Contents of Lecture II

Statistical Physics and Potts Model

MC simulations of systems described by the Gibbs canonical ensemble aim at calculating estimators of physical observables at a temperature T. In the following we choose units so that the becomes one, and $\beta = 1/T$. Let us consider the calculation of the **expectation value** of an **observable** $\mathcal{O}$. Mathematically all systems on a computer are discrete, because a finite word length has to be used. Hence, the expectation value is given by the sum

$$\widehat{\mathcal{O}} = \widehat{\mathcal{O}}(\beta) = \langle \mathcal{O} \rangle \;\; = \;\; Z^{-1} \sum_{k=1}^{K} \mathcal{O}^{(k)} \, e^{-\beta \, E^{(k)}} \tag{1}$$

$$\text{where} \;\; Z \;\; = \;\; Z(\beta) \;\; = \;\; \sum_{k=1}^{K} e^{-\beta \, E^{(k)}} \tag{2}$$

is the **partition function** and the index $k = 1, \ldots, K$ labels all **configurations** of the system, and $E^{(k)}$ is the (internal) energy of configuration k. The configurations are also called **microstates**. To distinguish the configuration index from other indices, it is put in parenthesis.

We introduce generalized Potts models in an external magnetic field on d-dimensional hypercubic lattices with periodic boundary conditions. Without being overly complicated, these models are general enough to illustrate the essential features we are interested in. In addition, various subcases of these models are by themselves of physical interest. Generalizations of the algorithmic concepts to other models are straightforward, although technical complications may arise.

We define the energy function of the system by

$$-\beta\, E^{(k)} = -\beta\, E_0^{(k)} + H\, M^{(k)} \tag{3}$$

where

$$E_0^{(k)} = -2 \sum_{<ij>} J_{ij}(q_i^{(k)}, q_j^{(k)})\; \delta(q_i^{(k)}, q_j^{(k)}) + \frac{2\, d\, N}{q} \tag{4}$$

$$\text{with} \quad \delta(q_i, q_j) = \begin{cases} 1 \text{ for } q_i = q_j \\ 0 \text{ for } q_i \neq q_j \end{cases} \quad \text{and} \quad M^{(k)} = 2 \sum_{i=1}^{N} \delta(1, q_i^{(k)}).$$

The sum $< ij >$ is over the nearest neighbor lattice sites and $q_i^{(k)}$ is called the **Potts spin** or **Potts state** of configuration k at site i. For the q-state Potts model $q_i^{(k)}$ takes on the values $1, \ldots, q$. The external magnetic field is chosen to interact with the state $q_i = 1$ at each site i, but not with the other states $q_i \neq 1$. The $J_{ij}(q_i, q_j)$, $(q_i = 1, \ldots, q; q_j = 1, \ldots, q)$ functions define the exchange coupling constants between the states at site i and site j. The energy function describes a number of physically interesting situations. With

$$J_{ij}(q_i, q_j) \equiv J > 0 \quad \text{(conventionally } J = 1) \tag{5}$$

the original model is recovered and $q = 2$ becomes equivalent to the Ising ferromagnet, see F.Y. Wu for a detailed review. The Ising case of Edwards-Anderson spin glasses and quadrupolar Potts glasses are obtained when the exchange constants are quenched random variables. Other choices of the J_{ij} include anti-ferromagnets and the fully frustrated Ising model.

For the **energy per spin** the notations ($H = 0$ and $H \neq 0$) are

$$e_{0s} = E_0/N \quad \text{and} \quad e_s = E/N \ . \tag{6}$$

A factor of two is introduced, so that e_{0s} agrees for $q = 2$ with the conventional Ising model definition, so that

$$\beta \;=\; \beta^{\text{Ising}} \;=\; \frac{1}{2}\beta^{\text{Potts}}. \tag{7}$$

For the $2d$ Potts models a number of exact results are known in the infinite volume limit, mainly due to work by Baxter. The phase transition temperatures are

$$\frac{1}{2}\beta_c^{\text{Potts}} \;=\; \beta_c \;=\; \frac{1}{T_c} \;=\; \frac{1}{2}\ln(1+\sqrt{q}), \quad q = 2, 3, \ldots \; . \tag{8}$$

At β_c the average energy per state is

$$e_{0s}^c = E_0^c/N = \frac{4}{q} - 2 - 2/\sqrt{q} \; . \tag{9}$$

The phase transition is second order for $q \leq 4$ and first order for $q \geq 5$ for which the exact infinite volume **latent heats** $\triangle e_{0s}$ and **entropy jumps** $\triangle s$ were also found by Baxter, while the interface tensions f_s were derived later.

