ELSEVIER

Physica A 254 (1998) 164–178

PHYSICA A

Multicanonical Monte Carlo simulations

Wolfhard Janke *

Institut für Physik, Johannes Gutenberg-Universität Mainz, D-55099 Mainz, Germany

Abstract

Canonical Monte Carlo simulations of disordered systems like spin glasses and systems undergoing first-order phase transitions are severely hampered by rare event states which lead to exponentially diverging autocorrelation times with increasing system size and hence to exponentially large statistical errors. One possibility to overcome this problem is the multicanonical reweighting method. Using standard *local* update algorithms it could be demonstrated that the dependence of autocorrelation times on the system size V is well described by a less divergent power law, $\tau \propto V^{\alpha}$, with $1<\alpha<3$, depending on the system. After a brief review of the basic ideas, combinations of multicanonical reweighting with *non-local* update algorithms will be discussed. With the multibondic algorithm, which combines multicanonical reweighting with cluster updates, the dynamical exponent α can be reduced to unity, the optimal value one would expect from a random walk argument. Asymptotically for large system sizes the multibondic algorithm therefore always performs better than the standard multicanonical method. Finally it is shown that a combination with multigrid update techniques improves the performance of multicanonical simulations by roughly one order of magnitude, uniformly for all system sizes.

1. Introduction

Advances in statistical physics have increased dramatically our understanding of the behavior of a large variety of macroscopic systems. Over the years emphasis has shifted to more and more complex systems, such as disordered media, macromolecules, and glassy magnets and materials. Analytic progress towards deeper understanding of these systems has been limited. The scarcity of exact results and the unreliability of mean-field type approaches call for development of powerful numerical tools for disordered complex systems.

One approach is computer simulations based on importance sampling Monte Carlo methods [1–5]. The idea is to generate a Markov chain through phase space such

* Fax: +49 6131 395441; e-mail: janke@miro.physik.uni-mainz.de.

0378-4371/98/$19.00
PII S0378-4371(98)00014-4

that the frequency of sampled system states ϕ coincides with a given probability distribution. Most simulation studies work with the canonical ensemble where the microscopic states ϕ are distributed according to the Boltzmann distribution,

$$\mathscr{P}_{\mathrm{can}}(\phi) \propto \exp(-\beta H(\phi))\,, \tag{1}$$

with $\beta = 1/k_B T$ being the inverse temperature in natural units and H the Hamiltonian of the system. While this choice has a strong physical motivation, it does not necessarily yield the most efficient performance of the simulations.

That non-Boltzmann sampling would be, in principle, a legitimate alternative has been known since the early days of Monte Carlo simulations [2]. Its practical significance was first realized by Torrie and Valleau [6], who proposed the use of so-called "umbrella sampling" techniques. Most of the early applications aimed at a reliable computation of free energies which are unacceptably poorly estimated in canonical Boltzmann sampling. In the following years attention slowly shifted to the problem of rare event sampling and related problems with quasi-ergodicity [7]. Still, it took many years and needed the work on the multicanonical scheme [8,9] before the idea of non-Boltzmann sampling turned into a widely accepted practical tool in computer simulation studies of phase transitions and complex physical systems.

In the next section a few typical phenomena are described where canonical Boltzmann sampling is extremely inefficient. The idea of multicanonical Monte Carlo sampling and some applications are discussed in Section 3. In Section 4 the multibondic method is described which combines multicanonical reweighting with non-local cluster updates. In Section 5 another combination with non-local multigrid update schemes is briefly mentioned, and Section 6 contains a summary of the main results.

2. Problems of canonical sampling

In order to see the potential problems of canonical Boltzmann sampling it is useful to consider reduced distribution functions which depend only on a few macroscopic observables $\{Q_i\}$ like the energy, the magnetization, or the overlap parameter in spin glass simulations. Formally this can be derived by introducing delta-function constraints in the canonical partition function and summing over the microscopic degrees of freedom. This gives a reduced distribution $P_{\mathrm{can}}(\{Q_i\}) \propto \exp(-\beta F(\{Q_i\}))$, where $F(\{Q_i\})$ is a coarse grained free energy or effective potential of the relevant macroscopic degrees of freedom. If we choose $Q_i = E$ to be the energy of the system, then $\beta F(E) = \beta E - S(E)$, where $S(E) = \ln \Omega(E)$ is the (microcanonical) entropy and $\Omega(E)$ the density-of-states function, and $P_{\mathrm{can}}(E) \propto \exp(-\beta F(E)) = \Omega(E)\exp(-\beta E)$ is just the ordinary energy distribution. For complex physical systems these effective potentials often have a complicated rugged landscape with many minima and maxima which usually become more pronounced with increasing system size. While the minima correspond to the most probable regions in phase space, the maxima represent "rare event states" which are hardly sampled in the canonical Monte Carlo process.

