Lecture V: Multicanonical Simulations.

1. Multicanonical Ensemble
2. How to get the Weights?
3. Example Runs ($2d$ Ising and Potts models)
4. Re-Weighting to the Canonical Ensemble
5. Energy and Specific Heat Calculation
6. Free Energy and Entropy Calculation
7. Summary

Multicanonical Ensemble

One of the questions which ought to be addressed before performing a large scale computer simulation is "What are suitable weight factors for the problem at hand?" So far we used the Boltzmann weights as this appears natural for simulating the Gibbs ensemble. However, a broader view of the issue is appropriate.

Conventional canonical simulations can by re-weighting techniques only be extrapolated to a vicinity of this temperature. For multicanonical simulations this is different. A single simulation allows to obtain equilibrium properties of the Gibbs ensemble over a range of temperatures. Of particular interest are two situations for which canonical simulations do not provide the appropriate implementation of importance sampling:

1. The physically important configurations are rare in the canonical ensemble.

2. A rugged free energy landscape makes the physically important configurations difficult to reach.

MC calculation of the interface tension of a first order phase transition provide an example where canonical MC simulation miss the important configurations.

Let $N = L^d$ be the lattice size. For first order phase transition pseudo-transition temperatures $\beta^c(L)$ exist so that the energy distributions $P(E) = P(E;L)$ become double peaked and the maxima at $E^1_{\max} < E^2_{\max}$ are of equal height $P_{\max} = P(E^1_{\max}) = P(E^2_{\max})$. In-between the maximum values a minimum is located at some energy $E_{\min}$. Configurations at $E_{\min}$ are exponentially suppressed like

$$P_{\min} = P(E_{\min}) = c_f \, L^p \, \exp(-f^s A) \tag{1}$$

where f^s is the interface tension and A is the minimal area between the phases, $A = 2L^{d-1}$ for an L^d lattice, c_f and p are constants (computations of p have been done in the capillary-wave approximation). The interface tension can be calculated by Binder's histogram method:

$$f^s(L) = -\frac{1}{A(L)} \ln R(L) \qquad \text{with} \qquad R(L) = \frac{P_{\min}(L)}{P_{\max}(L)} \tag{2}$$

and a finite size scaling (FSS) extrapolation of $f^s(L)$ for $L \to \infty$.

For large systems a canonical MC simulation will practically never visit configurations at energy $E = E_{\min}$ and estimates of the ratio $R(L)$ will be very inaccurate. The terminology supercritical slowing down was coined to characterize such an exponential deterioration of simulation results with lattice size.

Multicanonical simulations approach this problem by sampling, in an appropriate energy range, with an **approximation**

$$\widehat{w}_{mu}(k) = w_{mu}(E^{(k)}) = e^{-b(E^{(k)})\,E^{(k)} + a(E^{(k)})} \tag{3}$$

to the weights

$$\widehat{w}_{1/n}(k) = w_{1/n}(E^{(k)}) = \frac{1}{n(E^{(k)})} \tag{4}$$

where $n(E)$ is the spectral density. The function $b(E)$ defines the inverse **microcanonical temperature** and $a(E)$ the **dimensionless, microcanonical free**

energy. The function $b(E)$ has a relatively smooth dependence on its arguments, which makes it a useful quantity when dealing with the weight factors.

Instead of the canonical energy distribution $P(E)$, one samples a new multicanonical distribution

$$P_{mu}(E) = c_{mu}\, n(E)\, w_{mu}(E) \approx c_{mu}\ . \tag{5}$$

The desired canonical probability density is obtained by re-weighting

$$P(E) = \frac{c_\beta}{c_{mu}} \frac{P_{mu}(E)}{w_{mu}(E)}\, e^{-\beta E}. \tag{6}$$

This relation is rigorous, because the weights $w_{mu}(E)$ used in the actual simulation are exactly known. With the approximate relation (5) the average number of configurations sampled does not longer depend strongly on the energy and accurate

estimates of the ratio $R(L) = P_{\min}/P_{\max}$ become possible:

$$R(L) = R_{mu}(L)\,\frac{w_{mu}(E^i_{\max})\,\exp(-\beta^c\,E_{\min})}{w_{mu}(E_{\min})\,\exp(-\beta^c\,E_{\max})}\,, \quad (i = 1, 2)\ . \tag{7}$$

The statistical errors are those of the ratio R_{mu} times the exactly known factor.

The multicanonical method requires two steps:

1. Obtain a working estimate $\widehat{w}_{mu}(k)$ of the weights $\widehat{w}_{1/n}(k)$. Working estimate means that the approximation to (4) has to be good enough to ensure movement in the desired energy range, while deviations of $P_{mu}(E)$ from the constant behavior (5) are tolerable.

2. Perform a Markov chain MC simulation with the fixed weights $\widehat{w}_{mu}(k)$. The thus generated configurations constitute the multicanonical ensemble. Canonical expectation values are found by re-weighting to the Gibbs ensemble and standard jackknife methods allow reliable error estimates.

How to get the Weights?

To get the weights is at the heart of the method (and a stumbling block for beginners). Some approaches used are listed in the following:

1. Overlapping constrained (microcanonical) MC simulations. Potential problem: ergodicity.

2. Finite Size Scaling (FSS) Estimates. Best when it works! Problem: There may be no FSS theory or the prediction may be so inaccurate that the initial simulation will not cover the target region.

3. General Purpose Recursions. Problem: Tend to deteriorate with increasing lattice size (large lattices).

Multicanonical Recursion (A variant of Berg, J. Stat. Phys. 82 (1996) 323.)