Sampling and Re-weighting

For the Ising model it is straightforward to **sample statistically independent configurations**. We simply have to generate N spins, each either up or down with 50% likelihood. This is called **random sampling**. The figure depicts a thus obtained histogram for the $2d$ Ising model **energy per spin**
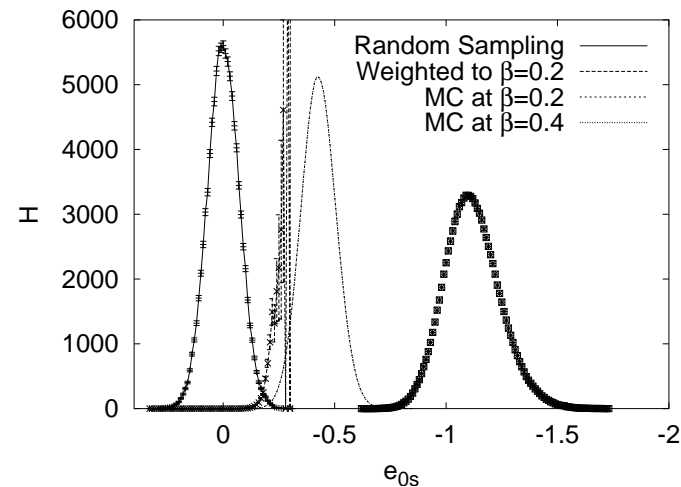

Figure 1: Energy histograms of $100\,000$ entries each for the Ising model on an 20×20 lattice: Random Sampling gives statistically independent configurations at $\beta = 0$. Histograms at $\beta = 0.2$ and $\beta = 0.4$ are generated by Markov chain MC. Re-weighting of the $\beta = 0$ random configurations to $\beta = 0.2$ is shown to fail (assignments a0301_02 and a0303_02).

Note that is is very important to distinguish the energy measurements on single configurations from the expectation value. The expectation value $\widehat{e}_{0s}$ is a single number, while e_{0s} fluctuates. From the measurement of many e_{0s} values one finds estimators of its moments. The mean is is denoted by $\overline{e}_{0s}$ and fluctuates.

The histogram entries at $\beta = 0$ can be re-weighted so that they correspond to other β values. We simply have to multiply the entry corresponding to energy E by $c_\beta \exp(-\beta E)$. Similarly histograms corresponding to the Gibbs ensemble at some value β_0 can be re-weighted to other β values.

Care has to be taken to ensure that the involved arguments of the exponential function do not become too large. This is done in by first calculating the mean energy and then implementing re-weighting with respect to the difference from the mean.

Re-weighting has a long history. For FSS investigations of second order phase transitions its usefulness of the re-weighting has been stressed by Ferrenberg and Swendsen (accurate determinations of peaks of the specific heat or of susceptibilities).

In the figure re-weighting is done from $\beta_0 = 0$ to $\beta = 0.2$. But, by comparison to the histogram from a Metropolis MC calculation at $\beta = 0.2$, the result is seen to be disastrous. The reason is easily identified: In the range where the $\beta = 0.2$ histogram takes on its maximum, the $\beta = 0$ histogram has not a single entry, *i.e.*, our naive sampling procedure misses the **important configurations** at $\beta = 0.2$. Re-weighting to new β values works only in a range $\beta_0 \pm \triangle\beta$, where $\triangle\beta \to 0$ in the infinite volume limit.

Important Configurations

Let us determine the important contributions to the partition function. The partition function can be re-written as a sum over energies

$$Z = Z(\beta) = \sum_E n(E) \, e^{-\beta \, E} \tag{10}$$

where the unnormalized spectral density $n(E)$ is defined as the number of microstates k with energy E (note, on the computer energy values are always discrete).

For a fixed value of β the energy probability density

$$P_\beta(E) = c_\beta\, n(E)\, e^{-\beta E} \tag{11}$$

is peaked around the average value $\widehat{E}(\beta)$, where c_β is a normalization constant so that the $\sum_E P_\beta(E) = 1$ holds.

Away from first and second order phase transitions, the width of the energy distribution is $\triangle E \sim \sqrt{V}$ follows from the fact that, away from phase transition points, the fluctuations of the $N \sim V$ lattice spins are essentially uncorrelated, so that the magnitude of a typical fluctuations is $\sim \sqrt{N}$. From this we find that the re-weighting range is $\triangle \beta \sim 1/\sqrt{V}$, as the energy is an extensive quantity $\sim V$ so that $\triangle \beta E \sim \sqrt{V}$ can stay within the fluctuation of the system.

Interestingly, the re-weighting range increases at second order phase transitions point, because critical fluctuations are larger than non-critical fluctuations. Namely, one has $\triangle E \sim V^x$ with $1/2 < x < 1$ and the requirement $\triangle \beta E \sim V^x$ yields $\triangle \beta \sim V^{x-1}$.

For first order phase transitions one has a latent heat $\triangle V \sim V$, but this does not mean that the re-weighting range becomes of order one. In essence, the fluctuations collapse, because the two phases become separated by an interfacial tension. One is back to fluctuations within either of the two phases, $i.e.$ $\triangle \beta \sim 1/\sqrt{V}$.

The important configurations at temperature $T = 1/\beta$ are at the energy values for which the probability density $P_\beta(E)$ is large. To sample them efficiently, one needs a procedure which generates the configurations with their Boltzmann weights

$$w_B^{(k)} = e^{-\beta E^{(k)}} \ . \tag{12}$$

The number of configurations $n(E)$ and the weights combine then so that the probability to generate a configuration at energy E becomes precisely $P_\beta(E)$ as given by equation (11).