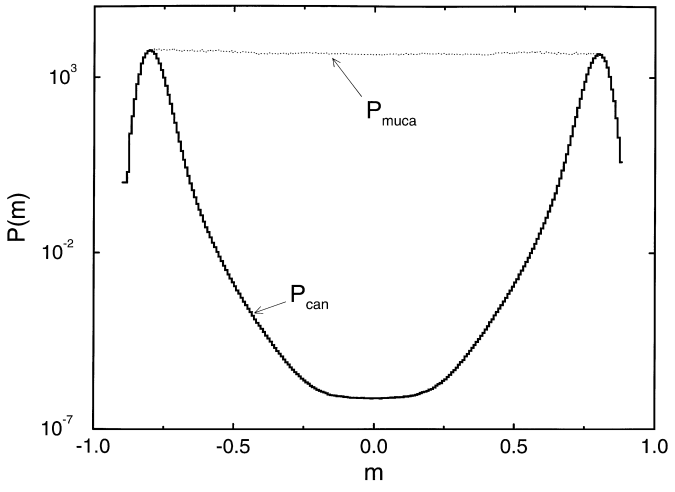

Fig. 1. Canonical and multicanonical probability distributions $P_{\rm can}(m)$ and $P_{\rm muca}(m)$ of the magnetization in the 2D ϕ^4 lattice model ($g=0.25,\ \mu^2=1.4,\ L=64$) on a logarithmic scale.

The simplest phenomenon illustrating the problems of canonical Boltzmann sampling are first-order phase transitions [10,11]. In large but finite systems mixed-phase configurations containing interfaces can occur which are identified as the rare event states. Compared with the pure phases their probability is suppressed by a factor $\propto \exp(-2\sigma L^{D-1})$, where σ is the (reduced) interface tension and L^{D-1} the cross-section of a D-dimensional system; see Fig. 1. The factor 2 reflects the fact that for the usually employed periodic boundary conditions at least two interfaces are present. In order to achieve equilibrium between the two (or more) probable regions in phase space the Monte Carlo process has to pass many times through these rare event states. This leads to autocorrelation times that are exponentially large in the system size, $\tau \propto \exp(2\sigma L^{D-1})$ – a phenomenon often termed *supercritical slowing down*. The autocorrelation times are easily visualized by looking at the temporal evolution of the observables. A plot of such a time series for the magnetization at the field-driven first-order phase transition in the 2D ϕ^4 lattice model is shown in Fig. 2a. We see that the systems stays for a long time in one of the two ordered phases (corresponding to the two peaks of $P_{\rm can}(m)$ shown in Fig. 1) and then suddenly jumps through the rare event region to the other side. The autocorrelation time τ is proportional to the average time spent in the pure phases during each visit.

Since this type of slowing down is directly linked with the shape of the canonical probability distribution, even sophisticated non-local update schemes such as cluster or multigrid techniques [3–5] cannot overcome this problem. One possible solution is to simulate auxiliary distributions whose shape is more regular and to reconstruct canonical expectation values at the end by reweighting. This is the basic idea underlying umbrella sampling [6] and the multicanonical scheme [8,9,12]. Related strategies, where one or more parameters are elevated to become dynamical degrees of freedom, are the