Multicanonical parameterization of the weights:

$$\begin{aligned} w(a) = e^{-S(E_a)} &= e^{-b(E_a)\,E_a + a(E_a)}\,, \\ \text{where}\quad b(E) &= \left[S(E+\epsilon) - S(E)\right]/\epsilon \quad (\epsilon \text{ smallest stepsize}) \\ \text{and}\quad a(E-\epsilon) &= a(E) + \left[b(E-\epsilon) - b(E)\right]\,E\,. \end{aligned}$$

Recursion:

$$\begin{aligned} b^{n+1}(E) &= b^n(E) + \hat{g}_0^n(E)\left[\ln H^n(E+\epsilon) - \ln H^n(E)\right]/\epsilon \\ \hat{g}_0^n(E) &= g_0^n(E) \,/\, \left[g^n(E) + \hat{g}_0^n(E)\right], \\ g_0^n(E) &= H^n(E+\epsilon)\,H^n(E) \,/\, \left[H^n(E+\epsilon) + H^n(E)\right], \\ g^{n+1}(E) &= g^n(E) + g_0^n(E),\ g^0(E) = 0\,. \end{aligned}$$

ANIMATION: 2d 10-state 80×80 Potts model (Iterations of 100 sweeps each).

F. Wang and Landau: Update are performed with estimators $g(E)$ of the density of states

$$p(E_1 \to E_2) = \min\left[\frac{g(E_1)}{g(E_2)}, 1\right] \ .$$

Each time an energy level is visited, they update the estimator multiplicatively:

$$g(E) \to g(E)\, f \,,$$

where, initially, $g(E) = 1$ and $f = f_0 = e^1$. Once the desired energy range is covered, the factor f is refined:

$$f_1 = \sqrt{f}, \ f_{n+1} = \sqrt{f_{n+1}}$$

until a value sufficiently close to $1.$ is reached. Sufficiently close means: The system keeps on cycling (next page) with fixed weights.

Now the usual MUCA production runs should be carried out (not their philosophy).

Example Runs ($2d$ Ising and Potts models)

Ising model on a 20×20 lattice: The multicanonical recursion is run in the range

$$\texttt{namin} = 400 \leq \texttt{iact} \leq 800 = \texttt{namax} \; . \tag{8}$$

The recursion is terminated after a number of so called tunneling events. A tunneling event is defined as an updating process which finds its way from

$$\texttt{iact} = \texttt{namin} \quad \text{to} \quad \texttt{iact} = \texttt{namax} \quad \text{and} \quad \text{back} \; .$$

This notation comes from applications to first order phase transitions. An alternative notation for tunneling event is random walk cycle. For most applications 10 tunneling events lead to acceptable weights.

For an example run of the Ising model we find the requested 10 tunneling events after 787 recursions and 64,138 sweeps (assignment a0501_01, a0501_02 for the $2d$ 10-state Potts model).

Performance

If the multicanonical weighting would remove all relevant free energy barriers, the updating process would become a free random walk. Therefore, the theoretically optimal performance for the second part of the multicanonical simulation is

$$\tau_{\mathrm{tun}} \sim V^2 .$$

Recent work about first order transitions by Neuhaus and Hager shows that that the multicanonical procedure removes only the leading free energy barrier, while at least one subleading barrier causes still a residual supercritical slowing done. Up to certain medium sized lattices the behavior $V^{2+\epsilon}$ gives a rather good effective description. For large lattices supercritical slowing down of the form

$$\tau_{\mathrm{tun}} \sim \exp\left(+\mathrm{const}\, L^{d-1}\right)$$

dominates again. The slowing down of the weight recursion with the volume size is expected to be even (slightly) worse than that of the second part of the simulation.

Re-Weighting to the Canonical Ensemble

Let us assume that we have performed a multicanonical simulation which covers the action histograms needed for a temperature range

$$\beta_{\mathrm{min}} \;\leq\; \beta = \frac{1}{T} \;\leq\; \beta_{\mathrm{max}} \; .$$

Given the multicanonical time series, where $i = 1, \ldots, n$ labels the generated configurations. The formula

$$\overline{\mathcal{O}} = \frac{\sum_{i=1}^{n} \mathcal{O}^{(i)} \, \exp\left[-\beta \, E^{(i)} + b(E^{(i)}) \, E^{(i)} - a(E^{(i)})\right]}{\sum_{i=1}^{n} \exp\left[-\beta \, E^{(i)} + b(E^{(i)}) \, E^{(i)} - a(E^{(i)})\right]} \; .$$

replaces the multicanonical weighting of the simulation by the Boltzmann factor. The denominator differs from the partition function Z by a constant factor which drops out.

For discrete systems it is sufficient to keep histograms when only functions of the energy are calculated. For an operator $\mathcal{O}^{(i)} = f(E^{(i)})$ and reweighting simplifies to

$$\overline{f} = \frac{\sum_E f(E)\, h_{mu}(E)\, \exp\left[-\beta\, E + b(E)\, E - a(E)\right]}{\sum_E h_{mu}(E)\, \exp\left[-\beta\, E + b(E)\, E - a(E)\right]}$$

where $h_{mu}(E)$ is the histogram sampled during the multicanonical production run and the sums are over all energy values for which $h_{mu}(E)$ has entries.