Importance Sampling and Markov Chain Monte Carlo

For the canonical ensemble **Importance sampling** generates configurations k with probability

$$P_B^{(k)} = c_B\, w_B^{(k)} = c_B\, e^{-\beta E^{(k)}} \tag{13}$$

where the constant c_B is determined by the normalization condition $\sum_k P_B^{(k)} = 1$. The **state vector** $(P_B^{(k)})$, for which the configurations are the vector indices, is called **Boltzmann state**. When configurations are **stochastically** generated with probability $P_B^{(k)}$, the **expectation value** becomes the **arithmetic average:**

$$\widehat{\mathcal{O}} = \widehat{\mathcal{O}}(\beta) = \langle \mathcal{O} \rangle = \lim_{N_K \to \infty} \frac{1}{N_K} \sum_{n=1}^{N_K} \mathcal{O}^{(k_n)} \; . \tag{14}$$

When the sum is truncated at some finite value of N_K, we obtain an **estimator of the expectation value**

$$\overline{\mathcal{O}} = \frac{1}{N_K} \sum_{k=1}^{N_K} \mathcal{O}^{(k_n)} \; . \tag{15}$$

Normally, we cannot generate configurations k directly with probability (13). But they may be found as members of the equilibrium distribution of a dynamic process. In practice Markov chains are used. A **Markov process** is a particularly simple dynamic process, which generates configuration k_{n+1} stochastically from configuration k_n, so that no information about previous configurations $k_{n-1}, k_{n-2}, \ldots$ is needed. The elements of the Markov chain **time series** are the configurations. Assume that the configuration k is given. Let the transition probability to create the configuration l in one step from k be given by $W^{(l)(k)} = W[k \to l]$. In essence, the matrix

$$W = \left(W^{(l)(k)} \right) \tag{16}$$

defines the Markov process. Note, that this **transition matrix** is a very big (never stored in the computer!), because its labels are the configurations. To achieve our goal to generate configurations with the desired probabilities, the matrix W is required to satisfy the following properties:

(i) **Ergodicity**:

$$e^{-\beta E^{(k)}} > 0 \quad \text{and} \quad e^{-\beta E^{(l)}} > 0 \quad \text{imply}: \tag{17}$$

an integer number $n > 0$ exists so that $(W^n)^{(l)(k)} > 0$ holds.

(ii) **Normalization**:

$$\sum_l W^{(l)(k)} = 1 . \tag{18}$$

(iii) **Balance**:

$$\sum_k W^{(l)(k)} e^{-\beta E^{(k)}} = e^{-\beta E^{(l)}} . \tag{19}$$

Balance means: The Boltzmann state (13) is an eigenvector with eigenvalue 1 of the matrix $W = (W^{(l)(k)})$.

In statistical physics the ensemble notation is frequently used. By definition, an **ensemble** is a collection of configurations so that to each configuration k a probability $P^{(k)}$ is assigned, $\sum_k P^{(k)} = 1$. The **Gibbs or Boltzmann ensemble** E_B is defined to be the ensemble with probability distribution (13).

An **equilibrium ensemble** E_{eq} of the Markov process is defined by its probability distribution P_{eq} satisfying

$$W P_{eq} = P_{eq}, \quad \text{in components} \quad P_{eq}^{(l)} = \sum_k W^{(l)(k)} P_{eq}^{(k)}. \quad (20)$$

Statement: Under the conditions (i), (ii) and (iii) the Boltzmann ensemble is the **only** equilibrium ensemble of the Markov process.

Proof: Let us first define a distance between ensembles. Suppose we have two ensembles E and E', each of which is a collection of many configurations. Denote the probability for configuration k in E by $P^{(k)}$ and in E' by $P'^{(k)}$. We define the **distance** between E and E' to be

$$||E - E'|| = \sum_k |P^{(k)} - P'^{(k)}|, \quad (21)$$

where the sum goes over all configurations. Suppose that E' resulted from the application of the transition matrix W to the ensemble E. We can compare the

distance of E' from the Boltzmann ensemble with the distance of E from the Boltzmann ensemble:

$$
\begin{aligned}
||E' - E_B|| &= \sum_l \left| \sum_k W^{(l)(k)} (P^{(k)} - P_B^{(k)}) \right| && \text{(using balance)} \\
&\leq \sum_l \sum_k \left| W^{(l)(k)} (P^{(k)} - P_B^{(k)}) \right| && \text{(using the triangle inequality)} \\
&= \sum_k \left| P^{(k)} - P_B^{(k)} \right| = ||E - E_B|| . && (22)
\end{aligned}
$$

The last line is obtained by making use of the normalization $\sum_l W^{(l)(k)} = 1$ and of $W^{(l)(k)} \geq 0$.

