Reconstruction of the finite size canonical ensemble from incomplete micro-canonical data

P.H. Lundow*and K. Markström†

August 12, 2005

Abstract

In this paper we discuss how partial knowledge of the density of states for a model can be used to give good approximations of the energy distributions in a given temperature range. From these distributions one can then obtain the statistical moments corresponding to eg the internal energy and the specific heat. These questions have gained interest apropos of several recent methods for estimating the density of states of spin models.

As a worked example we finally apply these methods to the 3-state Potts model for cubic lattices of linear order up to 128. We give estimates of eg latent heat and critical temperature, as well as the microcanonical properties of interest.

1 Introduction

When studying a statistical mechanical model the most complete information is given by the density of states function. From complete knowledge of the density of states one can immediately work with the microcanonical ensemble and of course also compute the partition function and through it have access to the canonical ensemble as well. The main problem here is that computing the density of states for systems of even very modest size is typically very hard. However, recently several sampling schemes which strive to approximate the density of states have appeared. One recent method was given [WL01] and in [WS02] several such methods were given, and in [HRA+04a] all of the later methods as well as several others were united in a common framework.

For work in the microcanonical ensemble the mentioned methods give all the information needed. Using them one can find the density of states in an energy interval around the critical region and that is all that is needed for most investigations of the critical properties of the model. The microcanonical ensemble is more refined than the canonical ensemble in that every equilibrium measure for the canonical ensemble is found among the equilibrium measures for the microcanonical ensemble, but for some models there are microcanonical equilibria which are not present in the canonical ensemble. For a fuller survey of

*KTH Physics, AlbaNova University Center, SE-106 91 Stockholm, Sweden, phl@kth.se.

†Department of Mathematics and Mathematical Statistics, Umeå University, SE-901 87 Umeå, Sweden, klas.markstrom@math.umu.se.

the mathematical theory of ensemble equivalence see [TET04] and its references. This means that all properties of the thermodynamic limit can be obtained via the microcanonical ensemble.

However, even in view of what has been said the canonical ensemble has its own interest for finite systems. Among other things it governs the behaviour of many sampling algorithms and for systems where we have ensemble nonequivalence its dynamic can be very interesting. In order to reconstruct the canonical ensemble one would in principle need to know the density of states for all values of the energy E. However, using methods as in [HRA$^+$04a] this is very costly, and also not needed for work in the microcanonical ensemble.

Our aim is to look at how density of states data from a restricted interval of energies can be used to get an approximation of the energy distribution of the canonical ensemble for some range of couplings K. Thanks to the strong concentration of the energy distributions we will see that one can obtain a very good approximation of the energy distribution and through its moments most of the standard thermodynamical properties. This will be demonstrated first in a case where we know the exact partition function, the Ising model on the 256×256 square lattice, and then for a case where we have ensemble nonequivalence: the 3-state potts model on the 3-dimensional cubic lattice. All in all we find that with data collected with the methods of [HRA$^+$04a] in mind one can get a good picture of the canonical ensemble as well as the microcanonical. In fact, thanks to knowing the density of states for a full interval of energies we will be able to reconstruct the canonical ensemble for all couplings in some interval rather than just those used in the sampling process.

2 Notation

Let us define what we need in terms of the Ising model. Later on, when the Potts model is our subject, we will redefine some quantities, but our general discussion will be held in terms of the Ising model. Let G be a graph on n vertices $V = \{1, \dots, n\}$ and m edges. A state is a function $s : V \to Q$ where $Q = \{+1, -1\}$ and we say that vertex i has spin s_i. The energy of a state is defined as $E(s) = \sum_{ij} s_i s_j$ where the sum is taken over all edges ij of the graph and we have $-m \le E \le m$. The magnetisation of s is defined as $M(s) = \sum_i s_i$ so that $-n \le M \le n$.

A normalised energy and magnetisation will often be used, here defined as $U = E/m$ and $\mu = M/n$ so that $-1 \le U, \mu \le 1$. The number of states having energy E and magnetisation M is denoted $a(E, M)$. The number of states at energy E, or, the density of states, is denoted $a(E)$, where, of course, $a(E) = \sum_M a(E, M)$.

From quotients of $a(E)$ we obtain what we will refer to as the coupling function

$$K(U) = \frac{1}{\ell} \log \frac{a(E)}{a(E + \ell)}$$

where $U = E/m$ and ℓ is the difference between two consecutive energies. The very fundamental entropy function

$$S(U) = \frac{\log a(E)}{n}$$

is of course related to the coupling function through

$$K(U) = -\frac{n}{m} S'(U)$$

See [HRA+04a] for proofs and further details.

The partition function is defined for all graphs as

$$\mathcal{Z}(K, H) = \sum_{E,M} a(E, M) \, \exp(K\, E + H\, M)$$

where K and H are the dimensionless coupling and external field respectively. When the external field is zero we simplify as

$$\mathcal{Z}(K) = \sum_{E} a(E) \, \exp(K\, E)$$

As a convention we will write our coupling dependent quantities in a calligraphic font, such as $\mathcal{Z}(K)$.

The central moments of a random variable X are defined

$$\sigma_i = \langle (X - \langle X \rangle)^i \rangle, \quad i = 0, 1, \ldots$$

where $\sigma_0 = 1$, $\sigma_1 = 0$ and $\sigma_2 = \mathsf{Var}\,(X)$. The standard deviation is written $\sigma = \sqrt{\sigma_2}$. The cumulants κ_i of a distribution, or, the i:th derivatives of $\log \mathcal{Z}$, can be expressed in terms of moments. For the first few we have

$$\begin{aligned}
\frac{\partial \log \mathcal{Z}(K)}{\partial K} &= \kappa_1 = \langle E \rangle \\
\frac{\partial^2 \log \mathcal{Z}(K)}{\partial K^2} &= \kappa_2 = \mathsf{Var}\,(E) = \sigma_2 \\
\frac{\partial^3 \log \mathcal{Z}(K)}{\partial K^3} &= \kappa_3 = \sigma_3 \\
\frac{\partial^4 \log \mathcal{Z}(K)}{\partial K^4} &= \kappa_4 = \sigma_4 - 3\,\sigma_2^2
\end{aligned}$$

The free energy is here defined as

$$\mathcal{F}(K) = \frac{1}{n} \, \log \mathcal{Z}(K)$$

and the reader should note that we have used a simplified version compared to its traditional form. The internal energy, specific heat and coupling dependent entropy are given by

$$\mathcal{U}(K) = \frac{1}{m} \, \frac{\partial \log \mathcal{Z}(K)}{\partial K} = \frac{1}{m} \, \langle E \rangle.$$

$$\mathcal{C}(K) = \frac{1}{m} \, \frac{\partial^2 \log \mathcal{Z}(K)}{\partial K^2} = \frac{1}{m} \, \mathsf{Var}\,(E)\,.$$

$$\mathcal{S}(K) = \mathcal{F}(K) - \frac{m}{n} \, K \, \mathcal{U}(K)$$

We would also like to study the higher derivatives in the form of skewness

$$\Gamma_1(K) = \frac{\partial^3 \log \mathcal{Z}(K)/\partial K^3}{\left(\partial^2 \log \mathcal{Z}(K)/\partial K^2\right)^{3/2}} = \frac{\sigma_3}{\left(\sigma_2\right)^{3/2}} = \frac{\kappa_3}{\left(\kappa_2\right)^{3/2}}$$

and (excess) kurtosis

$$\Gamma_2(K) = \frac{\partial^4 \log \mathcal{Z}(K)/\partial K^4}{\left(\partial^2 \log \mathcal{Z}(K)/\partial K^2\right)^2} = \frac{\sigma_4}{\sigma_2^2} - 3 = \frac{\kappa_4}{\kappa_2^2}$$

Note that for normal distributions they both evaluate to zero.