The computer implementation of these equations requires care. The differences between the largest and the smallest numbers encountered in the exponents can be really large. We can avoid large numbers by dealing only with logarithms of sums and partial sums:

For $C = A + B$ with $A > 0$ and $B > 0$ we can calculate

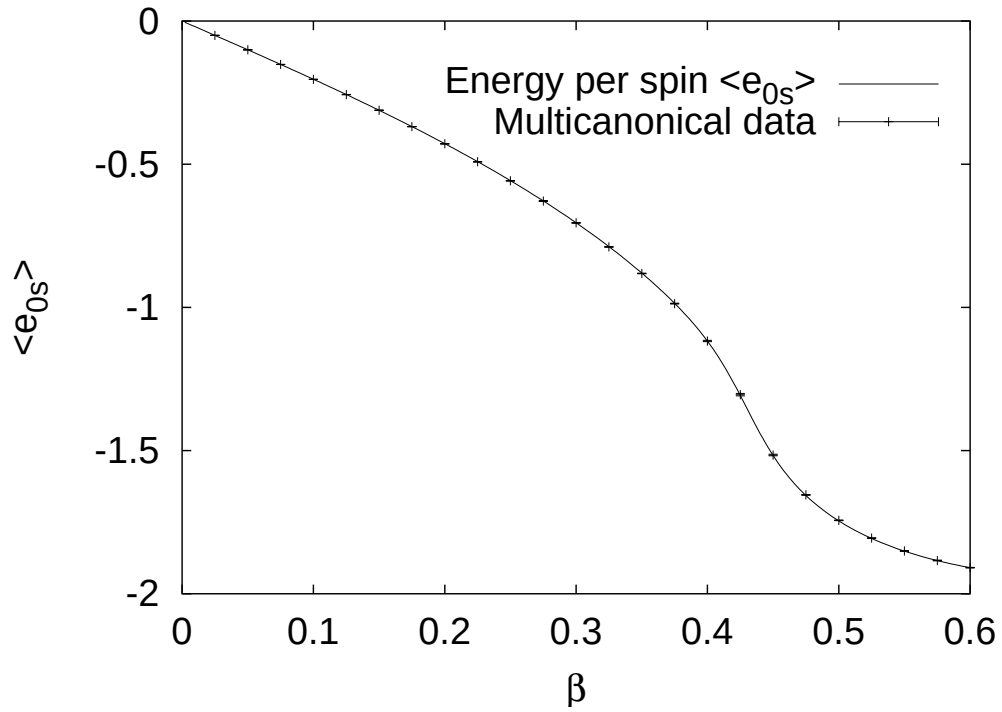
$$\ln C = \ln(A + B) \quad \text{from} \quad \ln A \quad \text{and} \quad \ln B$$

without ever storing either A or B or C:

$$\begin{aligned}
\ln C &= \ln\left[\max(A,B)\left(1 + \frac{\min(A,B)}{\max(A,B)}\right)\right] \\
&= \max(\ln A, \ln B) + \ln\{1 + \exp[\min(\ln A, \ln B) - \max(\ln A, \ln B)]\} \\
&= \max(\ln A, \ln B) + \ln\{1 + \exp[-|\ln A - \ln B|]\} \ .
\end{aligned}$$

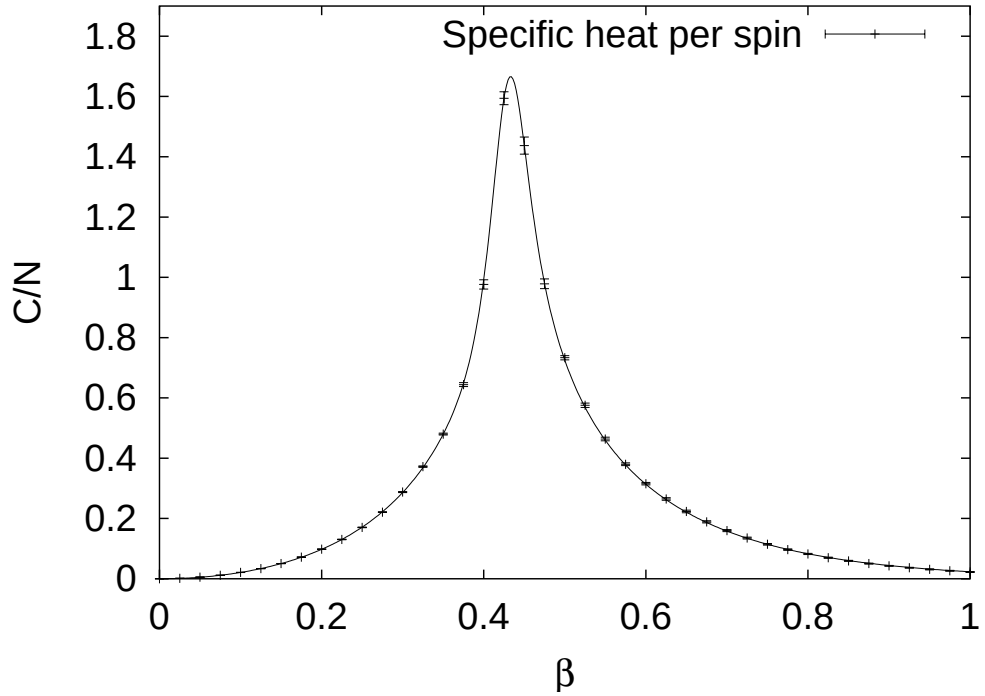
Energy and Specific Heat Calculation

We are now ready to produce multicanonical data for the **energy per spin** (with jackknife errors) of the $2d$ Ising model on a 20×20 lattice and to compare them with the exact results of Ferdinand and Fisher (assignment a0501_03):

