Multicanonical Ensemble: A New Approach to Simulate First-Order Phase Transitions

Bernd A. Berg [1],[2],[a] and Thomas Neuhaus [1]

[1]*Fakultät für Physik, Universität Bielefeld, D-4800 Bielefeld, Federal Republic of Germany*
[2]*Supercomputer Computations Research Institute, Tallahassee, Florida 32306*

Relying on the recently proposed multicanonical algorithm, we present a numerical simulation of the first-order phase transition in the 2D 10-state Potts model on lattices up to sizes 100×100. It is demonstrated that the new algorithm *lacks* an exponentially fast increase of the tunneling time between metastable states as a function of the linear size L of the system. Instead, the tunneling time diverges approximately proportional to $L^{2.65}$. On our largest lattice we gain more than 2 orders of magnitude as compared to a standard heat-bath algorithm. As a first physical application we report a high-precision computation of the interfacial free energy per unit area.

Critical slowing down is of crucial importance to computer simulations of phase transitions. For second-order phase transitions long autocorrelation times at criticality cause severe restrictions on the maximum lattice size for which one can obtain good statistics of thermodynamic quantities. For a number of spin systems this critical slowing down was overcome by the nonlocal cluster algorithm of Swendsen and Wang [1]; for a recent review see [2]. However, for first-order transition one encounters an even worse and different problem of critical slowing down. The interfacial free energy between disordered and ordered states has a finite value on the critical point for the infinite volume system. Configurations dominated by the presence of the interface will be exponentially suppressed by the Boltzmann factor in the canonical ensemble. On finite lattices this leads then to an exponentially fast suppression of the tunneling between metastable states of the system with increasing lattice size. To overcome this critical slowing down effect for first-order transitions, we recently proposed a *multicanonical* Monte Carlo (MC) algorithm [3]. The multicanonical MC algorithm is designed to enhance configurations which are dominated by the presence of the interface and therefore exponentially suppressed. In this way it is possible to avoid the exponentially fast growing slowing down at the first-order phase transition. In this paper we demonstrate this in the case of our example: the 2D 10-state Potts model.

The 2D 10-state Potts model [4] is defined by the partition function

$$Z(\beta) = \sum_{\text{configurations}} \exp(\beta S), \tag{1}$$

$$S = \sum_{\langle i,j \rangle} \delta_{q_i, q_j}, \tag{2}$$

$$q_i, q_j = 0, \dots, 9. \tag{3}$$

Recently there has been renewed interest in this model [5-7]. It serves as an excellent laboratory for finite-size-scaling (FSS) studies of temperature-driven strong first-order phase transitions. To calculate the interfacial free energy F^s between the disordered and the (ten) ordered

states has remained the hardest problem. The reason is the pronounced double-peak structure of the sampled action density $P_L(S)$ in the canonical ensemble near the critical point. The pseudocritical point β_L^c is defined such that both maxima are of equal height:

$$P_L^{1,\max} = P_L(S_L^{1,\max}) = P_L(S_L^{2,\max}) = P_L^{2,\max}. \tag{4}$$

In addition we have imposed the normalization condition $1 = P_L^{1,\max} = P_L^{2,\max}$. Figure 1 depicts the action densities for lattices with $L = 16, 24, 34, 50, 70$, and 100 on a logarithmic scale and we see that 4 orders of magnitude are involved: $P_L^{\min}/P_L^{\max} \approx 5.1 \times 10^{-5}$ for $L = 100$. With our conventions for $P_L(S)$ the interfacial free energy per unit area $F^s = F_\infty^s$ can now be defined [8] as the $L \to \infty$ limit of the quantity