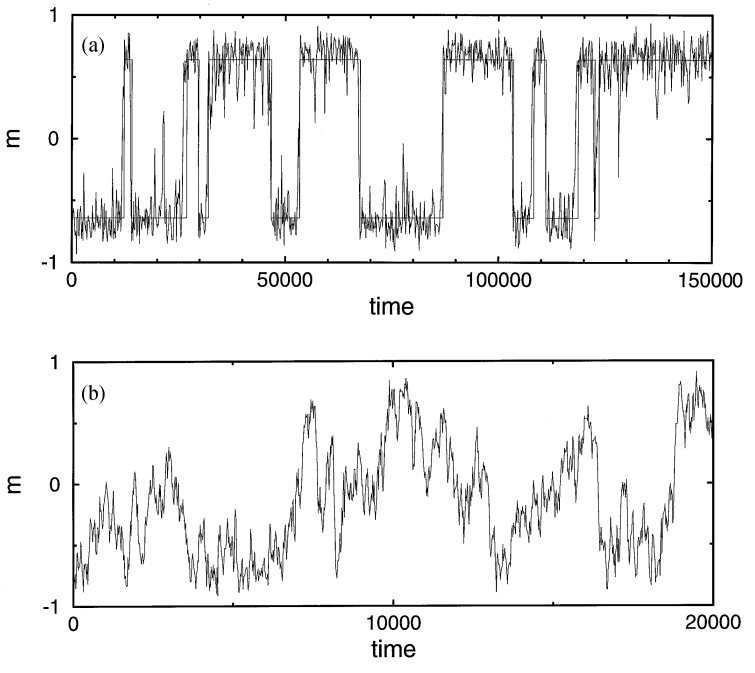

Fig. 2. Time evolution of magnetization measurements in (a) canonical and (b) multicanonical simulations of the 2D ϕ^4 lattice model.

expanded ensemble method [13], simulated [14] and parallel [15] tempering, and the dynamical-parameter method [16].

3. Multicanonical sampling

The multicanonical sampling method can be divided into two conceptually different approaches. The first is based on "enhancing the probability of rare event states", which is the typical strategy for dealing with first-order phase transitions. This allows the study of properties of the rare event states, for example interface tensions, which would be impossible with the expanded ensemble or tempering methods. The second approach can be best described by "avoiding rare events" which is closer in spirit to the alternative methods. In this variant one tries to connect the important parts of phase space by "easy paths" which go around the suppressed rare event regions.

In multicanonical simulations the canonical Boltzmann distribution (1) is replaced by

$$\mathscr{P}_{\mathrm{muca}}(\phi) \propto \exp(-\beta H(\phi) - f(\{Q_i(\phi)\}))\,, \tag{2}$$

where the reweighting factor $w(\{Q_i\}) \equiv \exp(-f(\{Q_i\}))$ is chosen in such a way that the probability distribution of the macroscopic variables $\{Q_i\}$ takes the desired form. Before discussing the choice of the variables $\{Q_i\}$ and the form of f, it should be

emphasized that, whatever these choices are, canonical expectation values of any observable $\mathcal{O}$ can always be recovered exactly by inverse reweighting,

$$\langle \mathcal{O} \rangle_{\mathrm{can}} = \langle \mathcal{O} \exp(f(\{Q_i\}))\rangle_{\mathrm{muca}} / \langle \exp(f(\{Q_i\}))\rangle_{\mathrm{muca}} \,. \tag{3}$$

The performance of the simulation, however, does depend crucially on the choice of $\{Q_i\}$ and the form of f – in the limiting case $f \equiv 0$ we are, of course, back to the troublesome canonical case.

The proper identification of the relevant set of Q_i's requires considerable physical intuition and insight into the specific system under study. While for disordered complex systems this is a serious problem, in studies of first-order phase transitions the proper choice is clear. At a temperature-driven transition the energy E is the relevant variable, and at a field-driven transition one should consider the magnetization M or order parameter Q. In the first case, $P_{\mathrm{can}}(E)$ exhibits a double-peak structure in the vicinity of the transition point, which becomes more and more pronounced with increasing system size. Here the weight function $f(E)$ in Eq. (2) is usually chosen such that the multicanonical distribution $P_{\mathrm{muca}}(E)$ of the energy is flat between the two peaks of the canonical distribution [17]. Similarly, in the so-called multi-magnetical variant [18,19] one aims at a flat magnetization distribution; see Fig. 1 for an example. At first sight it may appear natural to require that the macroscopic variables Q_i are uniformly sampled. The method is, however, by no means restricted to this choice, and it has in fact been argued that in certain applications non-uniform distributions are more appropriate [20].