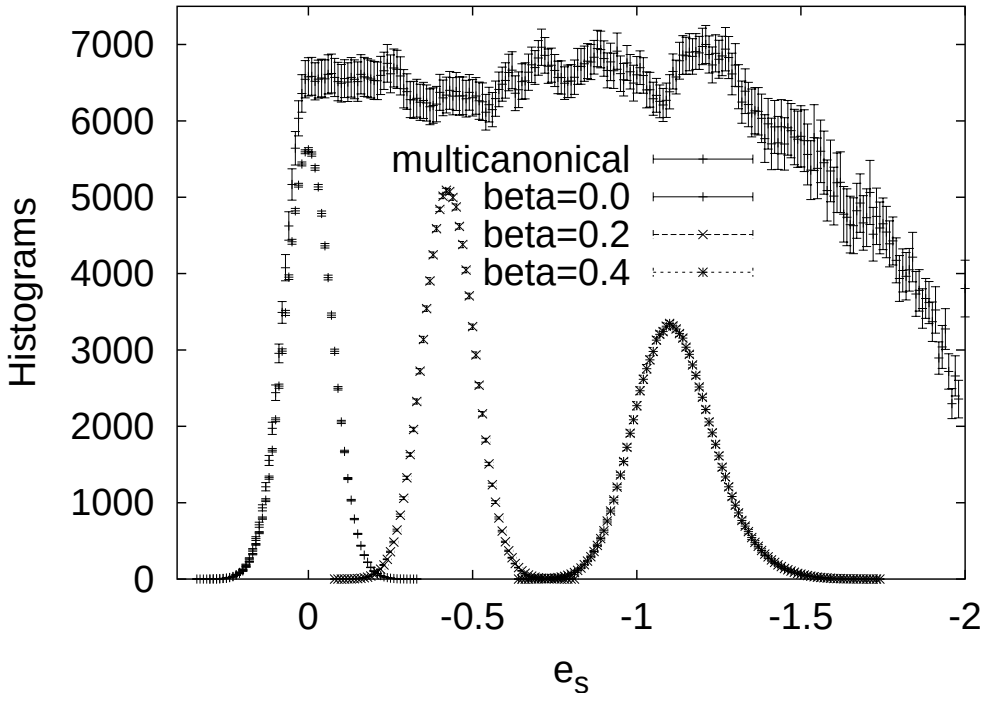


The same numerical data allow to calculate the **specific heat** defined by

$$C \;=\; \frac{d\,\widehat{E}}{d\,T} \;=\; \beta^2 \,\left(\langle E^2 \rangle - \langle E \rangle^2 \right) \;\; .$$

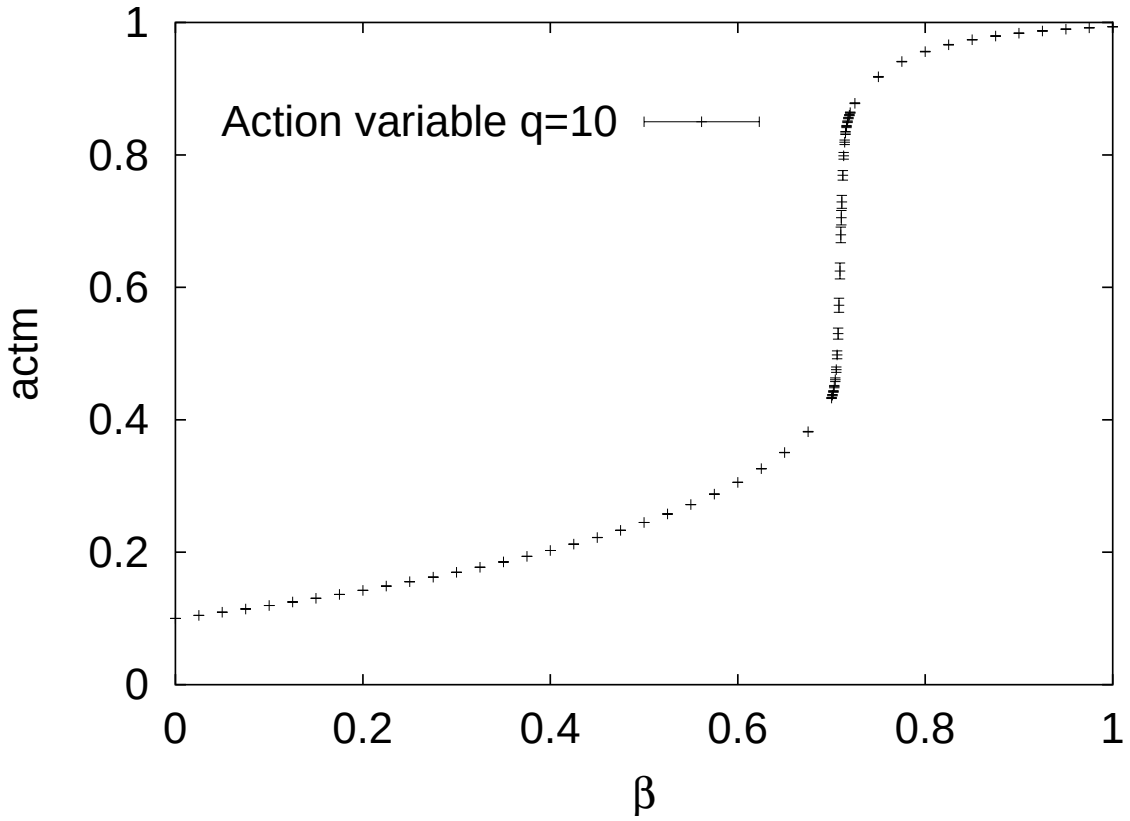


The energy histogram of the multicanonical simulation together its canonically re-weighted descendants at $\beta = 0$, $\beta = 0.2$ and $\beta = 0.4$ (assignment a0501_04):