$$F_L^s = -\frac{1}{L} \ln P_L^{\min}. \tag{5}$$

For a numerical calculation of F_L^s it is now clear that any algorithm which samples configurations with a probability $\sim P_L(S)$ would slow down proportional to $1/P_L^{\min}$. As for large lattices $P_L^{\min} \sim \exp(-F^s L^{d-1})$, it is expected that an appropriately defined tunneling time τ_L^t will be-
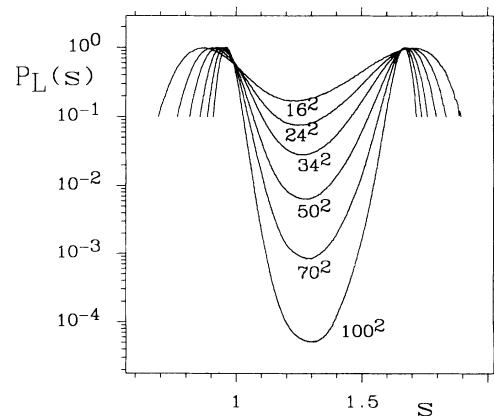

FIG. 1. Action density distributions $P_L(s)$ for lattices of size $L = 16, 24, 34, 50, 70$, and 100 on a logarithmic scale. The values of the maxima have been normalized to 1.

have as

$$\tau'_L = A_\tau L^\alpha e^{+F'L^{d-1}}. \tag{6}$$

The parameters A_τ and α can, in principle, be determined by a fit to the measured tunneling times.

The multicanonical MC algorithm samples configurations with the weight

$$\mathcal{P}_L^{MC}(S) \sim e^{(\alpha_L^k + \beta_L^k S)} \text{ for } S_L^k < S \le S_L^{k+1} \tag{7}$$

instead of sampling with the usual Boltzmann factor $P_L^\beta(S) \sim \exp(\beta_L^c S)$ corresponding to the canonical ensemble. Here we partitioned the total action interval $0 \le S \le 2V$ into $k = 0, \ldots, N$ ($N+1$ odd) intervals $I_k = (S_L^k, S_L^{k+1}]$. The idea of the multicanonical MC algorithm is to choose intervals I_k and values of β_L^k and α_L^k at the pseudocritical point β_L^c in such a way that the resulting multicanonical action density $\mathcal{P}_L(S)$ has an approximately flat behavior for values of the action in the interval $[S_L^{1,max}, S_L^{2,max}]$; that is to say, configurations dominated by the interface are no longer exponentially suppressed as they are in the metastable-unstable region of the canonical ensemble. Physically this can be achieved by choosing the β parameters β_L^k such that the system gets heated when it is in the ordered state of the metastable region, cooled when it is in the disordered state, and neither of these if it is in the unstable region. The parameters β_L^k hereby take the form $\beta_L^k = \beta_L^c + \delta\beta_L^k$, where the coupling constant difference $\delta\beta_L^k$ changes sign as a function of S and is responsible for the altered dynamics of the model. The parameters α_L^k are adjusted in such a way that $\mathcal{P}_L(S)$ is a steady function of S.

In [3] we demonstrated, that when the double-peak distribution $P_L(S)$ can be approximated by a double Gaussian, the multicanonical action density $\mathcal{P}_L(S)$ can be made arbitrarily flat by driving a control parameter $r > 1$ towards 1. In this case we choose action values S_L^k with $S_L^0 = 0$, $S_L^{N+1} = 2V$, $S_L^1 = S_L^{1,max}$, $S_L^N = S_L^{2,max}$ and in the interval $[S_L^{1,max}, S_L^{min})$ action values defined by the equation $P_L(S_L^k) = r^{1-k} S_L^{1,max}$ for $k = 1, \ldots, N/2$. An analog procedure is adopted in the interval $(S_L^{min}, S_L^{2,max}]$. Having defined the action values S_k and corresponding intervals I_k the setting

$$\beta_L^k = \begin{cases} \beta_L^c \text{ for } k = 0, N/2, N, \\ \beta_L^c + \ln(r)/(S_L^{k+1} - S_L^k) \text{ for } k = 1, \ldots, N/2 - 1, \\ \beta_L^c - \ln(r)/(S_L^{k+1} - S_L^k) \text{ for } k = N/2+1, \ldots, N-1, \end{cases} \tag{8}$$

and the recursion

$$\alpha_L^{k+1} = \alpha_L^k + (\beta_L^k - \beta_L^{k+1})S_L^{k+1}, \quad \alpha_L^0 = 0 \tag{9}$$

