm following for a two state Potts model without external field. Suppose we have a Potts Hamiltonian,

$$\mathcal{H} = -J \sum_{(i,j)} (\delta_{ij} - 1) \quad (69)$$

where the sum runs over nearest neighbours. The probability density function for this system is,

$$P(\nu) = \frac{\exp[-\beta\mathcal{H}(\nu)]}{Z} \quad (70)$$

with partition function

$$Z = \sum_{\nu} \exp[-\beta\mathcal{H}(\nu)] \quad (71)$$

where the sum runs over spin states ν .

As mentioned above, the Swendsen-Wang method relies on the mapping of the Potts model on a bond percolation model and thus the aim of the following is to derive the partition function of the percolation model from the partition function of the Potts model.

Suppose we have our original partition function Eq. (71). Now, consider the following reduced Hamiltonian,

$$\mathcal{H}_{lm} = -J \sum_{(i,j) \neq (l,m)} (\delta_{ij} - 1) \quad (72)$$

which denotes a summation over all pairs of spins *except* the pair l, m .

The original Hamiltonian can be written in terms of this reduced Hamiltonian as,

$$\mathcal{H} = \mathcal{H}_{lm} + \begin{cases} 0 & \text{for } l = m \\ J & \text{for } l \neq m \end{cases} \quad (73)$$

We can now rewrite the original partition sum in terms of the reduced hamiltonian,

$$Z = \sum_{\nu} \exp(-\beta \mathcal{H}_{lm}) \delta_{lm} + \exp[-\beta (\mathcal{H}_{lm} + J)] (1 - \delta_{lm}) \quad (74)$$

where the delta functions serve to select only terms with equal spins $s_l = s_m$ or terms with different spins $s_l \neq s_m$. Rewriting Eq. (74) by reordering terms yields,

$$Z = \sum_{\nu} \exp(-\beta \mathcal{H}_{lm}) \{ \delta_{lm} [1 - \exp(-\beta J)] + \exp(-\beta J) \} \quad (75)$$

which can be simplified to

$$Z = [1 - \exp(-\beta J)] \sum_{\nu} \exp(-\beta \mathcal{H}_{lm}) \delta_{lm} + \exp(-\beta J) \sum_{\nu} \exp(-\beta \mathcal{H}_{lm}) \quad (76)$$

The notation can now be made a little bit more compact; using $p = (1 - \exp(-\beta J))$ and $(1 - p) = \exp(-\beta J)$ Eq. (76) reduces to,

$$Z = p \sum_{\nu} \exp(-\beta \mathcal{H}_{lm}) \delta_{lm} + (1 - p) \sum_{\nu} \exp(-\beta \mathcal{H}_{lm}) \quad (77)$$

In Eq. (77) we see that Z is the sum of two terms, the first is a 'restricted' sum over states and it comes with a pre-factor p while the second is a full sum which comes with a pre-factor $(1 - p)$. We can iterate the process of excluding pairs of same spins in the partition sum so let us next consider another pair k, n . By applying the same procedure again we obtain,

$$Z = p^2 Z_{(l,m),(k,n)}^{S,S} + p(1 - p) Z_{(l,m),(k,n)}^{S,I} + (1 - p)p Z_{(l,m),(k,n)}^{I,S} + (1 - p)^2 Z_{(l,m),(k,n)}^{I,I} \quad (78)$$

where

$$Z_{(l,m),(k,n)}^{S,S} = \sum_{\nu} \exp(-\mathcal{H}_{lm,kn}) \delta_{lm} \delta_{kn}$$

and

$$Z_{(l,m),(k,n)}^{S,I} = \sum_{\nu} \exp(-\mathcal{H}_{lm,kn}) \delta_{lm}.$$

Similar expressions are obtained for the other terms in Eq. 78

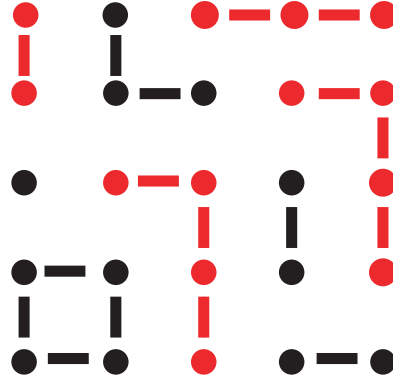


Fig. 10: *Counting bonds and clusters. This particular configuration contains $N_p = 40$ possible bonds, $N_b = 17$ real bonds and $N_c = 9$ clusters (a single isolated site also counts as a cluster).*