An important technical point is the procedure for constructing the multicanonical weight factor [21–27]. For a uniform multicanonical distribution the formal solution is $\exp(-f(\{Q_i\})) = P_{\mathrm{can}}(\{Q_i\})^{-1}$. Of course, the probability distribution on the r.h.s. is not known at the beginning and one has to proceed by iteration. Starting with the canonical weight, or some initial guess based on results for already simulated smaller systems together with finite-size scaling extrapolations, one performs a short simulation to get an improved estimate of the canonical distribution. When this is inverted one obtains a new estimate of the multicanonical weight factor, which then is used in the next iteration and so on. In this naive version only the simulation data of the last iteration are used in the construction of the improved weight factor. A more sophisticated procedure, in which the new weight factor is computed from all available data accumulated so far, was proposed in Ref. [24] and recently improved in Ref. [27].

In principle one could construct non-canonical probability distributions for a whole set of macroscopic variables Q_i. In practice, however, this is rarely done. First, the construction of the weights becomes much more tedious. And second, also the performance of the production runs would slow down quite dramatically. For an idealized flat multicanonical distribution in one macroscopic variable a random walk argument suggests that $\tau \propto V^{\alpha}$, with $\alpha = 1$. The important point is, of course, that the exponential supercritical slowing down of the canonical formulation is replaced by a much weaker power law. But even a power-law scaling can be disastrous if the exponent α is not small. And this would inevitably be the case for a flat higher-dimensional distribution since a much bigger space would have to be scanned by the random walk.

Most multicanonical simulations are therefore performed with $Q_i = E = H$ or $Q_i = M$. The time evolution of magnetization measurements in a multicanonical (or, more precisely, multimagnetical) simulation is shown in Fig. 2b. Being a reweighting approach, the multicanonical method can, in principle, be combined with any legitimate update algorithm. In an actual implementation difficulties may arise, however, from the fact that due to the non-linear function $f(H)$ the effective multicanonical Hamiltonian

$$H_{\mathrm{muca}} = H + f(H)/\beta \tag{4}$$

is implicitly non-local. For other choices of Q_i the situation is similar. Most studies therefore employed *local* update algorithms (Metropolis or heat-bath) which are straightforward to adapt to H_{muca} [9,28]. The numerical estimates for τ scale indeed with the expected power law, $\tau \propto V^{\alpha}$, but with non-trivial exponents $\alpha > 1$ ($\alpha \approx 1.3$ for 2D Potts models) [12,17,29].

An important application of "enhancing rare events" in multicanonical simulations is the estimation of interface tensions, using the histogram method [30],

$$2\sigma_L = \frac{1}{L^{d-1}} \ln(P_{\mathrm{can}}^{\max}/P_{\mathrm{can}}^{\min}), \tag{5}$$

and finite-size scaling extrapolations in L. Even though this method was proposed already some time ago, it could never really be exploited in canonical simulations because of the statistical noise of the interface configurations. The first multicanonical studies concentrated on relatively simple systems like Ising and Potts models. While in the 2D Ising model the exactly known interface tension between the two ordered phases at low temperatures could easily be reproduced [18], the 3D studies revealed severe problems with finite-size scaling extrapolations [19] which remained unnoticed in previous canonical simulations of smaller systems. At the time of the simulations of the 2D 7-state [28] and 10-state [9] Potts models, the resulting estimates of the order–disorder interface tension could only be compared with alternative numerical methods. Since the differences turned out to be huge [31], it was gratifying that shortly after the numerical studies an exact formula could be derived [32] which is in nice agreement with the multicanonical results. With later refinements in the data analysis impressive agreement could be achieved [33–36]. Other applications of multicanonical simulations are studies of the coexistence curve in a Lennard–Jones fluid [37] and of the liquid–vapor asymmetry in pure fluids [38]. Furthermore, multicanonical simulations have also been applied to the first-order phase transitions in SU(3) lattice gauge theory [39,40] and in the electroweak standard model [41,42]. On the theoretical side, the improved accuracy allowed significant tests of finite-size scaling theories for first-order phase transitions [43–45].