The algorithm **reduces the distance** of an ensemble from the equilibrium Boltzmann ensemble whenever ergodicity holds. First, we note that the inequality (22) is strict if $W^{(l)(k)}$ never vanishes. This is easily seen to follow from the normalization condition $\sum_k P^{(k)} = \sum P_B^{(k)} = 1$: Assume, that some

configuration (k_1) exists for which

$$P^{(k_1)} - P_B^{(k_1)} > 0 \tag{23}$$

holds, it follows that for at least one other configuration (k_2) exists for which

$$P^{(k_2)} - P_B^{(k_2)} < 0 \ . \tag{24}$$

Consequently, the contributions in $\sum_k W^{(l)(k)}(P^{(k)} - P_B^{(k)})$ flip signs and the equality of the $\leq$ sign in equation (22) can be excluded. Here we have used the (unrealistic) assumption that all matrix elements $W^{(l)(k)}$ are greater than zero. This restriction is easily removed. Ergodicity implies $(W^n)^{(k_1)(k_2)} > 0$ and $(W^n)^{(k_2)(k_1)} > 0$ for some integer $n > 0$. As $(W^n)^{(l)(k)}$ can be used instead of $W^{(l)(k)}$, it follows that the distance to the Boltzmann ensemble will be reduced by applying the Markov process sufficiently often. This concludes the proof of the statement. Obviously, the approach to the Boltzmann distribution will become slow if n is large.

There are many ways to construct a Markov process satisfying (i), (ii) and (iii).

A stronger condition than balance (19) is

(iii') **Detailed balance**:

$$W^{(l)(k)} e^{-\beta E^{(k)}} = W^{(k)(l)} e^{-\beta E^{(l)}} . \tag{25}$$

Using the normalization $\sum_k W^{(k)(l)} = 1$ detailed balance implies balance (iii).

At this point we have succeeded to replace the canonical ensemble average by a time average over an artificial dynamics. Calculating then averages over large times, like one does in real experiments, is equivalent to calculating averages of the ensemble. One distinguishes *dynamical universality classes*. The Metropolis and heat bath algorithms discussed in the following fall into the class of so called *Glauber dynamics*, which imitates the thermal fluctuations of nature to some extent. In a frequently used classification this is called *model A*. Cluster algorithms discussed constitute another universality class. Some recent attention has focused on dynamical universality classes of non-equilibrium systems.

The Metropolis Algorithm

Detailed balance still does not uniquely fix the transition probabilities $W^{(l)(k)}$. The **Metropolis algorithm** is a popular choice can be used whenever one knows how to calculate the energy of a configuration. Given a configuration k, the Metropolis algorithm proposes a configuration l with probability

$$f(l, k) \quad \text{normalized to} \quad \sum_l f(l, k) = 1 \ . \tag{26}$$

We derive a symmetry condition which ensures detailed balance.

The new configuration l is accepted with probability

$$w^{(l)(k)} = \min\left[1, \frac{P_B^{(l)}}{P_B^{(k)}}\right] = \begin{cases} 1 & \text{for} \quad E^{(l)} < E^{(k)} \\ e^{-\beta(E^{(l)} - E^{(k)})} & \text{for} \quad E^{(l)} > E^{(k)}. \end{cases} \tag{27}$$

If the new configuration is rejected, the old configuration has to be counted again. A beginners mistake is to count configurations only after accepted changes.

The **acceptance rate** is defined as the ratio of accepted changes over proposed moves. With this convention we do not count a move as accepted when it proposes the at hand configuration.

The Metropolis procedure gives rise to the transition probabilities

$$W^{(l)(k)} = f(l,k)\,w^{(l)(k)} \ \text{ for } \ l \neq k \tag{28}$$

$$\text{and} \ W^{(k)(k)} = f(k,k) + \sum_{l \neq k} f(l,k)\,(1 - w^{(l)(k)}) \,. \tag{29}$$

Therefore, the ratio $\left(W^{(l)(k)}/W^{(k)(l)}\right)$ satisfies detailed balance (25) if

$$f(l,k) = f(k,l) \ \text{ holds} \,. \tag{30}$$

Otherwise the probability density $f(l,k)$ is unconstrained. So there is an amazing flexibility in the choice of the transition probabilities $W^{(l)(k)}$. One can even use acceptance probabilities distinct from those of equation (27) and the proposal probabilities are then not necessarily symmetric anymore (e.g. Hastings). Also, the algorithm generalizes immediately to arbitrary weights.

The $O(3)$ σ Model and the Heat Bath Algorithm

We give an example of a model with a continuous energy function. The $2d$ version of the model is of interest to field theorists because of its analogies with the four-dimensional Yang-Mills theory. In statistical physics the d-dimensional model is known as the Heisenberg ferromagnet. Expectation values are calculated with respect to the partition function

$$Z \;=\; \int \prod_i ds_i \; e^{-\beta E(\{s_i\})} \;. \tag{31}$$

The spins $\;\vec{s}_i = \begin{pmatrix} s_{i,1} \\ s_{i,2} \\ s_{i,3} \end{pmatrix}\;$ are normalized to $\;(\vec{s}_i)^2 = 1$ \tag{32}

and the measure ds_i is defined by $\;\displaystyle\int ds_i \;=\; \frac{1}{4\pi} \int_{-1}^{+1} d\cos(\theta_i) \int_0^{2\pi} d\phi_i \;,$ (33)

where the polar (θ_i) and azimuth (ϕ_i) angles define the spin s_i on the unit sphere.