defines the multicanonical ensemble. In accordance with detailed balance, standard Metropolis and heat-bath updating algorithms have been generalized to the multicanonical situation [3,9]. Finally we obtain the canonical action density distribution $P_L(S)$ through a reweighting
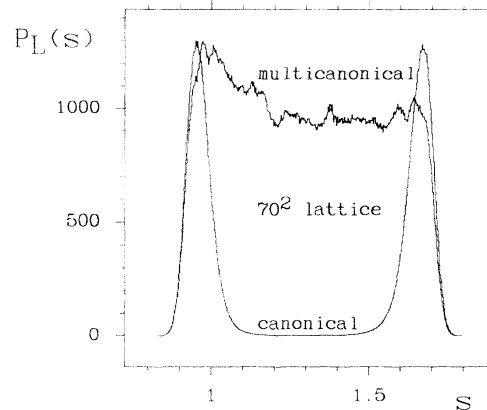

FIG. 2. Multicanonical action density distribution $\mathcal{P}_{70}(s)$ together with its reweighted distribution $P_{70}(s)$.

step similar to [10,11] from the multicanonical distribution $\mathcal{P}_L(S)$:

$$P_L(S) = e^{(\beta_L^c - \beta_L^k S - \alpha_L^k)}\mathcal{P}_L(S) \text{ for } S_L^k < S \le S_L^{k+1}. \tag{10}$$

As an example we show for our $L = 70$ system in Fig. 2 the multicanonical action density distribution $\mathcal{P}_{70}(S)$ together with the reweighted distribution $P_{70}(S)$. In practice the appropriate choice of the parameters in Eq. (7) is obtained by making from the given systems a FSS prediction of the density distribution $P_L(S)$ for the next larger system. A second run may then be performed with optimized parameters. It is our experience that the guess works normally so well that the second run is barely an improvement as compared with the first. On the smallest systems standard MC simulation provides initial data. Our statistics for this investigation was 4×10^6 heat-bath sweeps per run and lattice size. One sweep updates each spin of the lattice once.

We define the tunneling time τ'_L as the average number of sweeps needed to get from a configuration with action $S = S_L^{1,max}$ to a configuration with $S = S_L^{2,max}$ and back. With our statistics of 4×10^6 sweeps per run the system tunnels then in total $8 \times 10^6/\tau'_L$ times. Table I collects

TABLE I. The tunneling times τ'_L as a function of the lattice size L for the multicanonical MC algorithm (second row) and the heat-bath algorithm (third row). For some lattice sizes we display the results of several simulations, whose difference lies in slightly different coupling parameters.

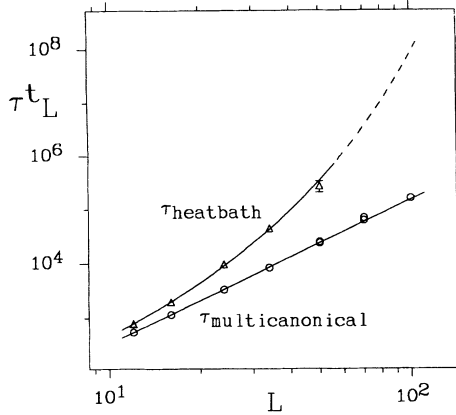| L	τ'_L (multicanonical)	τ'_L (heat bath)
12	542(4)	793(7)
12	$\cdots$	776(9)
16	1147(10)	1988(23)
24	3354(57)	9634(408)
34	8375(245)	43923(3151)
50	23763(1321)	270565(63222)
50	24932(1064)	$\cdots$
70	69492(6383)	$\cdots$
70	62218(5560)	$\cdots$
100	160334(16252)	$\cdots$

FIG. 3. Tunneling times for the multicanonical MC algorithm and the heat-bath algorithm in a double-log scale. The curves correspond to the fits in Eqs. (11) and (12). The dashed part of the curve indicates the extrapolation to the $L = 100$ lattice for the heat-bath algorithm. On the 100 lattice the system still tunnels 50 times between the metastable states during 4×10^6 sweeps, when the multicanonical simulation is used.