The second strategy of "avoiding rare events" can be best explained for the example of a simple magnet at low temperatures, as described by, e.g., the Ising model. Instead of enhancing the probability of interface configurations in between the two ordered phases, one constructs multicanonical weights such that the energy distribution is flat between the lowest and highest possible energies. This allows the system to travel freely

back and force in energy, thereby choosing one or the other low-energy branch when coming back from the disordered phase. Since the ordered phases are now connected by the path through the disordered phase, their relative probabilities are properly sampled without the need of going through the highly suppressed interface configurations.

While this appears obvious for the simple Ising model, the properties of this second approach are much more involved in spin glass simulations where the phase structure at low temperatures is a priori unknown. For the dynamical behavior of this algorithm in applications to the Edwards–Anderson Ising spin glass in two and three dimensions values of $\alpha \approx 2.2$ and $\alpha \approx 2.8$, respectively, were reported in Refs. [21,46,47]. Another interesting application of this type of multicanonical simulations are studies of tertiary protein structures [48,49].

The remaining slowing down problem in multicanonical simulations is still severe. In fact, it is even worse than the critical slowing down of *local* Monte Carlo algorithms at a second-order phase transition. In this context major progress has been made by the development of *non-local* updates schemes such as cluster and multigrid algorithms [3–5]. It appeared therefore worthwhile to enquire if also the performance of multicanonical simulations can be further improved by combining the reweighting idea with one of the more sophisticated *non-local* update algorithms. In the next two sections some results for combinations with cluster algorithms and multigrid schemes are reported.

4. Multibondic sampling

In this section we confine the discussion to q-state Potts models with partition function

$$Z_{\mathrm{Potts}} = \sum_{\{\sigma_i\}} e^{-\beta E}; \qquad E = -\sum_{\langle ij \rangle} \delta_{\sigma_i \sigma_j}; \quad \sigma_i = 1, \ldots, q\,, \tag{6}$$

but generalizations to any model allowing some kind of (embedded) cluster representation should be straightforward.

For Potts models the construction of cluster update algorithms can be easily derived from the equivalent Fortuin–Kasteleyn representation [50,51],

$$Z_{\mathrm{Potts}} = \sum_{\{\sigma_i\}} \sum_{\{b_{ij}\}} \prod_{\langle ij \rangle} [\delta_{b_{ij},0} + p \delta_{\sigma_i \sigma_j} \delta_{b_{ij},1}] \tag{7}$$

$$= \sum_{\{b_{ij}\}} p^B q^{N_c}\,, \tag{8}$$

where $p = e^{\beta} - 1$. Bonds with $b_{ij} = 0$ or 1 are interpreted as "passive" or "active" bonds, respectively. Sites connected by active bonds belong to the same cluster. $B = \sum_{\langle ij \rangle} b_{ij}$ counts the number of active bonds and N_c is the number of clusters (including one-site clusters) for a given bond configuration.

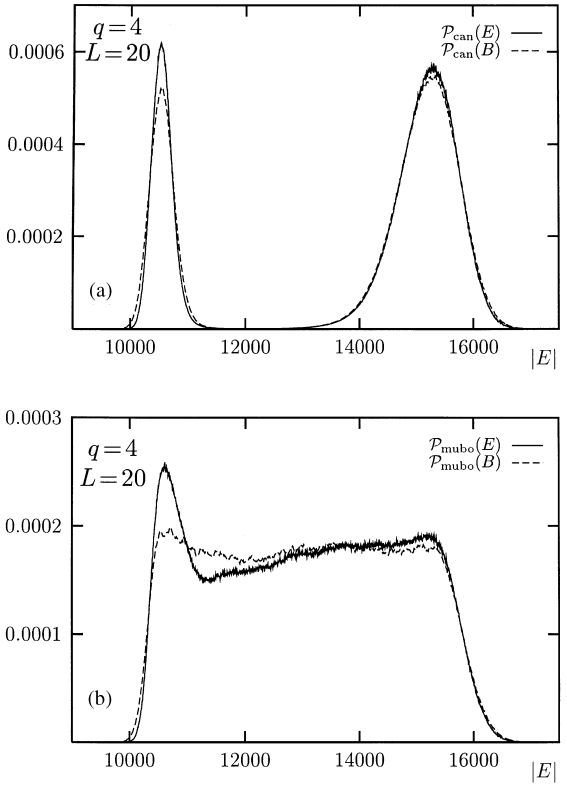

Fig. 3. (a) Canonical and (b) multibondic energy and bond histograms for the 3D 4-state Potts model at $\beta = 0.62857$ on a 20^3 lattice. The bond histograms are plotted vs $B(p+1)/p$ with $p = \exp(\beta) - 1$.