Derivatives with respect to the field H are of course obtained analogously. The magnetisation and susceptibility are defined respectively as

$$\mu(K,H) = \frac{1}{n}\frac{\partial \log \mathcal{Z}(K,H)}{\partial H} = \frac{1}{n}\langle M \rangle$$
$$\chi(K,H) = \frac{1}{n}\frac{\partial^2 \log \mathcal{Z}(K,H)}{\partial H^2} = \frac{1}{n}\mathsf{Var}\,(M)$$

However, what one usually want is the spontaneous magnetisation and susceptibility. As finite size approximations of these we use

$$\bar{\mu}(K) = \frac{1}{n}\langle |M| \rangle$$
$$\bar{\chi}(K) = \frac{1}{n}\mathsf{Var}\,(|M|)$$

and assume that these converge to the appropriate limits.

Given a lattice of side L with L^3 vertices we call L the linear order of he lattice. When necessary we will subscript the functions with the linear, as in $\mathcal{Z}_L$.

3 Distributions of energy

We will assume that our sampled data contains information on quotients of consecutive density of states, or rather, that we have estimated the coupling function $K(U)$ for an interval of energies $u \leq U \leq v$. We would like to reconstruct the distribution of energies for a given coupling K_0. The process is rather straightforward and follows more or less by definition, but we will derive it in some detail.

3.1 From coupling to distribution

We assume the Boltzmann distribution for the states, that is, if we sample at a coupling K_0 the probability for our system being in state s is

$$\mathsf{Pr}\,(s) = \frac{\exp(K_0\, E(s))}{\mathcal{Z}(K_0)}$$

and consequently the probability for our system being in a state of energy E is

$$\mathsf{Pr}\,(E) = \frac{a(E)\, \exp(K_0\, E)}{\mathcal{Z}(K_0)} \tag{1}$$

Recall that we defined $a(E) = \exp(n\, S(U))$. Then we obtain

$$\mathsf{Pr}\,(E) = \frac{\exp\left(n\, S(U) + m\, U\, K_0\right)}{\mathcal{Z}(K_0)} \tag{2}$$

By definition we also have

$$S(U) = \int_{-1}^{U} S'(x)\,\mathrm{d}x = \int_{-1}^{u} S'(x)\,\mathrm{d}x + \int_{u}^{U} S'(x)\,\mathrm{d}x = A - \frac{m}{n}\int_{u}^{U} K(x)\,\mathrm{d}x$$

and trivially

$$U = u + \int_{u}^{U} 1\,\mathrm{d}x$$

Plugging these identities into Equation 2 and applying only a modicum of algebraic manipulation it simplifies finally into

$$\mathsf{Pr}\,(E) = c\,\exp\left(m\int_{u}^{U} K_0 - K(x)\,\mathrm{d}x\right) \tag{3}$$

Since the outcome is a probability function the constant c is defined by normalising so that

$$\sum_{E} \mathsf{Pr}\,(E) = 1 \tag{4}$$

where the sum is taken over all energies E such that $u \leq E/m \leq v$. Finally, we note in passing that the derivative of the probability function with respect to U is $m\,(K_0 - K(U))\,\mathsf{Pr}\,(E)$. Thus the points where the sign of the derivative changes is determined by when $K_0 = K(U)$.

Note that we have only defined the function $K(U)$ at discrete points $U = E/m$ so we should be somewhat careful with how the integral is taken. If a function $f(x)$ is defined at $a = x_0 < x_1 < \ldots < x_p = b$ then we use a left-point rule for integration

$$\int_{a}^{b} f(x)\,\mathrm{d}x = \sum_{i=0}^{p-1} f(x_i)\,(x_{i+1} - x_i)$$

Having reconstructed the distribution of energies the moments and cumulants are easily retrieved. First the average

$$\langle E\rangle = \sum_{E} E\,\mathsf{Pr}\,(E)$$

and then the central moments

$$\sigma_i = \sum_{E} \left(E - \langle E\rangle\right)^i\,\mathsf{Pr}\,(E)$$

and from these we obtain the sought-after estimates of the derivatives by evaluating the cumulants of the distribution so that, for example

$$\kappa_4 = \sigma_4 - 3\,\sigma_2^2 = \frac{\partial^4 \log \mathcal{Z}}{\partial K^4}$$

Let us address the issue of derivatives with respect to the field as well. If we during our sampling process remembered to collect data on the magnetisation as well, then we can reconstruct the spontaneous magnetisation and susceptibility

as well. Our program should then collect raw moments on the form $\langle |M|^i \mid E \rangle$. Then the following holds

$$\bar{\mu}(K_0) = \frac{1}{n} \langle |M| \rangle = \frac{1}{n} \langle \langle |M| \mid E \rangle \rangle = \frac{1}{n} \sum_E \langle |M| \mid E \rangle \, \mathsf{Pr}(E) \tag{5}$$

$$\bar{\chi}(K_0) = \frac{1}{n} \mathsf{Var}(|M|) = \frac{1}{n} \left(\sum_E \langle |M|^2 \mid E \rangle \, \mathsf{Pr}(E) - \left(\sum_E \langle |M| \mid E \rangle \, \mathsf{Pr}(E) \right)^2 \right) \tag{6}$$

The following is a nice alternative way of writing the variance

$$\mathsf{Var}(|M|) = \langle \mathsf{Var}(|M| \mid E) \rangle + \mathsf{Var}(\langle |M| \mid E \rangle)$$

that is, the variance is the sum of the expectation of the variances and the variance of the expectations.

3.2 Accuracy of the reconstructed distributions

Since the canonical ensemble is always determined by the density of states we only have two sources of errors: the precision of the original data and the truncation error due to not having data from all energies.

In a perfect world the collected data comes from the entire interval of energies $-1 \leq U \leq 1$. However, normally it suffices for the interval to be wide enough to cover the energies at coupling K_0 with a high probability. In short, the distribution of energies corresponding to K_0 must stay in the interval $[u, v]$ with a probability close to 1. If $[u, v]$ only covers say, 99% or less of the energies you see at K_0, the normalising step in Equation 4 will produce erroneous results.