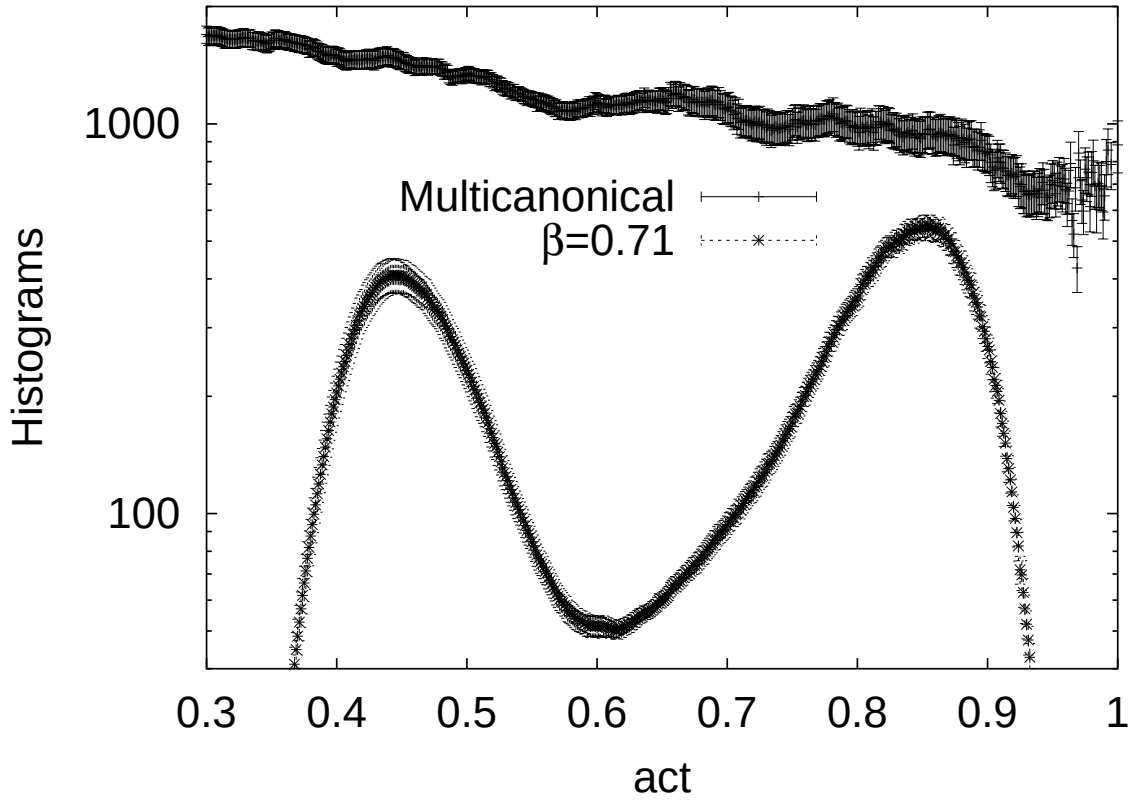


The normalization of the multicanonical histogram is adjusted so that it fits into the same figure with the three re-weighted histograms.

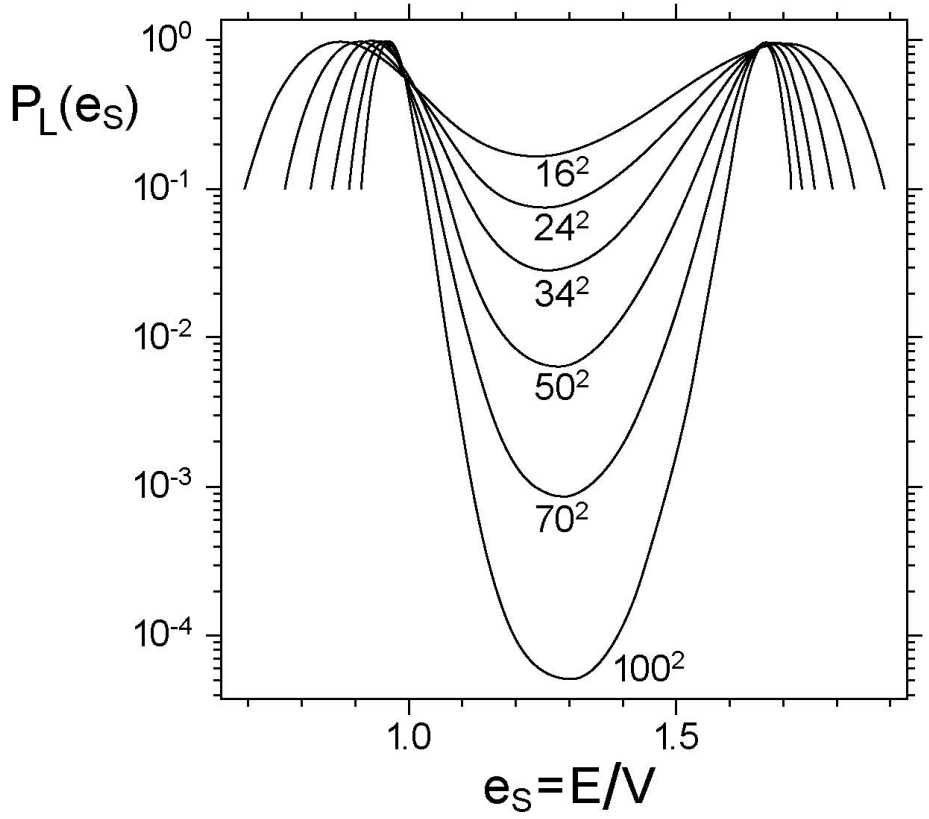
Similar data for the $2d$ 10-state Potts model, action variable `actm` data (assignment a0501_05):