The crucial observation [52] is that the canonical bond distribution $P_{\rm can}(B)$ is very similar to the energy distribution $P_{\rm can}(E)$. For an illustration see Fig. 3. In general this is suggested by the identity

$$\langle E \rangle = -\frac{p+1}{p} \langle B \rangle \,, \tag{9}$$

and similar relations for the higher-order moments [52–54]. From Fig. 3a it is clear that the probabilities of the strongly suppressed interface configurations can also be enhanced by reweighting the bond distribution instead of the energy distribution. This suggested the introduction of a "multibondic" partition function [52]

$$Z_{\rm mubo} = \sum_{\{\sigma_i\}} \sum_{\{b_{ij}\}} \prod_{\langle ij \rangle} [p\delta_{\sigma_i\sigma_j}\delta_{b_{ij},1} + \delta_{b_{ij},0}] \exp(-f(B)) \,, \tag{10}$$

where the choice $\exp(-f(B)) = P_{\rm can}(B)^{-1}$ would assure that $P_{\rm mubo}(B) = {\rm const}$. The construction of $f(B)$ proceeds similarly to the multicanonical case and again any reasonable approximation of $P_{\rm can}(B)$ can be used in practice. As is shown in Fig. 3b, once

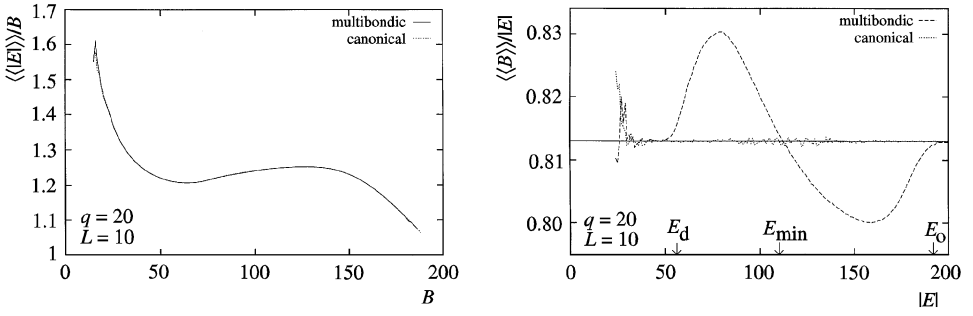

Fig. 4. Multibondic and canonical data for the scaled "microbondic" (fixed B) expectation values of the energy $\langle\langle|E|\rangle\rangle(B)/B$ (left plot) and the scaled "microcanonical" (fixed E) expectation values of the total bond number $\langle\langle B\rangle\rangle(E)/|E|$ (right plot) in the 2D 20-state Potts model on a 10^2 lattice at $\beta=\beta_t=\ln(1+\sqrt{20})$. The horizontal line on the r.h.s. is the exact canonical result $p/(p+1)$ with $p=\exp(\beta)-1$.

the multibondic bond distribution $P_{\mathrm{mubo}}(B)$ is approximately flat between the two peaks of the canonical distribution this is also the case for the energy distribution $P_{\mathrm{mubo}}(E)$.

The multibondic reweighting factor affects only the first step of the standard Swendsen–Wang cluster update algorithm [55]:

(1) If $\sigma_i\neq\sigma_j$, set $b_{ij}=0$ as usual. If $\sigma_i=\sigma_j$, then assign new values $b_{ij}^{\mathrm{new}}=0$ and 1 with relative probabilities $\exp(-f(B')):p\exp(-f(B'+1))$, where $B'=B-b_{ij}^{\mathrm{old}}$.

(2) Identify clusters of spins that are connected by "active" bonds ($b_{ij}=1$).

(3) Draw a random value $1,\ldots,q$ independently for each cluster and assign this value to all spins in a cluster.

The standard cluster algorithm is obviously recovered by setting $f\equiv 0$ in the first step.