Given a coupling K_0 that is close to the critical coupling K_c we expect the distribution to be anything but normal (ie gaussian). But, as we move away from K_c the distribution typically becomes close to normal. For example, at $K = 0$ the distribution is clearly approaching a normal one with increasing system size. It has been shown, see [ML73], that this also holds for Ising systems when K is greater than some $K_1 > K_c$. Since our approach is somewhat pragmatic we will only assume that far enough from K_c the energy distributions can be treated as roughly normal.

For how large or small values of K_0 does the procedure return a credible distribution on $[u, v]$? Treating the distribution as roughly normal it should be enough, for all practical purposes, to make sure that the end-points are at least 4, if possible 5 and preferably 6 standard deviations σ away.

The probability density of a normally distributed variable is

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)$$

We will translate it slightly to the right, ie take $f(x - p)$ for $p > 0$, and cut it off at $x = 0$. Now let X be a random variable with probability density

$$g(x) = \begin{cases} \frac{f(x-p)}{A(p)} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

where $A(p)$ is the mass of probability on $x > 0$, ie

$$A(p) = \int_0^\infty f(x-p)\,\mathrm{d}x$$

so that $g(x)$ becomes a cut-off, but otherwise normal looking, probability density on the real axis. For which p is the y-axis, ie the cut-off point, located k standard deviations σ away? Numerical calculations gave Table 1 below and in Figure 1 the distribution functions are shown for $k = 2, 3, 4$. In the Table we also list the errors

$$\epsilon_i = |\kappa_i(g) - \kappa_i(f)|$$

that is, we take the difference in cumulants for the cut-off density g and the normal density f translated k standard deviations. These errors are of course very idealised, being based on normal distributions, and should be considered rough guidelines. For any particular distribution we will see different errors and especially the higher cumulants will deviate from these.

k	p	ϵ_1	ϵ_2	ϵ_3	ϵ_4
2	1.728042230	$2{\cdot}10^{-1}$	$2{\cdot}10^{-1}$	$2{\cdot}10^{-1}$	$2{\cdot}10^{-1}$
3	2.973669402	$2{\cdot}10^{-2}$	$1{\cdot}10^{-2}$	$4{\cdot}10^{-2}$	$8{\cdot}10^{-2}$
4	3.998789790	$1{\cdot}10^{-3}$	$5{\cdot}10^{-4}$	$2{\cdot}10^{-3}$	$7{\cdot}10^{-3}$
5	4.999979927	$2{\cdot}10^{-5}$	$7{\cdot}10^{-6}$	$4{\cdot}10^{-5}$	$2{\cdot}10^{-4}$
6	5.999999885	$6{\cdot}10^{-7}$	$4{\cdot}10^{-8}$	$2{\cdot}10^{-7}$	$1{\cdot}10^{-6}$

Table 1: Peak location p and cumulant errors for a cut-off normal distribution with $\langle X \rangle = k\,\sigma$.

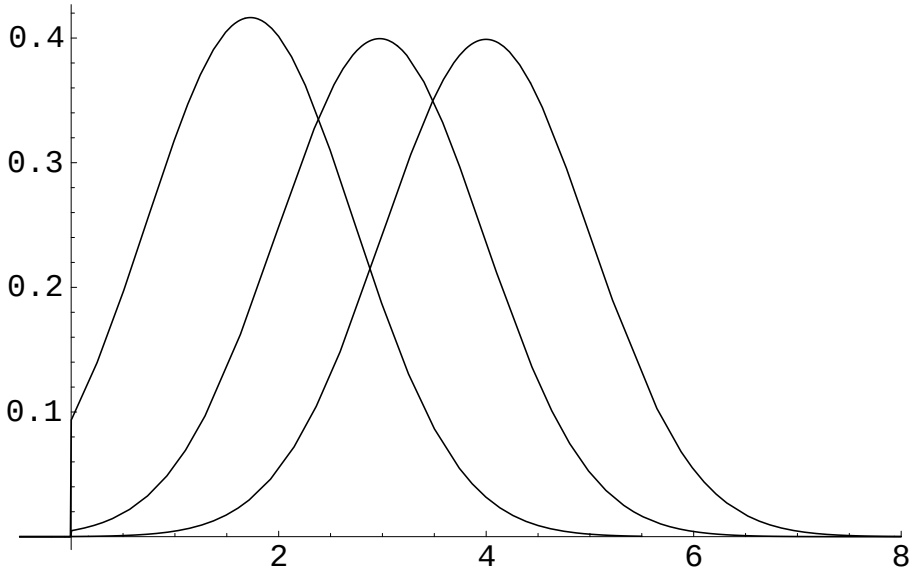


Figure 1: Cut-off normal distributions such that $\langle X \rangle = k\,\sigma$ for $k = 2, 3, 4$.

4 The 2-dimensional Ising model

We will employ the 2D Ising model as a test bed for our method. Recall that the critical coupling is $K_c = \operatorname{arctanh}(\sqrt{2}-1) \approx 0.4407$ and that the critical energy is $U_c = 1/\sqrt{2} \approx 0.7071$. In [HRA$^+$04b] we computed the exact partition function for the 256×256-lattice with periodic boundary. However, since the largest density of states $a(0)$ has 19726 digits we will take the liberty of doing all actual computations with 50 digits numerical precision instead.

Suppose now that we have collected data on $K(U)$ for $u = 0.6 \leq U \leq 0.8 = v$, an interval comprising 6554 energies. In Figure 2 we plot $K(U)$ and $K'(U)$ for $L = 256$. From the exact (50 digits) coupling function on the interval $[0.6, 0.8]$

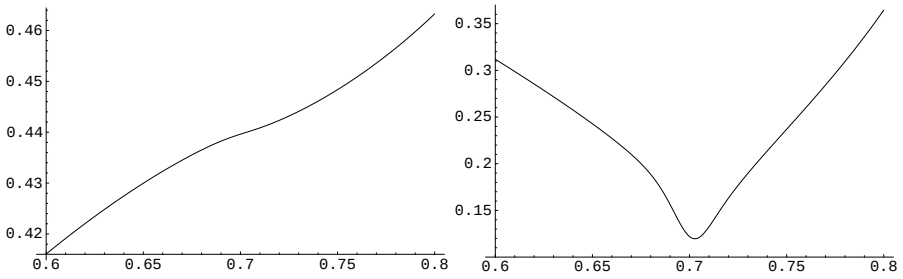


Figure 2: $K(U)$ and $K'(U)$ for 256×256-lattice.

we reconstruct the distribution, ie the probability density function, of energies at K_c using Equation 3. Let ϵ_i denote the relative error of the i:th cumulant where we compare the cumulant κ_i of the reconstructed energy distribution with the i:th derivative of $\log \mathcal{Z}$, ie

$$\epsilon_i = \left| \frac{\kappa_i}{\partial^i \log \mathcal{Z} / \partial K^i} - 1 \right|$$