The energy is

$$E = \sum_{<ij>} \left(1 - \vec{s}_i \vec{s}_j\right) , \tag{34}$$

where the sum goes over the nearest neighbor sites of the lattice. We would like to update a single spin $\vec{s}$. The sum of its $2d$ neighbors is

$$\vec{S} = \vec{s}_1 + \vec{s}_2 + \ldots + \vec{s}_{2d-1} + \vec{s}_{2d} .$$

Hence, the contribution of spin $\vec{s}$ to the action is $2d - \vec{s}\vec{S}$. We propose a new spin $\vec{s}\,'$ with the measure (33) by drawing two uniformly distributed random numbers

$$\phi^r \in [0, \pi) \quad \text{for the azimuth angle and}$$

$$\cos(\theta^r) = x^r \in [-1, +1) \quad \text{for the cosine of the polar angle.}$$

This defines the probability function $f(\vec{s}\,', \vec{s})$ of the Metropolis process, which accepts the proposed spin $\vec{s}\,'$ with probability

$$w(\vec{s} \to \vec{s}\,') = \begin{cases} 1 & \text{for} \quad \vec{S}\vec{s}\,' > \vec{S}\vec{s}, \\ e^{-\beta(\vec{S}\vec{s} - \vec{S}\vec{s}\,')} & \text{for} \quad \vec{S}\vec{s}\,' < \vec{S}\vec{s}. \end{cases}$$

If sites are chosen with the uniform probability distribution $1/N$ per site, where N is the total number of spins, it is obvious that the procedure fulfills detailed balance. It is noteworthy that the procedure remains valid when the spins are chosen in the systematic order $1, \ldots, N$, then $1, \ldots, N$ again, and so on. Balance (19) still holds, whereas detailed balance (25) is violated (exercise).

The heath bath algorithm

Repeating the Metropolis algorithm again and again for the same spin $\vec{s}$ leads to the equilibrium distribution of this spin, which reads

$$P(\vec{s}\,'; \vec{S}) = \text{const}\, e^{\beta \vec{S} \vec{s}\,'} \quad \text{with} \quad \int P(\vec{s}\,'; \vec{S})\, ds' = 1 \ .$$

One would prefer to choose $\vec{s}\,'$ directly with the probability

$$W(\vec{s} \to \vec{s}\,') \; = \; P(\vec{s}\,'; \vec{S}) \; = \; \text{const}\, e^{\beta \vec{s}\,' \vec{S}} \ ,$$

as $\vec{s}\,'$ is then immediately Boltzmann distributed with respect to its neighbor spins. The algorithm, which creates this distribution, is called the **heat bath algorithm.**

Implementation of this algorithm becomes feasible when the energy function is sufficiently simple to allow for an explicit calculation of the probability $P(\vec{s}\,'; \vec{S})$. This is an easy task for the $O(3)$ σ-model. Let

$$\alpha = \text{angle}(\vec{s}\,', \vec{S}), \quad x = \cos(\alpha) \quad \text{and} \quad S = \beta|\vec{S}| \ .$$

For $S = 0$ a new spin $\vec{s}\,'$ is simply obtained by random sampling. We assume in the following $S > 0$. The Boltzmann weight becomes $\exp(xS)$ and the normalization constant follows from

$$\int_{-1}^{+1} dx\, e^{xS} \ = \ \frac{2}{S}\, \sinh(S) \ .$$

Therefore, the desired probability is

$$P(\vec{s}\,'; \vec{S}) \ = \ \frac{S}{2\sinh(S)}\, e^{xS} \ =: f(x)$$

and the method of the first lecture can be used to generate events with the probability density $f(x)$.

With

$$y = F(x) = \int_{-1}^{x} dx' \, f(x') = \int_{-1}^{x} dx' \, \frac{S}{2 \sinh(S)} \, e^{x'S} = \frac{\exp(+xS) - \exp(-S)}{2 \sinh(S)}$$

a uniformly distributed random number $y^r \in [0,1)$ translates into

$$x^r = \cos \alpha^r = \frac{1}{S} \ln \left[\exp(+S) - y^r \exp(+S) + y^r \exp(-S) \right] . \qquad (35)$$

Finally, one has to give $\vec{s}\,'$ a direction in the plane orthogonal to S. This is done by choosing a random angle β^r uniformly distributed in the range $0 \leq \beta^r < 2\pi$. Then, $x^r = \cos \alpha^r$ and β^r completely determine $\vec{s}\,'$ with respect to $\vec{S}$. Before storing $\vec{s}\,'$ in the computer memory, we have to calculate coordinates of $\vec{s}\,'$ with respect to a Cartesian coordinate system, which is globally used for all spins of the lattice. This is achieved by a linear transformation. We define

$$\cos \theta = \frac{S_3}{S}, \ \sin \theta = \sqrt{1 - \cos^2 \theta}, \ \cos \phi = \frac{S_1}{S \sin \theta} \ \text{ and } \ \sin \phi = \frac{S_2}{S \sin \theta} \ .$$