the measured tunneling times. For our smaller systems we have also carried out standard heat-bath MC runs at β_L^c and the associated tunneling times are also reported in Table I. For the larger systems standard MC runs would not tunnel often enough to allow for a reliable direct calculation of their tunneling times. This is of course due to the exponential slowing down of the standard MC simulation. In Fig. 3 we display on a log-to-log scale the divergence of the tunneling times τ_L^t for the multicanonical MC algorithm (circles) and the heat-bath algorithm (triangles). There is clearly a different behavior of the two algorithms involved. While for the multicanonical MC algorithm the increase of the tunneling time is consistent with a power law, the heat-bath algorithm displays an exponentially fast growing tunneling time. Performing a χ^2 fit we obtain the following fits:

$$\tau_L^t(\text{multicanonical}) = 0.73(3)L^{2.65(2)}$$

with $\chi^2/n_{\text{d.o.f.}} = 1.09$,   (11)

$$\tau_L^t(\text{heat bath}) = 1.46L^{2.15}e^{+0.080L}$$

with $\chi^2/n_{\text{d.o.f.}} = 2.2$.   (12)

The quality of the fits as indicated by the χ^2 values ($n_{\text{d.o.f.}}$ denotes the number of degrees of freedom) is reasonable. In case of the heat-bath algorithm we could not reliably determine the errors from a three-parameter fit. The ratio $R = \tau_L^t(\text{heat bath})/\tau_L^t(\text{multicanonical})$ is a direct measure for the relative efficiency of the two algorithms. Using the fits we extrapolate its value to the $L = 100$ system and estimate a factor $R \approx 500$ for this case. The multicanonical algorithm approximately slows down like $\sim V^{2.325}$ with respect to the number of updates per degree of freedom. This is only slightly worse than the optimal performance $\sim V^2$ which was estimated in [3] based on a random-walk picture. For the heat-bath algorithm the inefficiency of the algorithm prohibits a very accurate estimate of F^s from the behavior of the tunneling time according to Eq. (6). The fitted value in Eq. (11) is, however, close to the determination of the next paragraph [Eq. (14)].

Our multicanonical data allow the so far most precise determination of the interfacial free energy per unit area for the 2D 10-state Potts model. For this purpose we determine maxima and minima of the $P_L(S)$ distributions by self-consistent straight line fits over suitable S ranges. Together with the central values of their associated ranges, our F_L^s values are collected in Table II. Performing the FSS fit according to [8]

$$F_L^s = F^s + c/L,$$ (13)

we obtain consistent results for lattices of size $L = 16$, 24, 34, 50, 70, and 100, as displayed in Fig. 4 (we have $\chi^2/n_{\text{d.o.f.}} = 0.54$). We estimate the infinite volume value of the interfacial free energy per unit area to be

$$F^s = 0.09781 \pm 0.00075.$$ (14)

This value may, however, depend weakly on the analytical form of the FSS fit [8] and even with our large lattices we may still face additional systematic errors of a similar order as the quoted statistical error. Future simulations on even larger lattice sizes might therefore be of interest.

In summary, we have introduced a multicanonical ensemble for the numerical simulation of first-order phase

TABLE II. The pseudocritical couplings β_L^c, locations of the maxima and minima of the action density distribution and interfacial free energies, as determined from the multicanonical distributions.