The relative errors $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4$ are negligibly small, less than $1{\cdot}10^{-30}$. However, this distribution lives clearly in the middle of our energy interval $[0.6, 0.8]$, the lower bound being 14σ below and the upper bound 12σ above the mean. In Table 2 we compute relative errors of the cumulants when the coupling corresponds to a cut-off distribution with $\langle E \rangle$ located $k\,\sigma$ from the lower bound $u = 0.6$ for $k = 2, 3, 4, 5, 6$ and in Table 3 we do the corresponding at the other end of the interval so that $\langle E \rangle$ is $k\,\sigma$ from the upper bound $v = 0.8$. Figure 3 shows the probability densities at $K = 0.423780$, $K = K_c$ and $K = 0.454942$.

k	K	ϵ_1	ϵ_2	ϵ_3	ϵ_4
2	0.418736	$7{\cdot}10^{-4}$	$2{\cdot}10^{-1}$	$1{\cdot}10^{+1}$	$9{\cdot}10^{+1}$
3	0.420644	$4{\cdot}10^{-5}$	$1{\cdot}10^{-2}$	$1{\cdot}10^{0}$	$4{\cdot}10^{+1}$
4	0.422226	$9{\cdot}10^{-7}$	$4{\cdot}10^{-4}$	$6{\cdot}10^{-2}$	$2{\cdot}10^{0}$
5	0.423780	$7{\cdot}10^{-9}$	$4{\cdot}10^{-6}$	$8{\cdot}10^{-4}$	$4{\cdot}10^{-2}$
6	0.425338	$2{\cdot}10^{-11}$	$1{\cdot}10^{-8}$	$3{\cdot}10^{-6}$	$2{\cdot}10^{-4}$

Table 2: Relative errors of cumulants for cut-off distribution with $\langle X \rangle = u + k\,\sigma$.

k	K	ϵ_1	ϵ_2	ϵ_3	ϵ_4
2	0.460385	$5{\cdot}10^{-4}$	$2{\cdot}10^{-1}$	$6{\cdot}10^{0}$	$5{\cdot}10^{+1}$
3	0.458331	$2{\cdot}10^{-5}$	$1{\cdot}10^{-2}$	$8{\cdot}10^{-1}$	$2{\cdot}10^{+1}$
4	0.456623	$5{\cdot}10^{-7}$	$3{\cdot}10^{-4}$	$3{\cdot}10^{-2}$	$1{\cdot}10^{0}$
5	0.454942	$4{\cdot}10^{-9}$	$3{\cdot}10^{-6}$	$3{\cdot}10^{-4}$	$1{\cdot}10^{-2}$
6	0.453254	$7{\cdot}10^{-12}$	$7{\cdot}10^{-9}$	$9{\cdot}10^{-7}$	$4{\cdot}10^{-5}$

Table 3: Relative errors of cumulants for cut-off distribution with $\langle X\rangle = v - k\,\sigma$.

We also computed the cumulant errors at K_c with the upper and lower bound of the energy interval located $6\,\sigma$ away. For $L = 32, 64, 128, 256$ the errors are quite small, $\epsilon_1 < 1{\cdot}10^{-12}$, $\epsilon_2 < 2{\cdot}10^{-10}$, $\epsilon_3 < 2{\cdot}10^{-8}$ and $\epsilon_4 < 6{\cdot}10^{-8}$. For $L \leq 16$, the errors become larger, but on the other hand for such small graphs it is easy to collect a complete set of $K(U)$-data instead of only a short interval.

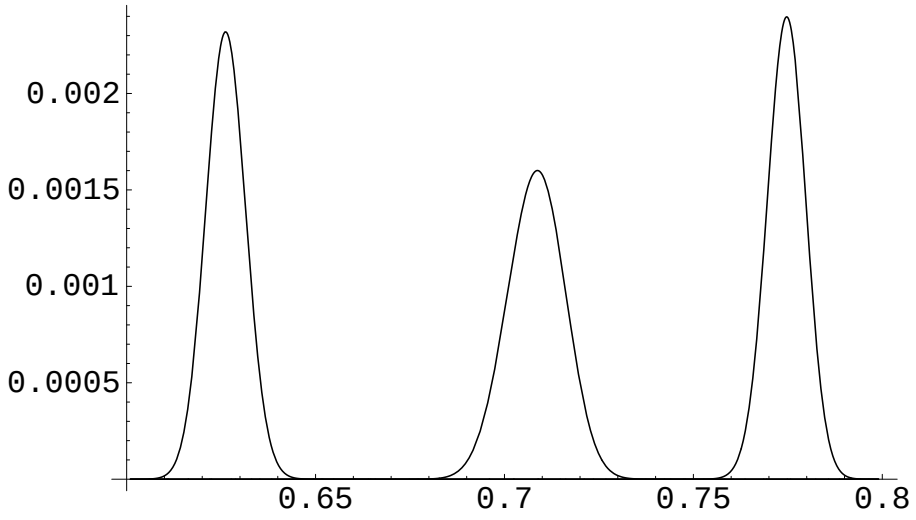


Figure 3: Probability densities of energies at $K_1 = 0.423780$ (left), K_c (middle) and $K_2 = 0.454942$ (right) for 256×256-lattice. At K_1 and K_2 the interval bounds are 5σ away. Probability $\mathsf{Pr}\,(E)$ on the y-axis and energy $U = E/m$ on the x-axis.

Regarding the magnetisation and susceptibility we have no way of comparing the reconstructed values with exact values. We have simply run the Metropolis method at 10 different temperatures in the vicinity of K_c ($0.42 \leq K \leq 0.46$) and collected magnetisation moments at each energy level. Using these data and the exact K-function we can reconstruct the magnetisation and susceptibility at any temperature in that region using Equation 5 and 6 adding data as prescribed in [HRA$^+$04a]. See Figure 4 for a snap-shot of the result. The reconstructed curves agrees very well with the sampled data.

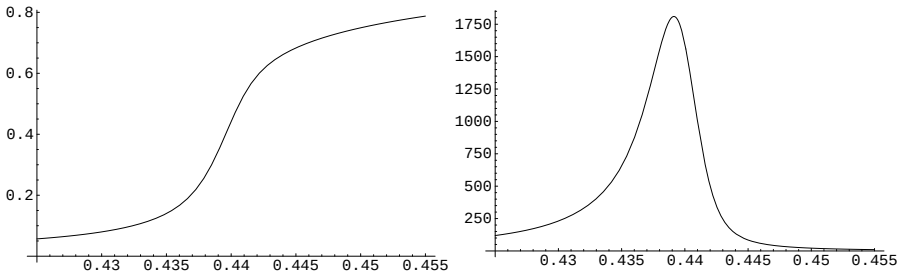


Figure 4: $\bar{\mu}(K)$ and $\bar{\chi}(K)$ for 256×256-lattice.