Recursive application of this procedure for (all) other pairs gives a binomial expression in which each term is a product of p 's and $(1 - p)$'s and a restricted sum of Z . Since by recursive application all neighbouring pairs are taken out of the hamiltonian \mathcal{H} , there are no terms left in any of the exponentials and the Z 's contain just sums of delta functions. These sums turn out to count the number of possible bond configurations. For a Potts model with q states this term is q_c^N with N_c the number of clusters in the specific configuration of bonds. We can now replace the original partition sum by,

$$Z = \sum_{\gamma} p^{N_b} (1 - p)^{N_p - N_b} 2^{N_c} \quad (79)$$

where N_p is the total number of nearest neighbours and N_b the number of bonds in the configuration. Figure 10 illustrates how to count bonds and clusters. In this configuration of 25 spins there is a total of $N_p = 40$ possible bonds, there are 9 clusters (unconnected spins count as a single cluster) and there are in total 17 bonds.

In order to create a new configuration using the bond model we first draw all the bonds between neighbours that have the same spin (Fig. 11 A). We then delete within each of the clusters each of the bonds with probability $\exp(-\beta J)$ (Fig. 11 B). Finally, each of the clusters is assigned a new spin value at random (Fig. 11 C). The process repeats now from the start on with creating clusters of equal spin value (Fig. 11 D). The sequence of states in (Fig. 11 D-E-F) illustrates that this procedure satisfies microscopic reversibility. The transition probabilities to go from A to C are given by the product of the probability to be in A , to go from A to B and to go from A to C . We can then immediately write down the condition of detailed balance,

$$\begin{aligned} T(A \Rightarrow C) &= P(A)T(A \Rightarrow B)T(B \Rightarrow C) = \\ T(D \Rightarrow A) &= P(D)T(D \Rightarrow E)T(E \Rightarrow A) \end{aligned} \quad (80)$$

The probabilities $T(B \Rightarrow C)$ and $T(E \Rightarrow A)$ are equal (in these transition states we have always the same number of clusters to have new spins assigned) so that,

$$\begin{aligned} T(A \Rightarrow C) &= P(A)T(A \Rightarrow B) = \\ T(D \Rightarrow A) &= P(D)T(D \Rightarrow E) \end{aligned} \quad (81)$$

from which we see that,

$$\frac{P(A)}{P(D)} = \frac{T(A \implies C)}{T(D \implies A)} = \frac{T(A \implies B)}{T(D \implies E)} \quad (82)$$

In order to flip the spins we do the following:

1. create bonds between all spins of same sign
2. destroy bonds with probability p
3. orient all spins in a cluster collectively and randomly
4. remove all bonds and go to 1

Finally, we note that the Swendsen-Wang methods satisfies detailed balance and generates the proper equilibrium distribution. In Fig. 12 we show the dynamical critical exponent as a function of the system size for a two-dimensional Ising model. Using the Swendsen-Wang algorithm, the dynamical critical exponent was reduced from 2.1 (Metropolis) to 0.35 [38], clearly a significant performance boost. Further improvement on this algorithm was established through the so-called Wolf algorithm [41] in which the dynamical exponents could be further reduced to zero. For the interested reader we refer to [6, 31, 41] for more details.

4.2 Parallel tempering

Parallel tempering is a special Monte Carlo simulation technique that was developed to sample systems with many local minima in the free energy [42, 43]. The method is based on carrying out multiple simulations at the same time, but each at slightly different conditions. Apart from the usual trial moves for the individual simulations, there are additional moves, namely the swapping of configurations between the different simulations. A simple example is that of n - NVT simulations that are performed at a slightly different temperature. Apart from the usual displacement steps in each of the separate simulation boxes there are also attempts to switch configurations between systems at a different temperature. The method is not exclusive for swapping configurations at different temperature but can also be implemented to swap configurations at different chemical potential, pressure et cetera.

For now we will just focus on the situation of n NVT systems differing in temperature. The partition function of this extended system is the product of the partition sums of the (independent) sub-systems,

$$Q_{ext.} = \prod_{i=1}^n Q(N, V, T_i) = \prod_{i=1}^n \frac{1}{\Lambda_i^{3N} N!} \int d\mathbf{r}_i^N \exp[-\beta_i \mathcal{U}(\mathbf{r}^N)] \quad (83)$$

where the subscript i indicates the different sub-systems with different temperature.