| L	β_L^c	$S_L^{1,\text{max}}$	S_L^{min}	$S_L^{2,\text{max}}$	F_L^s
12	1.40738(09)	116	169	243	0.1071(06)
16	1.41534(12)	216	309	429	0.1086(07)
24	1.42100(08)	523	723	978	0.1058(08)
34	1.42338(09)	1072	1466	1945	0.1039(13)
50	1.42481(06)	2358	3162	4192	0.1027(11)
50	1.42469(06)	2357	3186	4190	0.1006(10)
70	1.42536(06)	4661	6257	8190	0.0983(20)
70	1.42541(05)	4660	6250	8178	0.1007(12)
100	1.42576(04)	9602	13060	16686	0.0986(18)
100	1.42577(04)	9577	13093	16711	0.0994(15)
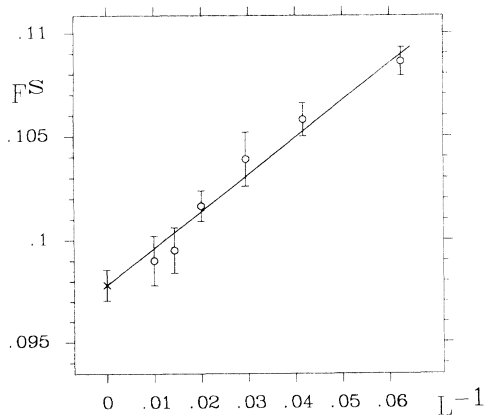
FIG. 4. FSS estimate of the interfacial free energy F^s. Averages are used for those lattices for which we have two data sets.

transitions, which eliminates an exponentially fast increase of the tunneling time between the ordered and disordered states in the critical region of the system. This finding is achieved by replacing the usual equilibrium dynamics of the canonical ensemble, through a new equilibrium dynamics, where the ordered and disordered states of the system get heated and cooled in a well controlled way. Thus configurations dominated by the presence of the interface are enhanced during the simulation.

The multicanonical MC algorithm gives a general framework for the numerical studies of first-order phase transitions in statistical mechanics as well as for field theoretic models. From the numerical point of view the interesting question will be what improvement factors can be achieved as compared with standard algorithms for certain models on certain lattice sizes. We expect the answer to this question to be determined by the value of the quantity $Q = F_L^s \times L^{d-1}$, where the strength of the first-order phase transition is indicated by the magnitude of F_L^s and the d is the dimensionality of the system. In the case of the 2D 10-state Potts model we find at values of $Q \sim 10$ approximately an improvement of 2 to 3 orders of magnitude, while at values of $Q \sim 1$ the improvement is marginal.

An implementation of the multicanonical MC algorithm for non-Abelian gauge theories is straightforward and we think that future investigations of the QCD deconfining phase transition will benefit from this. Beyond first-order phase transition, it may well be that multicanonical algorithms could be of use for other numerical calculations in statistical mechanics, such as estimates of the free energy or spin-glass simulations.

[a]On leave of absence from Department of Physics, The Florida State University, Tallahassee, FL 32306.
[1] R. H. Swendsen and J.-S. Wang, Phys. Rev. Lett. **58**, 86 (1987).
[2] R. H. Swendsen, J.-S. Wang, and A. M. Ferrenberg, "New Monte Carlo Methods for Improved Efficiency of Computer Simulations in Statistical Mechanics" (unpublished).
[3] B. Berg and T. Neuhaus, University of Bielefeld Report No. BI-TH 91/08 (to be published).
[4] R. B. Potts, Proc. Cambridge Philos. Soc. **48**, 106 (1952); F. Y. Wu, Rev. Mod. Phys. **54**, 235 (1982).
[5] A. Billoire, R. Lacaze, A. Morel, S. Gupta, A. Irbäck, and B. Peterson, Nucl. Phys. **B358**, 231 (1991).
[6] J. Lee and J. M. Kosterlitz, Phys. Rev. Lett. **65**, 137 (1990); Phys. Rev. B **43**, 3265 (1991).
[7] C. Borgs, R. Kotechý, and S. Miracle-Solé, J. Stat. Phys. **62**, 529 (1991); C. Borgs and W. Janke, Freie Universität Berlin Report No. HEP 6/91 (to be published).
[8] K. Binder, Phys. Rev. A **25**, 1699 (1982); Z. Phys. B **43**, 119 (1981).
[9] B. Baumann, Nucl. Phys. **B285**, 391 (1987).
[10] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **61**, 2635 (1988); **63**, 1658(E) (1989), and references given therein.
[11] B. Baumann and B. Berg, Phys. Lett. **164B**, 131 (1985).