Unit vectors of a coordinate frame K', with $\hat{z}$ in the direction of $\hat{S}$ and $\hat{y}$ in the $x - y$ plane, are then defined by

$$\hat{z}' = \begin{pmatrix} \sin\theta\,\cos\phi \\ \sin\theta\,\sin\phi \\ \cos\theta \end{pmatrix}, \quad \hat{x}' = \begin{pmatrix} \cos\theta\,\cos\phi \\ \cos\theta\,\sin\phi \\ -\sin\theta \end{pmatrix} \quad \text{and} \quad \hat{y}' = \begin{pmatrix} -\sin\phi \\ \cos\phi \\ 0 \end{pmatrix}.$$

Expanding $\vec{s}\,'$ in these units vectors, we have

$$\vec{s}\,' = \sin\alpha^r\,\cos\beta^r\,\hat{x}\,' + \sin\alpha^r\,\sin\beta^r\,\hat{y}\, + \cos\alpha^r\,\hat{z}\,'$$

$$= \begin{pmatrix} \sin\alpha^r\,\cos\beta^r\,\cos\theta\,\cos\phi - \sin\alpha^r\,\sin\beta^r\,\sin\phi + \cos\alpha^r\,\sin\theta\,\cos\phi \\ \sin\alpha^r\,\cos\beta^r\,\cos\theta\,\sin\phi + \sin\alpha^r\,\cos\beta^r\,\cos\phi + \cos\alpha^r\,\sin\theta\,\sin\phi \\ -\sin\alpha^r\,\cos\beta^r\,\sin\theta + \cos\alpha^r\,cos\theta \end{pmatrix} \quad (36)$$

and the three numbers of the column vector (36) are stored in the computer.

Potts Model Heat bath code

The heat bath algorithm chooses a state q_i directly with the local Boltzmann distribution defined by its nearest neighbors. The state q_i can take on one of the values $1, \ldots, q$ and, with all other states set, determines a value of the energy function (3). We denote this energy by $E(q_i)$ and the Boltzmann probabilities are

$$P_B(q_i) \;=\; \text{const } e^{-\beta\, E(q_i)} \tag{37}$$

where the constant is determined by the normalization condition

$$\sum_{q_i=1}^{q} P_B(q_i) \;=\; 1 \;. \tag{38}$$

In equation (37) we can define $E(q_i)$ to be just the contribution of the interaction of q_i with its nearest neighbors to the total energy and absorb the other contribution into the overall constant. The $E(q_i)$ values depend only on how the nearest neighbors of the spin q_i partition into the values $1, \ldots, q$. For low values of q and

the dimension d the most efficient implementation of the heatbath algorithm is to tabulate all possibilities. However, here we prefer to give a generic code which works for arbitrary values of q and d.

For this we calculate the cumulative distribution function of the heat bath probabilities

$$P_{HB}(q_i) = \sum_{q'_i=1}^{q_i} P_B(q'_i) \ . \tag{39}$$

The normalization condition (38) implies $P_{HB}(q) = 1$. Comparison of these cumulative probabilities with a uniform random number x^r yields the heat bath update $q_i \to q'_i$. Note that in the heat bath procedure the original value q_i^{in} does not influence the selection of q_i^{new}.

Start and equilibration

Under repeated application of one of our updating procedures the probability of states will approach the Boltzmann distribution. However, initially we have to start with a microstate which may be far off the Boltzmann distribution. Far off means, that the Boltzmann probability (at temperature T) for a typical state of the initially generated distribution can be very, very small. Suppression factors like 10^{-10000} are well possible. Although the weight of states decreases with $1/n$ where n is the steps of the Markov process, one should exclude the initial states from the equilibrium statistics. In practice this means we should allow for a certain number of sweeps `nequi` to equilibrate the system.

Many ways to generate start configurations exist. Two natural and easy to implement choices are:

1. Generate a random configuration corresponding to $\beta = 0$. This defines a **random** or **disordered start** of a MC simulation.