The acceptance rule for a particle displacement is the same as in the canonical ensemble,

$$A(o \implies n) = \min(1, \exp\{-\beta [\mathcal{U}(\mathbf{r}_n^N) - \mathcal{U}(\mathbf{r}_o^N)]\}) \quad (84)$$

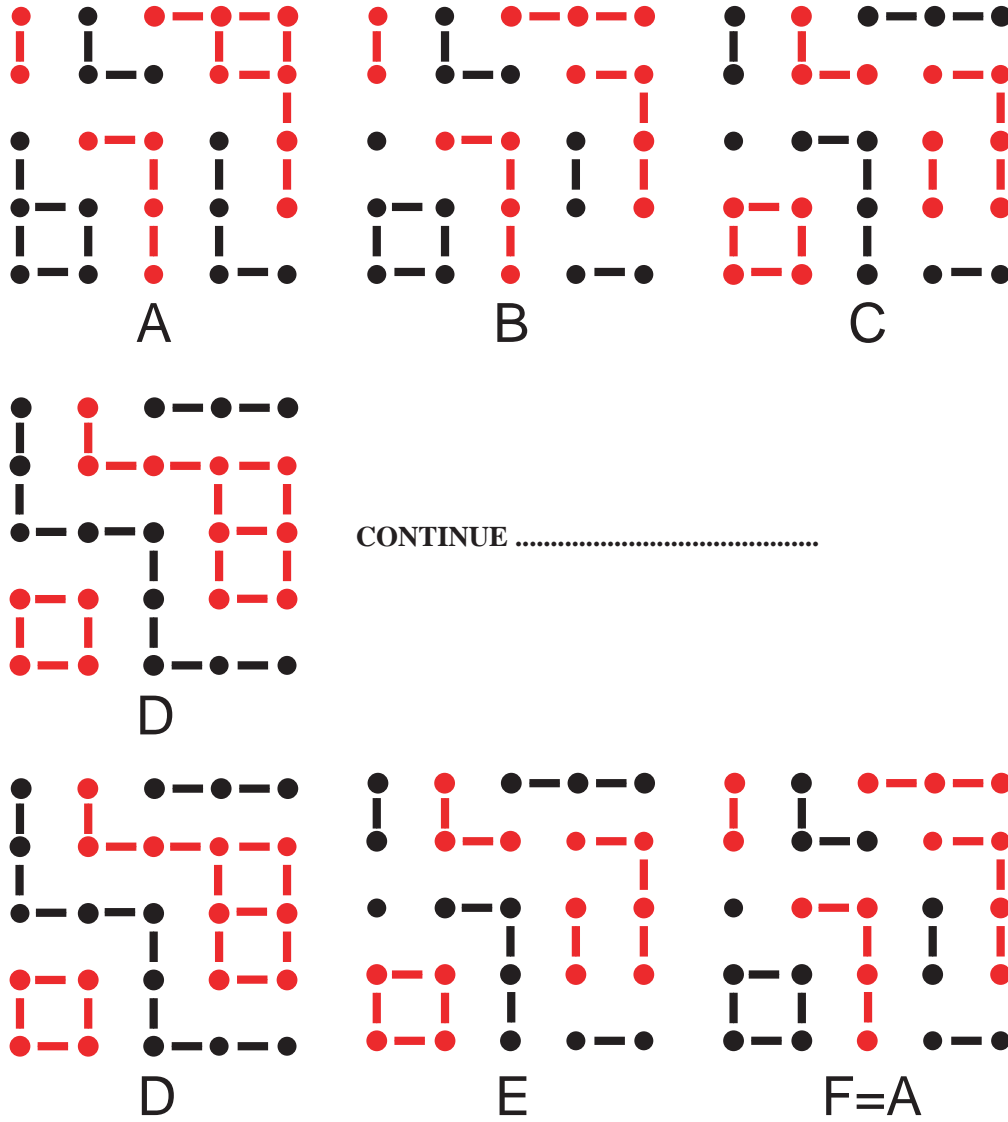


Fig. 11: Top row. A: all spins of equal sign are connected in clusters. B: bonds within clusters are destroyed with probability p . C: this new state originates from B by randomly selecting a new spin state for each of the clusters. Middle row: D originates from C by connecting neighbouring spins that have the same sign. The process then continues like in the top row. Bottom row: this illustrates microscopic reversibility. From D there is always a path to go back to A.