5 The free energy

By definition we have that

$$\mathcal{F}(K) = \mathcal{F}(0) + \frac{m}{n} \int_0^K \mathcal{U}(x)\,\mathrm{d}x$$

where the constant $\mathcal{F}(0) = \log 2$ for the Ising model, and $\mathcal{F}(0) = \log q$ for the q-state Potts model. Having evaluated $\mathcal{U}(K)$ for a number of values of K this is of course easily accomplished. Unfortunately, this formulation implies that we have collected data so that the energy distribution can be reconstructed for $K \geq 0$. For smaller graphs this is of course perfectly alright but for large graphs this was exactly what we wanted to avoid. However, due to the well-behaved nature of the internal energy $\mathcal{U}(K)$ we can circumvent this problem. Suppose we have reconstructed the internal energy for two system sizes L_1 and L_2, where $L_1 < L_2$. Suppose further that we have $\mathcal{U}_{L_1}(K)$ for $0 \leq K \leq b_1$ and $\mathcal{U}_{L_2}(K)$ for $a \leq K \leq b_2$ where $0 \leq a \leq K_c \leq b_2 \leq b_1$. Then, for $a \leq K \leq b_2$ we have

$$\mathcal{F}(K) = \mathcal{F}(0) + \frac{m}{n} \int_0^a \mathcal{U}_{L_1}(x)\,\mathrm{d}x + \frac{m}{n} \int_a^K \mathcal{U}_{L_2}(x)\,\mathrm{d}x + \epsilon$$

where ϵ is an error term.

How big is the error? Let f and g be continuous functions on the interval $[a, c]$ with $a < b < c$. Then the following elementary calculation gives an estimate:

$$\begin{aligned}
\int_a^c g(x)\,\mathrm{d}x &= \int_a^b g(x)\,\mathrm{d}x + \int_b^c g(x)\,\mathrm{d}x = \\
&= \int_a^b g(x) - f(x) + f(x)\,\mathrm{d}x + \int_b^c g(x)\,\mathrm{d}x = \\
&= \int_a^b f(x)\,\mathrm{d}x + \int_b^c g(x)\,\mathrm{d}x + \epsilon
\end{aligned}$$

where ϵ is the error term

$$\epsilon = \int_a^b g(x) - f(x)\,\mathrm{d}x$$

which gives the very simple but useful estimate

$$|\epsilon| \leq (b - a) \max_{a \leq x \leq b} |g(x) - f(x)| \tag{7}$$

Since the internal energy function is an increasing and, in fact, convex function, it is easy to establish the maximum. The integration is numerical so it is important to evaluate $\mathcal{U}(K)$ at points chosen densely enough, with special attention to values close to K_c where $\mathcal{U}(K)$ is expected to change rapidly.

5.1 A worked example for the 2D Ising model

Here our goal is to compute the free energy at $K = K_c$ for the 256×256 2D Ising model by using a sequence of system sizes, $L = 32, 64, 128, 256$, and formulate the method as

$$\mathcal{F}_{256}(K_c) = \log 2 + 2\int_0^{0.15} \mathcal{U}_{32}(x)\,\mathrm{d}x + 2\int_{0.15}^{0.30} \mathcal{U}_{64}(x)\,\mathrm{d}x + \\ +2\int_{0.30}^{0.40} \mathcal{U}_{128}(x)\,\mathrm{d}x + 2\int_{0.40}^{K_c} \mathcal{U}_{256}(x)\,\mathrm{d}x$$

though the integration will of course be numerical. To evaluate the $\mathcal{U}_L(x)$ at the couplings indicated by the integral boundaries we use the exact (to 50 digits precision) K-functions at intervals wide enough to keep the endpoints $6\,\sigma$ away for each energy distribution. Picking simple values gives the following energy intervals, coupling intervals and step lengths used for evaluating $\mathcal{U}_L(K)$.

$$\begin{array}{llll} L = 32, & -0.15 \le U \le 0.30, & 0.00 \le K \le 0.15, & h = 0.0010 \\ L = 64, & 0.05 \le U \le 0.45, & 0.15 \le K \le 0.30, & h = 0.0005 \\ L = 128, & 0.30 \le U \le 0.60, & 0.30 \le K \le 0.40, & h = 0.0002 \\ L = 256, & 0.50 \le U \le 0.75, & 0.40 \le K \le 0.441, & h = 0.0001 \end{array}$$

The intervals for K were chosen to give good overlaps, and are of course model dependent.

Numerical integration of these data points, using eg the trapezoidal method gives that $\mathcal{F}_{256}(K_c) \approx 0.92970521$. The correct answer is $\mathcal{F}_{256}(K_c) = 0.92970516\ldots$ giving an error of $5.5{\cdot}10^{-8}$.

The error contribution given by Equation 7 is only of the order of $6{\cdot}10^{-12}$. The main error source is actually the numerical integration. As is well-known, numerical evaluation of $\int_a^b f(x)\,\mathrm{d}x$ with the trapezoidal rule gives the error term

$$\epsilon = -\frac{(b-a)\,h^2}{12}\,f''(\xi), \quad a < \xi < b$$

where h is the step length. Since the function we integrate is $\mathcal{U}(K)$ its second derivative is

$$\mathcal{U}''(K) = \frac{1}{m}\,\frac{\partial^3 \log \mathcal{Z}(K)}{\partial K^3} = \frac{\kappa_3}{m}$$

and it is at a very little extra cost we evaluate the third cumulant when we already have the distribution.

For example, the error contributed from

$$2\int_{a=0}^{b=0.15} \mathcal{U}_{32}(x)\,\mathrm{d}x$$

is at most

$$\frac{2\,(b-a)\,h^2}{12}\,\max_{a<x<b} \mathcal{U}''_{32}(x) = \frac{2\,(0.15-0)\,0.001^2}{12} 1.67 \approx 4.2{\cdot}10^{-8}$$

and the errors contributed by the other integrals are at most respectively $3.4 10^{-8}$ for $L = 64$, $1.6 \cdot 10^{-8}$ for $L = 128$ and $5.8 \cdot 10^{-8}$ for $L = 256$ and they sum up to $1.5 \cdot 10^{-7}$ which is clearly larger than the actual error we received.

6 An example with a first order phase transition: The 3-dimensional 3-state Potts model

For this model we need to redefine some of our quantities. A state is here a function $s : V \to Q$ where Q is a set of q distinct elements, eg $Q = \{1, \ldots, q\}$. The energy is defined as $E(s) = \sum_{ij} \delta(s_i, s_j)$, where $\delta(x, y)$ is the Kronecker-delta, so that $0 \leq E \leq m$. We normalise as before and let $U = E/m$ so that $0 \leq U \leq 1$. Let $\eta_j = \sum_i \delta(s_i, j)$, ie the number of vertices having spin j. For the Potts model, the definition of magnetisation M varies slightly in the literature. For example, $M = \max(\eta_1, \ldots, \eta_q)$ is sometimes used, but $M = \eta_1^2 + \ldots + \eta_q^2$ is the one used here. This means that $n^2/q \leq M \leq n^2$ and we normalise by taking $\bar{\mu} = \mu = M/n^2$ so that $1/q \leq \bar{\mu} \leq 1$. Having defined these quantities their physical versions follow accordingly.