2. Generate a configuration for which all Potts spins take on the same q-value. This is called an **ordered start** of a MC simulation.
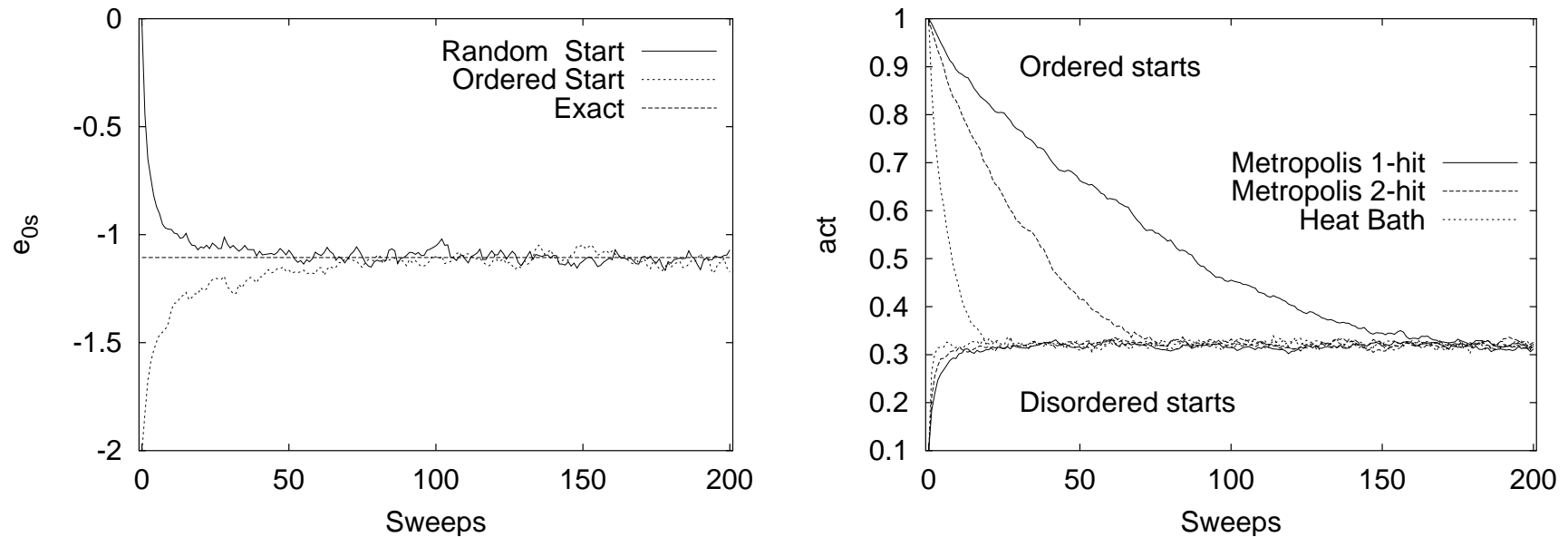
Examples of initial time series:

Figure 2: Left: Two Metropolis time series of 200 sweeps each for a $2d$ Ising model on a 80×80 lattice at $\beta = 0.4$ are shown. Ordered and disordered starts are used. The exact mean value $\widehat{e}_{0s} = -1.10608$ is also indicated (assignment a0302_01). Right: $q = 10$ Potts model time series of 200 sweeps on an 80×80 lattice at $\beta = 0.62$. Measurements of the action variable after every sweep are plotted for ordered and disordered starts (assignment a03003_05).

Consistency Checks

For the $2d$ Ising model we can test against the exact finite lattice results of Ferdinand and Fisher. We simulate an 20^2 lattice at $\beta = 0.4$.

Now we use a statistics of $10\,000$ sweeps for reaching equilibrium and assume that this is an overkill. A more careful analysis is the subject of the next lecture.

The statistics for measurement is chosen 32 times larger: 64 bins of $5\,000$ sweeps each. The number 64 is taken, because according to the student distribution the approximation to the Gaussian approximation is then already excellent, while the binsize of $5\,000$ $(\gg 200)$ is argued to be large enough to neglect correlations between the bins. With this statistics we find (assignment a0303_06)

$$\overline{e}_{0s} = -1.1172\,(14)\ \ (\text{Metropolis})\ \ \text{versus}\ \ \widehat{e}_s = -1.117834\ \ (\text{exact})\ . \tag{40}$$

Performing the Gaussian difference test gives a perfectly admissible value

$$Q = 0.66\ .$$

For the $2d$ 10-state Potts model at $\beta = 0.62$ we test our Metropolis versus our heat bath code on a small 20×20 lattice. For the heat bath updating we use the same statistics as previously for the $2d$ Ising model.

For the Metropolis updating we increase these numbers by a factor of four to $40\,000$ sweeps for reaching equilibrium and $64 \times 20\,000$ sweeps for measurements. This increase is done, because we expect the performance of Metropolis updating for the 10-state model to be far worse than for the 2-state model: At low temperature the likelihood to propose the most probable (aligned) Potts spin is $1/2$ for the 2-state model, but only $1/10$ for the 10-state model and $\beta = 0.62$ is sufficiently close to the ordered phase, so that this effect is expected to be of relevance. The results of our simulations are (assignment a0303_10)

$$\texttt{actm} = 0.321772\,(75)\ \text{(Metropolis)}\quad \text{versus}\quad \texttt{actm} = 0.321661\,(70)\ \text{(heat bath)}$$

and $Q = 0.28$ for the Gaussian difference test. Another perfectly admissible value.