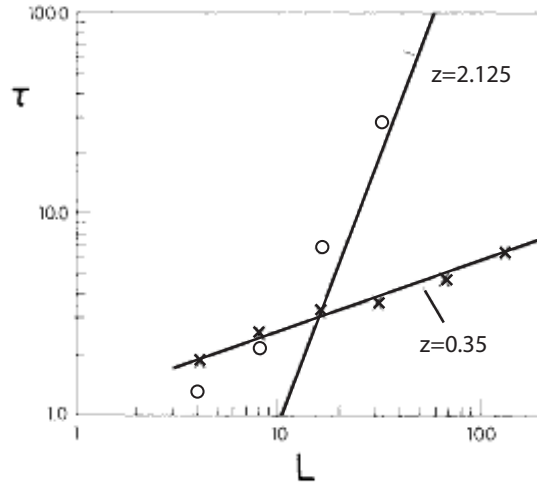


Fig. 12: Log-log plots of the correlation times for Monte Carlo simulations of the 2d Ising model at the critical temperature as a function of the system size. The circles indicate data for a standard Monte Carlo simulation and seem to follow a scaling with exponent 2.1. The crosses are results using the cluster flipping model described here and follow a powerlaw with a much smaller exponent of 0.35. Data were taken from Ref. [38].

The acceptance rule for the configuration swap is easily obtained from the condition of detailed balance,

$$\begin{aligned} P(\mathbf{i}, \beta_i)P(\mathbf{j}, \beta_j) \times A[(\mathbf{i}, \beta_i)(\mathbf{j}, \beta_j) \Rightarrow (\mathbf{j}, \beta_i)(\mathbf{i}, \beta_j)] &= \\ P(\mathbf{j}, \beta_i)P(\mathbf{i}, \beta_j) \times A[(\mathbf{j}, \beta_i)(\mathbf{i}, \beta_j) \Rightarrow (\mathbf{i}, \beta_i)(\mathbf{j}, \beta_j)] & \end{aligned} \quad (85)$$

where we have like in Eq. (46) assumed that the probability to create a move is symmetric in both directions. We then find that

$$A[(\mathbf{i}, \beta_i)(\mathbf{j}, \beta_j) \Rightarrow (\mathbf{i}, \beta_j)(\mathbf{j}, \beta_i)] = \min(1, \exp\{[\beta_j - \beta_i][\mathcal{U}(\mathbf{r}_i^N) - \mathcal{U}(\mathbf{r}_j^N)]\}) \quad (86)$$

Configuration swaps are only possible if the temperature differences are not too large — the probability density functions for the different temperatures need to overlap in (at least) their tails. Parallel tempering provides an improved sampling of regions of phase space that are difficult to reach.

Figure 13 shows an example of a system where parallel tempering is used to improve sampling. A single particle moves in a one-dimensional potential landscape that is characterized by deep valleys and high barriers (Fig. 13 A). In both cases, the particle was initially put at $x = 0$ where the potential is zero. In this case we used a system at 5 different temperatures but we show only the lowest and highest temperature. In Fig. 13 B we show the probability density functions of the particle position without parallel tempering for a particle at $k_B T/A = 4 \cdot 10^{-4}$ and $k_B T/A = 2.0$ where A is the height of the highest barrier. The low-temperature particle (in red) explores essentially only the potential landscape nearby its initial position, whereas the high temperature particle hops over any of the barriers without great difficulty and samples the whole configuration space. If we now switch on the configuration swaps between high and low temperature all 'potential valleys' become accessible to the low temperature particle and (Fig. 13 C) sampling is greatly improved.