6.1 The sampled data

The data were generated and collected by using the sampling method described in detail in [HRA$^+$04a] and we refer to that paper for further details. Since this model is conjectured to have a first order phase transition and cluster methods thus are expected to have exponential mixing time [CJF$^+$99] we opted for a highly optimised single spin Metropolis method. We used up to a few hundred independent spin systems, which after slowly being brought to the right coupling were given a few days or weeks, depending on their size, of continuous running for mixing. The length of the sampling runs were of the same order.

For the smaller lattices we collected at least 10000 measurements per energy level, often orders of magnitude more. For the larger lattices, say $L \geq 32$, this quickly becomes difficult. For $L = 96, 128$ we did not manage to fill out all energy levels inside the energy jump at critical the coupling, though these empty levels are *very* few relatively speaking. From these data we then constructed the coupling function $K(U)$.

The coupling functions $K(U)$, from which the energy distributions are generated, are shown in the left plot of Figure 5 and the right plot shows the magnetisations $\mu(U)$. Note that the K-functions behave rather different from that of the Ising model in Figure 2. Here the K-functions have their own set of critical points and they are listed in Table 4. Let U^- and U^+ be the locations of the maximum and the minimum respectively of $K(U)$. The corresponding values of K at these points are denoted K^- and K^+ respectively. We define the latent heat as $U^\pm = U^+ - U^-$. Let also μ^+ and μ^- denote the magnetisation at respectively U^+ and U^-.

6.2 The reconstructed ensemble

Next we used the U-dependent functions from the previous section to reconstruct the K-dependent quantities in the way described earlier in the paper. Let us here stress that none of the functions plotted here were sampled directly, ie we

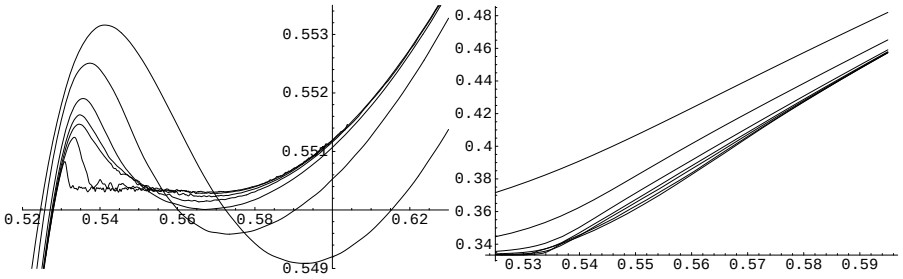


Figure 5: $K(U)$ (left) and $\mu(U)$ (right) for $L \geq 16$.

L	U^-	U^+	$U^\pm$	K^-	K^+	μ^-	μ^+
6	0.54502	0.60677	0.06174	0.553398	0.548924	0.4174	0.5125
8	0.54138	0.59245	0.05101	0.553159	0.549088	0.3930	0.4779
12	0.53851	0.57979	0.04128	0.552756	0.549367	0.3694	0.4442
16	0.53721	0.57327	0.03606	0.552507	0.549588	0.3580	0.4264
24	0.53603	0.56884	0.03281	0.552133	0.549861	0.3474	0.4119
32	0.53564	0.56717	0.03153	0.551905	0.550012	0.3427	0.4057
48	0.53478	0.56645	0.03167	0.551625	0.550151	0.3380	0.4012
64	0.53457	0.56796	0.03340	0.551465	0.550233	0.3363	0.4029
96	0.53337	0.56816	0.03479	0.551242	0.550284	0.3344	0.4019
128	0.53090	0.56785	0.03694	0.550826	0.550306	0.3336	0.4004

Table 4: Critical points and values of the combinatorial functions.

did not keep track of the variance and expectation of the energy in the sampling runs and everything here is based on the microcanoncial data.

First we will show a quick gallery of pictures of the physical, ie coupling dependent, quantities that were defined in Section 2. In Figure 6 we show the free energy $\mathcal{F}$ and the internal energy $\mathcal{U}$ in a narrow region around K_c for the larger lattices. The dramatic jump in the energy of course leads to a similar behaviour in the entropy $\mathcal{S}(K)$ shown in Figure 7. This is also seen in the magnetisation $\mu(K)$ in the same figure. We find that everything agrees well with the expected first order nature of the phase transition.

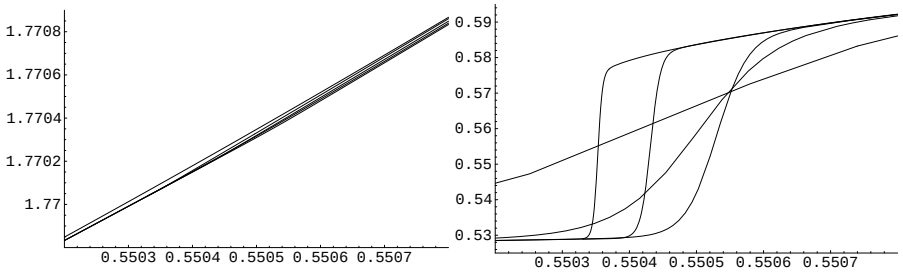


Figure 6: Free energy $\mathcal{F}(K)$ (left) and internal energy $\mathcal{U}(K)$ (right) for $L \geq 32$.

The specific heat $\mathcal{C}(K)$ is shown for $L = 64$ in the left plot of Figure 8. The maximum of this quantity grows very fast with L and to be able to compare

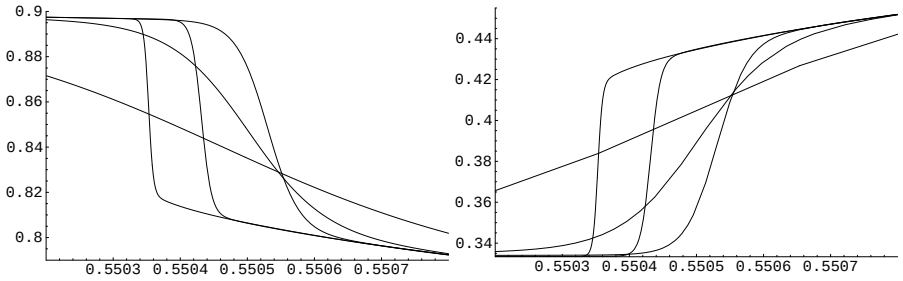


Figure 7: Entropy $\mathcal{S}(K)$ (left) and magnetisation $\mu(K)$ (right) for $L \geq 32$.

them for several L the right plot shows their logarithm. This can also be said

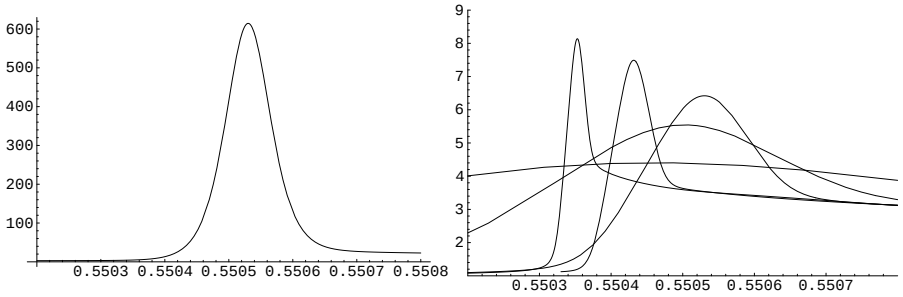


Figure 8: Specific heat $\mathcal{C}(K)$ for $L = 64$ (left) and its logarithm for $L \geq 32$ (right).