$3d$ **3-state Potts model**

To illustrate features of a first order phase transition for the $3d$ 3-state Potts model on an 24^3 lattice. We use the 1-hit Metropolis algorithm and simulate at $\beta = 0.275229525$, perform $20\,000$ sweeps for reaching equilibrium, then $64 \times 10\,000$ sweeps with measurements. From the latter statistics we show the action variable histogram and its error bars (assignment a0303_10):
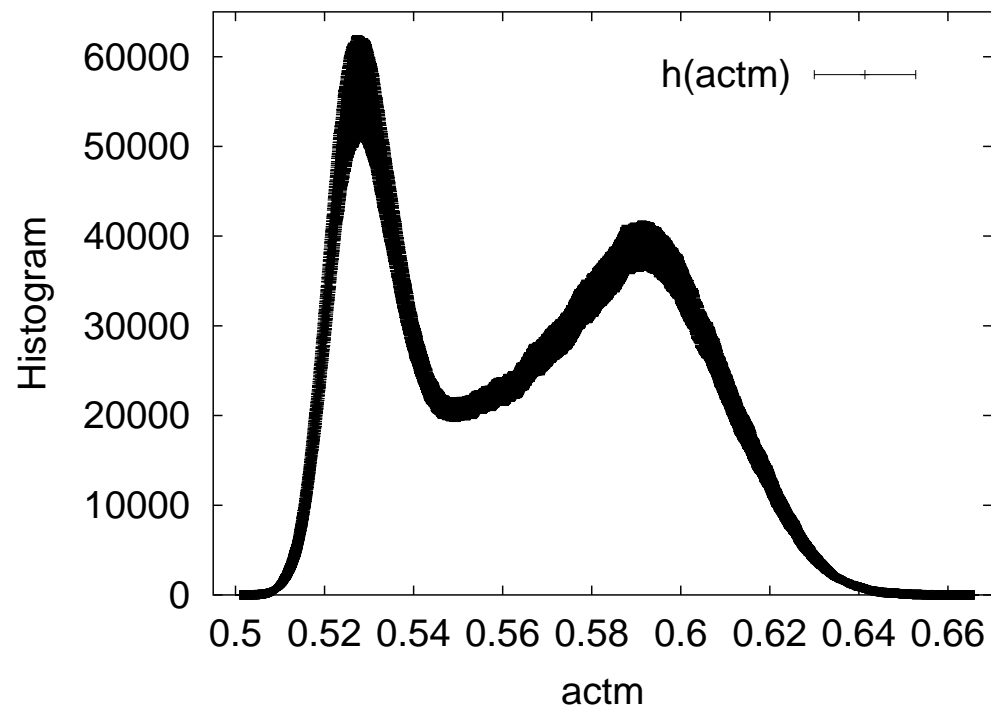
The histogram exhibits a **double peak** structure which is typically obtained when systems with first order transitions are simulated on finite lattices in the neighborhood of so called **pseudo-transition temperatures**. These are finite lattice temperature definitions, which converge with increasing system size towards the infinite volume transition temperature. Equal heights of the maxima of the two peaks is a popular definition of a pseudo-transition temperature for first order phase transitions. Our β value needs to be re-weighted to a slightly higher value to arrange for equal heights (assignment a0303_10).

Self-Averaging Illustration for the $O(3)$ model

We compare the peaked distribution function of the mean action per link for different lattice sizes. The property of **self-averaging** is observed: The larger the lattice, the smaller the confidence range. The other way round, the peaked distribution function is very well suited to exhibit observables for which self-averaging does not work, as for instance encountered in spin glass simulations.
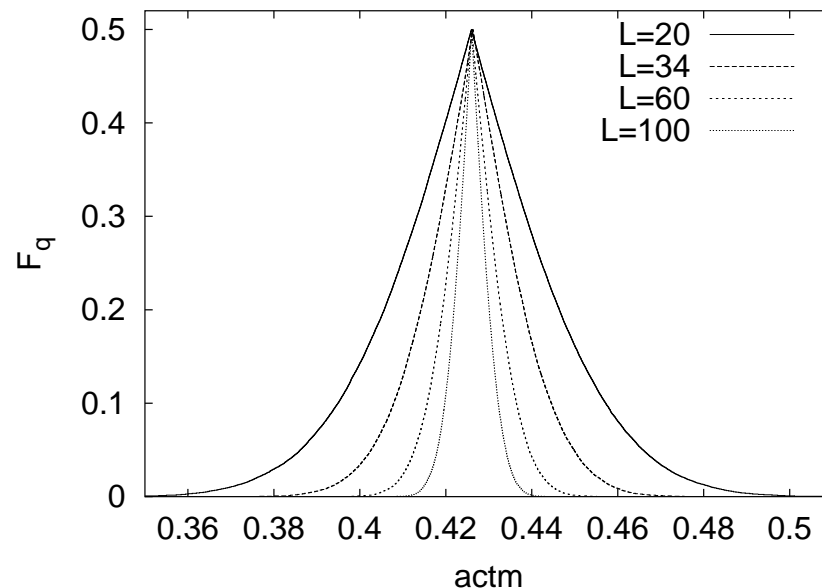

Figure 3: $O(3)$ σ-model at $\beta = 1.1$ (assignments a0304_06 and a0304_08).