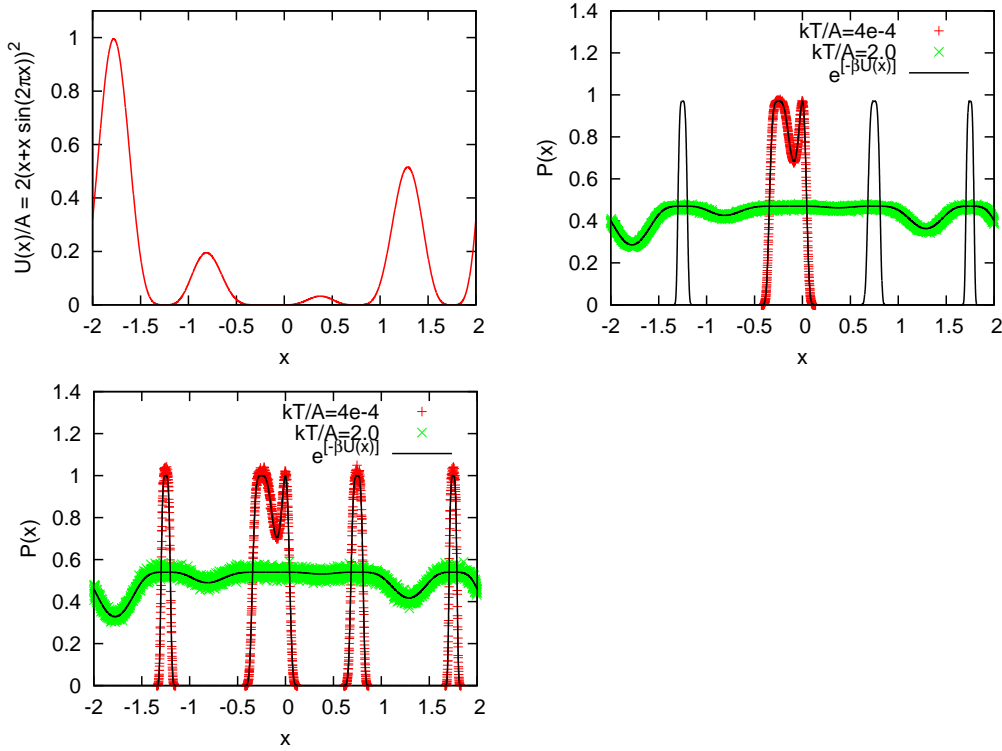


Fig. 13: Illustration of parallel tempering. *A*: one-dimensional potential in which a particle is confined. *B*: probability density function (not normalized) to find a particle at position x without parallel tempering. Red: low temperature, green: high temperature, black: theoretical curve. *C*: same as for *B*, but now with parallel tempering.

4.3 Umbrella sampling

In many cases, Metropolis sampling of the configuration space provides accurately the desired averages. In some cases however, the Metropolis method fails and one has to utilize alternative methods.

One of these cases is when the states of interest have a low probability, like for instance the top of a free energy barrier, or, when one is interested in free energy differences. In that case the umbrella sampling technique [44] is a convenient tool. In order to derive the umbrella sampling method, let us start by rewriting the average of Eq. (20) in a modified fashion by introducing a weight function $w(\mathbf{r}^N)$,

$$\langle A \rangle = \frac{\int d\mathbf{r}^N w^{-1}(\mathbf{r}^N) A(\mathbf{r}^N) \exp[-\beta \mathcal{U}(\mathbf{r}^N)] w(\mathbf{r}^N)}{\int d\mathbf{r}^N w^{-1}(\mathbf{r}^N) \exp[-\beta \mathcal{U}(\mathbf{r}^N)] w(\mathbf{r}^N)} \quad (87)$$

If we now define the biased Boltzmann factor $\exp[-\beta \mathcal{U}(\mathbf{r}^N)] w(\mathbf{r}^N)$ we can recognize the ratio of two averages in Eq. (87),

$$\langle A \rangle = \frac{\frac{\int d\mathbf{r}^N w^{-1}(\mathbf{r}^N) A(\mathbf{r}^N) \exp[-\beta \mathcal{U}(\mathbf{r}^N)] w(\mathbf{r}^N)}{\int d\mathbf{r}^N \exp[-\beta \mathcal{U}(\mathbf{r}^N)] w(\mathbf{r}^N)}}{\frac{\int d\mathbf{r}^N w^{-1}(\mathbf{r}^N) \exp[-\beta \mathcal{U}(\mathbf{r}^N)] w(\mathbf{r}^N)}{\int d\mathbf{r}^N \exp[-\beta \mathcal{U}(\mathbf{r}^N)] w(\mathbf{r}^N)}} \quad (88)$$

which is just,

$$\langle A \rangle = \frac{\langle A/w \rangle_b}{\langle 1/w \rangle_b} = \langle A/w \rangle_b \langle w \rangle. \quad (89)$$

The averages with subscript b denote averages in the biased ensemble that have a limiting distribution $\exp(-\beta \mathcal{U})w(\mathbf{r}^N)$. As we see from Eq. (89) the average $\langle A \rangle$ is thus simply the average of A/w in the biased ensemble times the average of w in the 'normal' ensemble.