about the skewness and kurtosis in Figure 9. These quantities changes sign in a number of critical points so taking logarithms is not advisable. The distributions

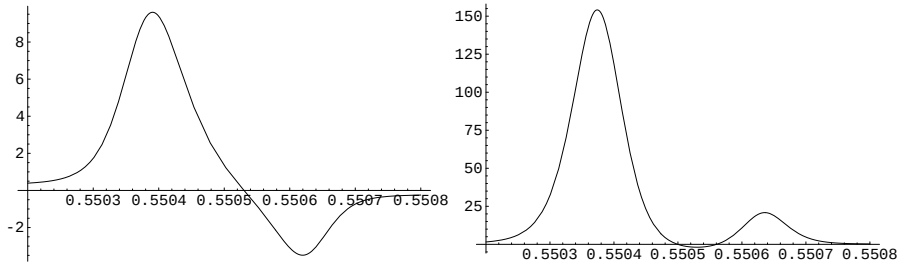


Figure 9: Skewness $\Gamma_1(K)$ (left) and kurtosis $\Gamma_2(K)$ (right) for $L = 64$.

go through a sharply bimodal phase as the coupling moves past K_c. Define K^* as the point where the specific heat has its maximum. What do the distributions at this point look like? In the left plot of Figure 10 the probability densities $p(x)$ of the normalised variable $x = (E - \langle E \rangle)/\sigma(E)$ at K^* are shown, while the right plot shows the distribution functions (accumulated densities) defined as

$$\Phi(x) = \int_{-\infty}^{x} p(t)\,\mathrm{d}t.$$

Note, by the way, that the peaks in the probability densities are very close $\pm\sigma$.

In Table 5 the data connected with K^* are listed.

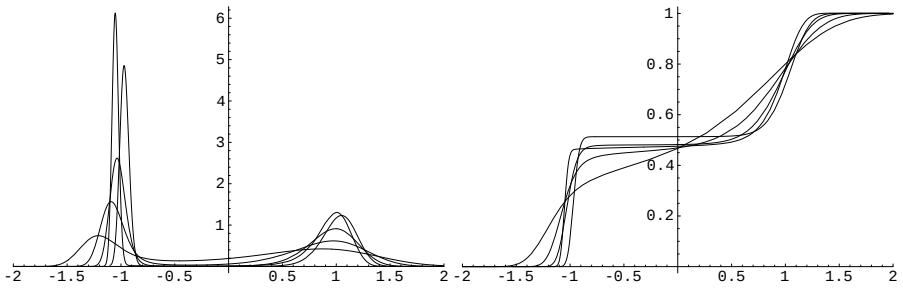


Figure 10: Normalised probability densities (left) and distribution functions (right) vs σ at K^* for $L \geq 32$.

L	K^*	$\mathcal{C}(K^*)$	$\mathcal{U}(K^*)$	$\mathcal{F}(K^*)$	$\mathcal{S}(K^*)$	μ^*
6	0.555045	5.1279	0.63766	1.78444	0.72265	0.5557
8	0.552821	7.2889	0.61575	1.77697	0.75577	0.5130
12	0.551143	12.539	0.59190	1.77220	0.79355	0.4643
16	0.550665	19.584	0.57983	1.77093	0.81306	0.4381
24	0.550460	41.929	0.56831	1.77034	0.83185	0.4118
32	0.550462	81.880	0.56318	1.77028	0.84026	0.3996
48	0.550502	255.97	0.55921	1.77032	0.84678	0.3901
64	0.550530	614.58	0.55821	1.77036	0.84842	0.3877
96	0.550432	1790.0	0.55550	1.77020	0.85291	0.3821
128	0.550353	3422.7	0.55279	1.77008	0.85738	0.3763

Table 5: Critical points K^* and values of the physical quantities.

6.3 Asymptotics

In this section we will see how some of the values in the tables above scale with the linear order. First we wish to establish the critical coupling K_c. We have three separate sequences of critical points which should all converge to K_c, namely K^*, K^+ and K^-. The coefficients of the fits descibred below are collecrted in Table 6.3.

The sequences K^+ and K^- from Table 4 have the nice feature that they are monotone; K^+ is increasing and K^- is decreasing. Unfortunately though, the K^--sequence appears slightly blemished for $L = 128$, as is the U^--sequence. Even so, after discarding that particular point and assuming that the sequences stay monotone also for larger L we then have upper and lower bounds on K_c. Then $0.550306 \leq K_c \leq 0.551242$, a rather wide interval. The story is different for the K^*-sequence, it sometimes increases, sometimes decreases, but we see nothing amiss with the value for $L = 128$.

To establish a K_c we have attempted a simple fit of the form $c_0 + c_1\, L^{-\lambda}$ for some coefficients c_0, c_1 and exponent λ to the data for K^*. To find the parameters we used Mathematica's non-linear fitting function. Our fitting function applied to this sequence gives acceptable fits. We will try to estimate the error

in this approach by fitting a curve to data of the form $L_{\min} \leq L \leq 128$ with $L_{\min} = 6, 8, 12$, ie for three sets of data. Leaving out more points gives the fitted curve an unconvincing look. This gave

$$K_c = 0.550425 \pm 0.000025$$

which agrees with the previous interval. This estimate is a little lower than that of [JV97] (who also provide a nice table of previous results) but their data is based on rather small graphs, $L \leq 36$. On the other hand, our estimate ends up right in the middle of the (rather wide) interval given by [GE94].

From the interval above we choose the mid-point as our limit, ie we set $K_c = c_0 = 0.550425$, and fit all points to determine the remaining parameters. Using the same limit we fitted curves to the K^- (discarding $L = 128$) and K^+ data. We received the curves shown with the points in the left plot of Figure 11.

A different behaviour is expected from the three sequences of energies; U^+, U^- and U^*. Here $U^+ \to U_c^+$ and $U^- \to U_c^-$ and the difference $U^\pm = U^+ - U^-$ should converge to the latent heat $U_c^\pm$, whereas $\mathcal{U}^*$ should converge to some value $\mathcal{U}_c$ between U^- and U^+. Again we see a possibly too big jump in the data for U^- at $L = 128$ so we will discard this point. Applying the process we described above we received

$$\begin{aligned} U_c^- &= 0.5322 \pm 0.0013 \\ U_c^+ &= 0.5670 \pm 0.0004 \\ \mathcal{U}_c &= 0.5513 \pm 0.0008 \end{aligned}$$

Again, using the mid-points as limits we received the curves shown with the data points in the right plot of Figure 11. Using these estimates of U_c^+ and U_c^- also provides us with estimates of the asymptotic latent heat $U^\pm$. This resulted in

$$U_c^\pm = U_c^+ - U_c^- = 0.03480 \pm 0.0017$$

which is clearly smaller than the estimate 0.0538 (after division by 3) of [JV97], but, as we recall, their estimates were based on much smaller systems.