The canonically re-weighted histogram at $\beta = 0.71$ together with the multicanonical histogram using suitable normalizations (the ordinate is on a logarithmic scale):



The multicanonical method allows to estimate the interface tension of the transition by following the minimum to maximum ratio over many orders of magnitude (Berg and Neuhaus 1992):



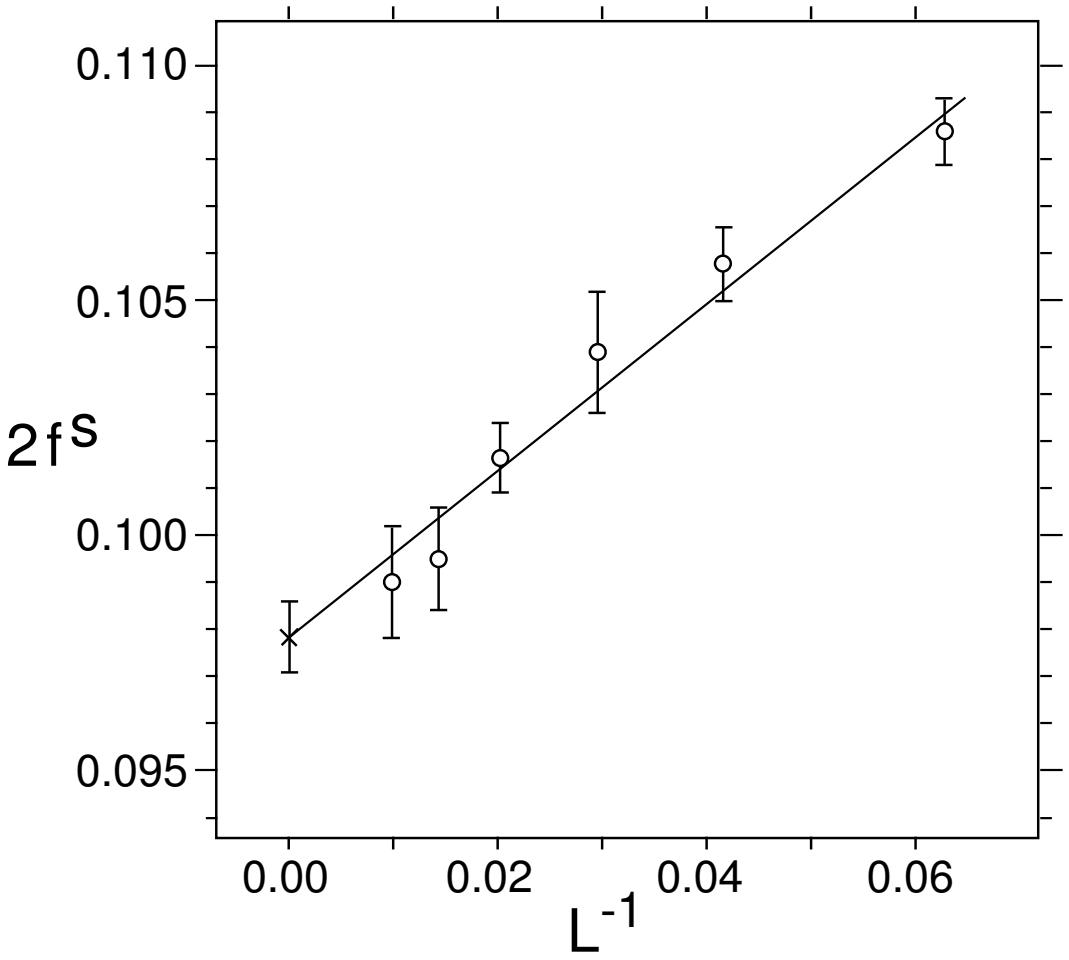


Figure 1: Interface tension fit (Berg and Neuhaus 1992).

Free Energy and Entropy Calculation

At $\beta = 0$ the Potts partition function is $Z_0 = q^N$. MUCA simulations allow for proper normalization of the partition function, if $\beta = 0$ is included in the temperature range. The properly normalized partition function allows to calculate the properly normalized

$$\text{Helmholtz free energy} \quad F \;=\; -\beta^{-1}\,\ln(Z)$$

and the

$$\text{Entropy} \quad S \;=\; \frac{F-E}{T} \;=\; \beta\,(F-E)$$

of the canonical ensemble. Here E is the expectation value of the internal energy and the last equal sign holds because of our choice of units for the temperature. For the $2d$ Ising model as well as for the $2d$ 10-state Potts model, we show in the next figure multicanonical estimates of the free energy density per site