In order to obtain good statistics for $\langle A \rangle$, the biased Boltzmann factor $\exp(-\beta \mathcal{U})w$ should overlap with A/w while the function w should overlap with the bare Boltzmann factor $\exp(-\beta \mathcal{U})$. The name umbrella sampling originates from this bridging property of w . The optimal choice for w would be $\exp(\beta \mathcal{F})$ since this would yield a flat distribution over which is finally integrated. However, \mathcal{F} is not known beforehand (if it would be, the whole simulation would not be necessary) and so some rough estimate for w has to be constructed in order to start the simulation.

It has been proven computationally convenient to divide the integration interval in multiple sub-intervals and to use separate weight functions instead of only one for the whole interval [5, 29]. This can be seen as follows. Suppose we want to sample an interval $\Delta\Phi = \Phi_{max} - \Phi_{min}$ which we subdivide in n intervals of length $\Delta\Phi/n$. Associated with the random walk in Φ -space there is a diffusion constant D_Φ . The time to sample one interval parallel to Φ is $\tau_{\parallel} = (\Delta\Phi/n)^2/D_\Phi$ and thus the total time to sample all intervals is $n \tau_{\parallel} = (\Delta\Phi)^2/nD_\Phi$ which seems to indicate that the total simulation time is inversely proportional with the number of intervals. However, the total simulation time also depends on the sampling time τ_{\perp} in directions orthogonal to Φ . When this time becomes larger than the time to sample parallel to Φ the efficiency of the algorithm decreases. The optimum should be when $\tau_{\parallel} \approx \tau_{\perp}$. In practice it is not really so easy to estimate these times beforehand so some trial and error will be involved in choosing the optimal parameters.

Estimating the free energy difference of a liquid and a crystal As an example we will in the following consider the free energy barrier for crystallization of soft spheres interacting via a soft-sphere pair potential,

$$u(r) = \epsilon \left(\frac{\sigma}{r} \right)^{12} \quad (90)$$

studied by Van Duijneveldt and Frenkel using umbrella sampling [45]. Similar studies were carried out more recently for the crystallization of spheres in a narrow slit [46] and for estimating the nucleation rate in Lennard-Jones fluids [47].

The problem can be formulated as follows. Suppose we have a fluid of soft repulsive particles close to, or at, the phase transition to the crystalline state. In Landau theory of phase transitions the system is characterized by an order parameter Φ which takes different values in the different phases that the system can be in. In the one-phase region of the phase diagram the probability density function $P(\Phi)$ takes only values that are sharply peaked around the average value for that phase. In the two phase region of the phase diagram we will observe a bi-modal order parameter distribution function. The probability distribution function is calculated in a simulation by collecting statistics of the occurrence of a particular value of Φ in a histogram. The Helmholtz free energy in the Landau formulation is related to the order parameter distribution function via [46, 48, 49],

$$\mathcal{F}(\Phi) = C - \beta^{-1} \ln\{P(\Phi)\}. \quad (91)$$

In this particular case the order parameter Φ has to distinguish between the fluid and crystal phase and thus an obvious choice is the bond order parameter Q_6 for which the explicit expressions can be found in for instance [45, 50]. The bond order parameter measures the degree of orientational order for a group of particles centered around a central particle. It takes for instance the values of 0 for the fluid phase, 0.35 for a simple cubic lattice, 0.51 for a body centered cubic lattice and 0.57 for a face centered cubic lattice. This makes it possible to identify these structures by just a number.

The actual simulations were carried out at constant pressure and in this ensemble the order parameter is related to the Gibbs free-energy,

$$\mathcal{G}(Q_6) = C - \beta^{-1} \ln\{P(Q_6)\}. \quad (92)$$

The order parameter interval of interest $0.05 < Q_6 < 0.5$ was divided in sub-intervals. The first simulation for each sub-interval was performed without a weight function for measuring a first estimate of the distribution function $P(Q_6)$. From this distribution function the Gibbs free energy was calculated using Eq. (92) and fitted to a polynomial $\mathcal{G}'(Q_6) = a_0 + a_1 Q_6 + a_2 Q_6^2 + \dots$. Subsequently, the simulation was repeated for the same sub-interval with a weight function $w(Q_6) = \exp[\beta \mathcal{G}'(Q_6)]$. This simulation was then used to determine $P(\Phi)$ more accurately. After all the simulations for all intervals were completed, the distribution functions were shifted by arbitrary factors to achieve continuity at the borders of each of the intervals. In Fig. (14) we show just one of the results of the calculation of the Gibbs free energy of the soft sphere system as a function of Q_6 .