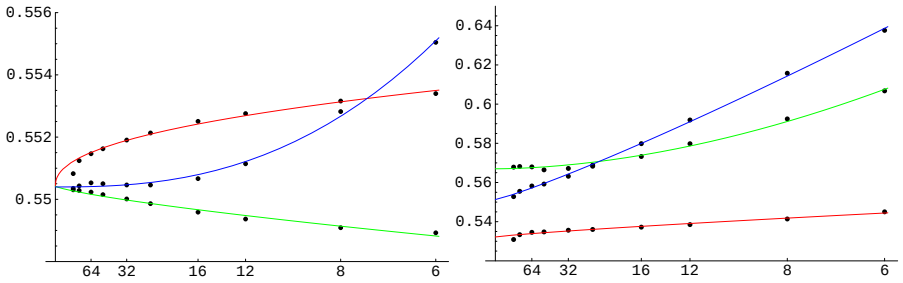


Figure 11: Left: couplings K^-, K^*, K^+ (downwards) vs $1/L$. Right: energies U^+, U^*, U^- (downwards at y-axis) vs $1/L$.

Applying this procedure to the free energy, where $\mathcal{F}(K^*) \to \mathcal{F}_c$, and the entropy, where $\mathcal{S}(K^*) \to \mathcal{S}_c$, we obtained

$$\begin{aligned} \mathcal{F}_c &= 1.77018 \pm 0.00005 \\ \mathcal{S}_c &= 0.86020 \pm 0.0015 \end{aligned}$$

Note the considerably larger error in $\mathcal{S}_c$ but also that the value is consistent with taking

$$\mathcal{S}_c = \mathcal{F}_c - 3\,K_c\,\mathcal{U}_c = 0.85983 \pm 0.0025$$

though we receive a slightly larger error estimate. The data points and the fitted curves are shown in Figure 12.

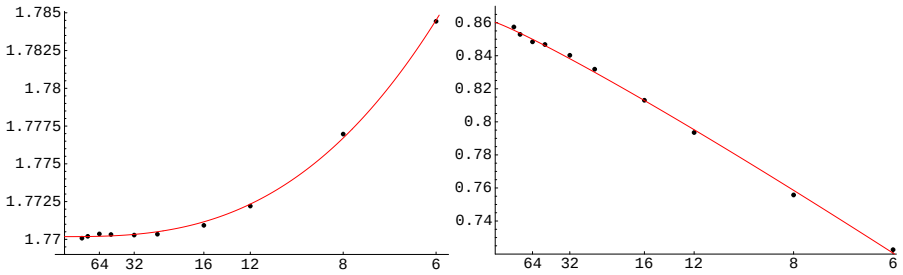
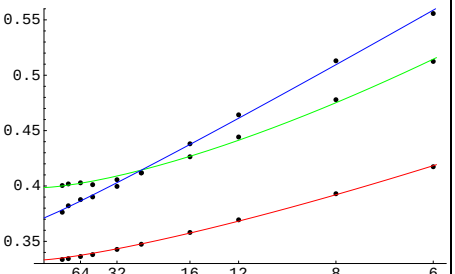


Figure 12: Free energy $\mathcal{F}(K^*)$ (left) and entropy $\mathcal{S}(K^*)$ (right) vs $1/L$ together with fitted curves.

For the magnetisations μ^-, μ^+ and μ^* we should see a behaviour analogous to that of the energies and we have treated them as such. We assume that $\mu^- \to \mu_c^- = 1/3$, $\mu^* \to \mu_c$ and $\mu^+ \to \mu_c^+$. Continuing with our trusted approach we received

$$\mu_c^+ = 0.3986 \pm 0.0015$$
$$\mu_c = 0.3711 \pm 0.0027$$

and the fitted curves are shown with the data points in Figure 6.3.



	c_0	c_1	λ
K^*	0.550425	0.465327	2.5679
K^-	0.550425	0.006770	0.4407
K^+	0.550425	-0.006277	0.7618
U^-	0.5322	0.05402	0.8295
U^+	0.5670	1.02718	1.8039
$\mathcal{U}^*$	0.5513	0.67094	1.1376
$\mathcal{F}(K^*)$	1.77018	1.92325	2.7342
$\mathcal{S}(K^*)$	0.86020	-1.00861	1.1037
μ^-	0.3333	0.8463	1.2827
μ^*	0.3711	1.2524	1.0593
μ^+	0.3986	1.5332	1.4412

Figure 13: Left: magnetisations μ^+, μ^*, μ^- (downwards at y-axis) vs $1/L$. Right: coefficients for our fitted curves

References

[CJF+99] C.Borgs, J.Chayes, A. Frieze, J.H.Kim, P.Tetali, E.Vigoda, and V.Vu, *Torpid mixing of some mcmc algorithms in statistical phys-*

ics, Proceedings of FOCS '99 (1999), 218–229., Preprint available at http://www.math.cmu.edu/~af1p/papers.html.

[GE94] A. J. Guttmann and I. G. Enting, *Series studies of the Potts model. III. The* 3*-state model on the simple cubic lattice*, J. Phys. A **27** (1994), no. 17, 5801–5812.

[HRA+04a] Roland Häggkvist, Anders Rosengren, Daniel Andrén, Petras Kundrotas, Per Håkan Lundow, and Klas Markström, *A Monte Carlo sampling scheme for the Ising model*, J. Statist. Phys. **114** (2004), no. 112, 455–480.

[HRA+04b] Roland Häggkvist, Anders Rosengren, Daniel Andrén, Petras Kundrotas, Per Håkan Lundow, and Klas Markström, *Computation of the Ising partition function for two-dimensional square grids*, Phys. Rev. E **69** (2004), no. 4, 046104.

[JV97] Wolfhard Janke and Ramon Villanova, *Three-dimensional* 3*-state Potts model revisited with new techniques*, Nuclear Phys. B **489** (1997), no. 3, 679–696.

[ML73] Anders Martin-Löf, *Mixing properties, differentiability of the free energy and the central limit theorem for a pure phase in the Ising model at low temperature*, Comm. Math. Phys. **32** (1973), 75–92.

[TET04] Hugo Touchette, Richard S. Ellis, and Bruce Turkington, *An introduction to the thermodynamic and macrostate levels of nonequivalent ensembles*, Phys. A **340** (2004), no. 1-3, 138–146, News and expectations in thermostatistics.

[WL01] F. Wang and D.P. Landau, *Efficient, multiple-range random walk algorithm to calculate the density of states*, Phys. Rev. Let. **86** (2001), no. 10, 2050–2053.

[WS02] Jian-Sheng Wang and Robert H. Swendsen, *Transition matrix Monte Carlo method*, J. Statist. Phys. **106** (2002), no. 1-2, 245–285.