$$f \;=\; F/N\ .$$

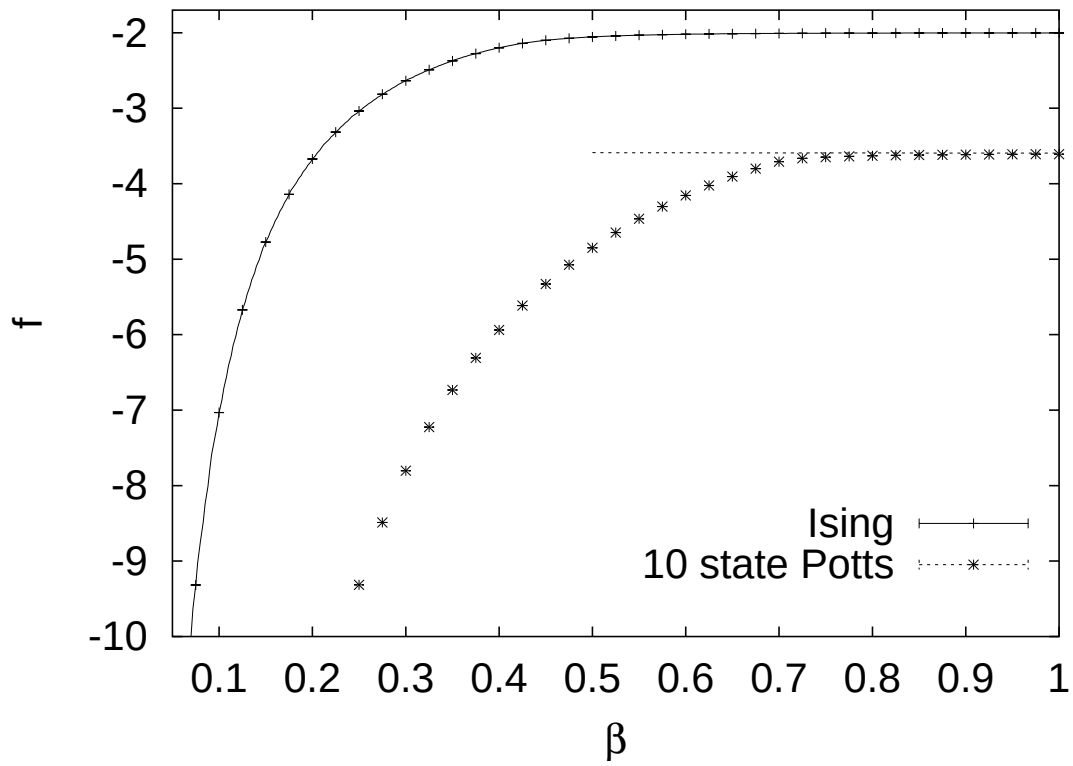


Figure 2: Free energies from multicanonical simulations of the Ising and 10-state Potts models on an 20×20 lattice (assignments a0501_03 and a0501_05). The full lines are the exact results of Ferdinand and Fischer for the Ising model and the asymptotic behavior $f_{\rm as} = \frac{2\,d}{q} - 2\,d - \beta^{-1}\,\frac{\ln(q)}{N}$ for the 10-state Potts model.

Entropy Density $s = S/N$:

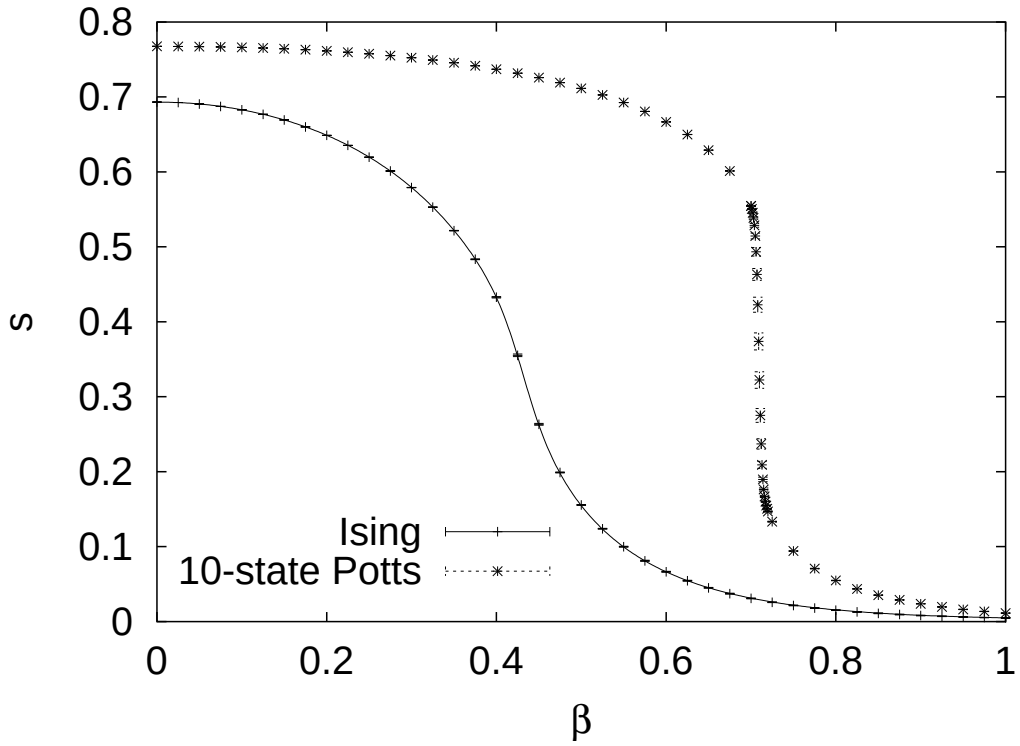


Figure 3: Entropies from multicanonical simulations of the Ising and 10-state Potts models on an 20×20 lattice (assignments a0501_03 and a0501_05). The full line is the exact result of Ferdinand and Fischer for the Ising model.