Using the pre-hand knowledge that the fluid and fcc phase are in equilibrium at the given pressure and temperature we know that their Gibbs energy is the same. The barrier between the two phases can be seen to about $5 k_B T$ and the intermediate state between liquid and fcc has a strong bcc character.

The strength of the umbrella sampling method lies in the fact that one can really constrain the system to a certain range of order parameters by adding a weight function that depends on the order parameter to the Boltzmann factor. In the case described here the weight function was chosen such that it flattened the distribution function whereas in other simulations it was constructed for instance using a form $w(\Phi) = k_\Phi (\Phi - \Phi_0)^2$ to constrain the system to a Φ -window centered around Φ_0 [47].

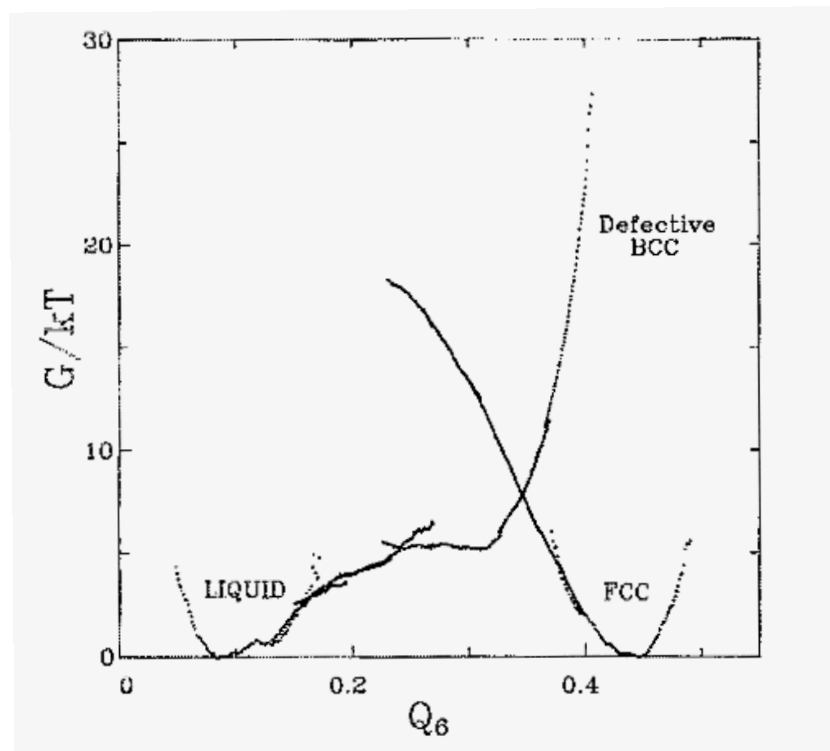


Fig. 14: Gibbs free energy near coexistence ($P^* = P\sigma^3/\epsilon = 24.21$ and $T^* = k_B T/\epsilon = 1$) for a 128-particle r^{-12} system as a function of Q_6 . The three different branches correspond to the liquid, the defect rich bcc crystal and the fcc crystal. In the drawing it was assumed that the liquid and fcc phases are in equilibrium and have the same Gibbs free energy [51]. All data were taken from Ref. [45].

References

- [1] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).
- [2] K. Binder, ed., *Monte Carlo methods in statistical physics* (Berlin, Heidelberg, New York: Springer Verlag, 1979).
- [3] M. P. Allen and D. J. Tildesley, *Computer simulations of liquids* (Clarendon Oxford, 1987).
- [4] K. Binder and D. W. Heermann, *Monte Carlo simulation in statistical physics* (Berlin: Springer, 1992).
- [5] D. Frenkel and B. Smit, *Understanding Molecular Simulation* (Academic Press San Diego, 1996).
- [6] D. P. Landau and K. Binder, *A guide to Monte Carlo simulations in statistical physics* (Cambridge University Press, 2000).
- [7] J.-P. Hansen and I. R. McDonald, *Theory of simple liquids* (Academic Press, London, 1986).
- [8] B. J. Alder and T. E. Wainwright, *J. Chem. Phys.* **27**, 1208 (1957).
- [9] P. Pusey and W. van Megen, *Nature* **320**, 990 (1984).
- [10] J. E. Lennard-Jones, *Physica* **4**, 941 (1937).
- [11] K. Kremer and G. S. Grest, *J. Chem. Phys.* **92**, 5057 (1990).
- [12] H. S. Seung and D. R. Nelson, *Phys. Rev. A* **38**(2), 1005 (1988).
- [13] H. Noguchi and G. Gompper, *Phys. Rev. Lett.* **93**, 258102 (2004).
- [14] A. Stroobants, H. N. W. Lekkerkerker, and D. Frenkel, *Phys. Rev. Lett.* **57**(12), 1452 (1986).
- [15] S. D. Zhang, P. A. Reynolds, and J. S. van Duijneveldt, *J. Chem. Phys.* **117**(21), 9947 (2002).
- [16] B. Smit, *J. Chem. Phys.* **96**, 8639 (1992).
- [17] D. C. Rapaport, *The art of molecular dynamics simulation* (Cambridge University Press, 1995).
- [18] K. K. Mon and K. Binder, *J. Chem. Phys.* **96**, 6989 (1992).
- [19] P. D. Coddington, *Int. J. Mod. Phys. C* **5**, 547 (1994).
- [20] S. Kirkpatrick and E. Stoll, *J. Comp. Phys.* **40**, 517 (1981).
- [21] URL <http://www.math.sci.hiroshima-u.ac.jp/m-mat/MT/emt.html>.