For the $2d$ Ising model one may also compare directly with the number of states. Up to medium sized lattices this integer can be calculated to all digits by analytical methods (Beale). MC results are only sensitive to the first few (not more than six) digits and one finds no real advantages over using other physical quantities.

Time series analysis

Typically, one prefers in continuous systems time series data over keeping histograms, because one avoids then dealing with the discretization errors. Even in discrete systems time series data are of importance, as one often wants to measure more physical quantities than just the energy. Then the RAM storage requirements may require to use a time series instead of histograms. We illustrate this point, using the Potts magnetization:

In assignments a0501_08 and a0501_09 we create the same statistics on 20×20 lattices as before, including time series measurements for the energy and also for the Potts magnetization. For energy based observables the analysis of the histogram and the time series data give consistent results.

For zero magnetic field, $H = 0$, the expectation value of the Potts magnetization on a finite lattice is

$$M_{q0} \;=\; \langle\, \delta_{q_i,q_0} \,\rangle \;=\; \frac{1}{q}\,,$$

independently of the temperature. For the multicanonical simulation each Potts state is visited with probability $1/q$ also at low temperatures. In contrast to this, the expectation value of the magnetization squared

$$M_{q0}^2 \;=\; q \left\langle \left(\frac{1}{N} \sum_{i=1}^{N} \delta_{q_i,q_0} \right)^2 \right\rangle$$

is a non-trivial quantity. At $\beta = 0$ its $N \to \infty$ value is $M_{q0}^2 = q\,(1/q)^2 = 1/q$, whereas it approaches 1 for $\beta \to \infty$. For $q = 2$ and $q = 10$ the next figure shows our numerical results and we see that the crossover of M_{q0}^2 from $1/q$ to 1 happens in the neighborhood of the critical temperature. A FSS analysis would reveal that, a singularity develops at β_c, which is in the derivative of M_{q0}^2 for the second order phase transitions $(q \leq 4)$ and in M_{q0}^2 itself for the first order transitions $(q \geq 5)$.

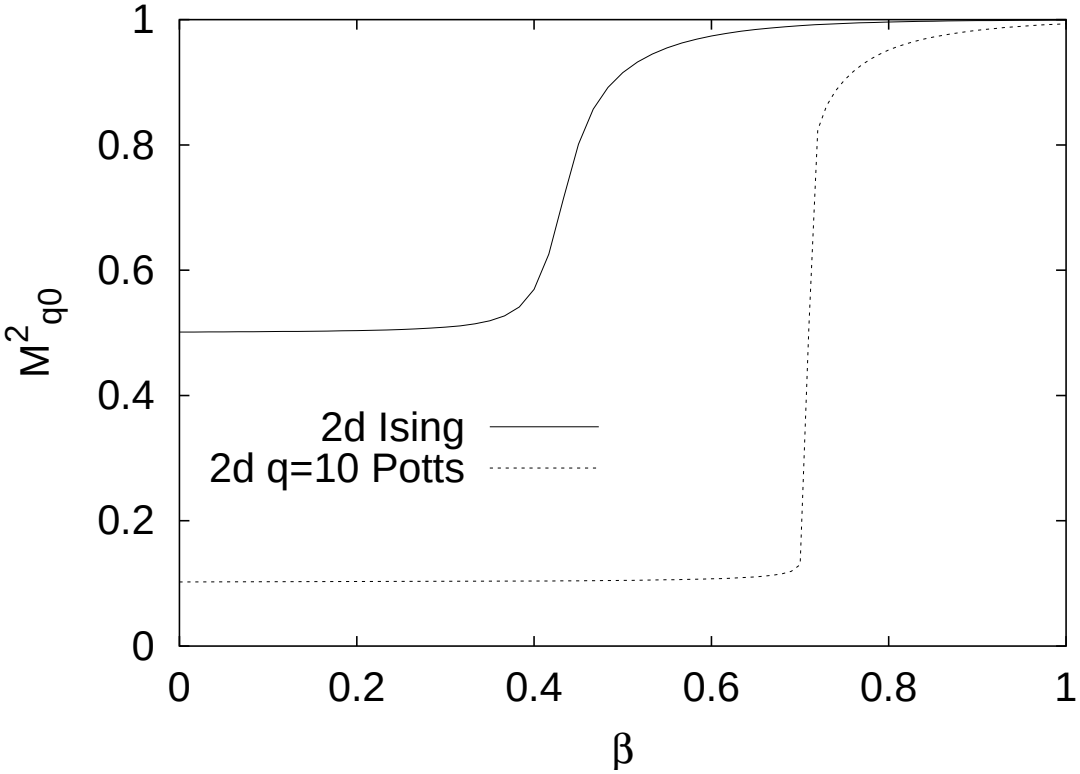


Figure 4: The Potts magnetization per lattice site squared for the $q = 2$ and $q = 10$ Potts models on an 20×20 lattice (assignments a0501_08 and a0501_09).

Summary

- We considered Statistics, Markov Chain Monte Carlo simulations, the Statistical Analysis of Markov chain data and, finally, Multicanonical Sampling.

- It is a strength of computer simulations that one can generate artificial (not realized by nature) ensembles, which enhance the probabilities of rare events on may be interested in, or speed up the dynamics.

- Nowadays Generalized Ensembles (umbrella, multicanonical, 1/k, ...) have found many applications. Besides for first order phase transitions in particular for Complex Systems such as biomolecules, where they accelerate the dynamics.