- [22] H. Flyvbjerg and H. G. Petersen, J. Chem. Phys. **91**(1), 461 (1989).
- [23] K. Huang, *Statistical Mechanics* (John Wiley & Sons, Inc., 1963).
- [24] R. K. Pathria, *Statistical Mechanics* (Pergamon Press, 1972).
- [25] W. Lenz, Phys. Zeitschrift **21**, 613 (1920).
- [26] E. Ising, Z. der Physik **31**, 253 (1925).
- [27] L. Onsager, Phys. Rev. **65**, 117 (1944).
- [28] A. B. Zamalodchikov, Int. J. Mod. Phys. A **4**, 4235 (1989).
- [29] D. Chandler, *Introduction to modern statistical mechanics* (Oxford University Press, 1987).
- [30] H. E. Stanley, *Introduction to phase transitions and critical phenomena*, The international series of monographs on physics (Oxford University Press, 1971).
- [31] J. J. Binney, N. J. Dowrick, A. J. Fisher, and M. E. J. Newman, *The theory of critical phenomena* (Clarendon press Oxford, 1992).
- [32] R. J. Glauber, J. Math. Phys. **4**, 294 (1963).
- [33] W. W. Wood, J. Chem. Phys. **48**, 415 (1968).
- [34] I. R. McDonald, Mol. Phys. **48**, 41 (1972).
- [35] Eppinga and D. Frenkel, Mol. Phys. **52**, 1303 (1984).
- [36] G. E. Norman and V. S. Filinov, High. Temp. (USSR) **7**, 216 (1969).
- [37] M. Mezei, Mol. Phys. **40**, 901 (1980).
- [38] R. H. Swendsen and J. S. Wang, Phys. Rev. Lett. **58**, 86 (1987).
- [39] P. W. Kasteleyn and C. M. Fortuin, J. Phys. Soc. Japan (Suppl) **26**, 11 (1969).
- [40] C. M. Fortuin and P. W. Kasteleyn, Physica **57**, 536 (1972).
- [41] U. Wolf, Phys. Rev. Lett. **62**, 361 (1989).
- [42] A. P. Lyubartsev, A. A. Martsinovski, S. V. Shevkunov, and P. N. Vorontsov-Vel'yaminov, J. Chem. Phys. **96**, 1776 (1992).
- [43] E. Marinari and G. Parisi, Europhys. Lett. **19**, 451 (1992).
- [44] G. M. Torrie and J. P. Valleau, J. Comp. Phys. **23**, 187 (1977).
- [45] J. S. van Duijneveldt and D. Frenkel, J. Chem. Phys. **96**, 4655 (1992).
- [46] R. Radhakrishnan and K. E. Gubbins, Mol. Phys. **96**, 1249 (1999).
- [47] P. R. ten Wolde, M. J. Ruiz-Montero, and D. Frenkel, J. Chem. Phys. **104**, 9932 (1996).

- [48] L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Pergamon Press London, 1980), 3rd ed.
- [49] P. M. Chaikin and T. C. Lubensky, *Principles of condensed matter physics* (Cambridge University Press, 1995).
- [50] P. J. Steinhardt, D. R. Nelson, and M. Ronchetti, Phys. Rev. B **28**, 784 (1983).
- [51] H. Ogura, M. Matsuda, T. Ogawa, N. Ogita, and A. Ueda, Prog. Theor. Phys. **58**, 419 (1977).