The properties of the multibondic cluster update can be visualized by measuring the conditional "microbondic" (fixed B) and "microcanonical" (fixed E) probability distributions $P_B(E)$ and $P_E(B)$, respectively. Since for a given bond configuration with bond number B the update of the spins proceeds in precisely the same way as in the canonical cluster algorithm, $P_B(E)$ should look the same in multibondic and canonical simulations. This is demonstrated on the l.h.s. of Fig. 4 where the first "microbondic" moment $\langle\langle|E|\rangle\rangle(B)/B$ is shown for both types of simulations. A plot of the variance $\langle\langle(E-\langle\langle E\rangle\rangle)^2\rangle\rangle/B$ looks qualitatively similar.

The influence of the reweighting factor in the first step of the multibondic update should show up, of course, when considering the "microcanonical" distribution $P_E(B)$. This can be clearly seen on the r.h.s. of Fig. 4 which displays the first "microcanonical" moment $\langle\langle B\rangle\rangle(E)/|E|$, as obtained in multibondic and canonical simulations. The canonical data follow the horizontal line. This is an exact result which can be derived as follows. Starting from an arbitrary spin configuration with energy E, active bonds can only appear between neighbors with equal spin values. Out of these candidate bonds we decide bond by bond with probability $1:p$ whether this bond is actived or not. Therefore the total number of active bonds, $B\leqslant|E|$, is distributed according to a

binomial probability distribution,

$$\mu_E(B) = \binom{|E|}{B} \left(\frac{1}{p+1}\right)^{|E|-B} \left(\frac{p}{p+1}\right)^{B}, \tag{11}$$

and, on the average,

$$\langle\langle B \rangle\rangle(E) = \sum_B B\mu_E(B) = \frac{p}{p+1}|E| \tag{12}$$

out of the $|E|$ candidate bonds will actually be activated.

Recalling step 1 of the multibondic update prescription it is now easy to see that in this case the probability distribution (11) is modified to

$$\tilde{\mu}_E(B) = \frac{1}{\mathcal{N}} \binom{|E|}{B} \left(\frac{1}{p+1}\right)^{|E|-B} \left(\frac{p}{p+1}\right)^{B} \exp(-f(B)), \tag{13}$$

where $\mathcal{N}$ is a normalization factor. This suggests an alternative update procedure. Given a spin configuration with energy E, one could determine a new bond configuration $\{b_{ij}\}$ also by first drawing the total number of active bonds B ($<|E|$) directly from the distribution (13), and then randomly activating B out of the $|E|$ candidate bonds with $\delta_{\sigma_i\sigma_j} = 1$.

In Ref. [52] we compared the performance of multibondic (mubo) and multicanonical (muca) simulations for the two-dimensional Potts model (6) with $q = 7$, 10, and 20. A quantity that allows direct comparison with previous work is the flipping time $\tau_E^{\text{flip}} = N_{\text{up}}/4$, with N_{up} denoting the average number of update sweeps for a round-trip between $E < E_{\min}$ and $E > E_{\max}$, where $E_{\min,\max}$ are the peak locations of the canonical energy distribution. The results for $q = 7$ are shown in Fig. 5. The multibondic algorithm clearly outperforms the multicanonical simulations and is slightly better than the "demon" algorithm [33] which also uses cluster updates as an important ingredient. From least-squares fits to the ansatz $\tau_E^{\text{flip}} = aV^{\alpha}$ we obtained $\alpha = 1.27(2) \approx 1.3$ for multicanonical heat-bath and $\alpha = 0.92(2)$ for multibondic cluster simulations. For $q = 10$ the exponents are similar, $\alpha = 1.32(2) \approx 1.3$ (muca) and $\alpha = 1.05(1)$ (mubo), but due to a larger prefactor for the multibondic algorithm we can take advantage of the asymptotic improvement only for large lattice sizes $L > 50$ [52,56]. For $q = 20$ the standard algorithm is clearly favored for all reasonable lattice sizes. For more details see Ref. [53].

The same systematics was recently observed [57] at the first-order phase transition of the three-dimensional Potts model (6) with $q = 4$ and 5; see Fig. 5. Here the exponents for the multicanonical simulations are $\alpha = 1.14(2)$ ($q = 4$) and 1.11(1) ($q = 5$), and for the multibondic algorithm we again obtained smaller values consistent with unity, $\alpha = 0.96(1)$ ($q = 4$) and 0.98(1) ($q = 5$).

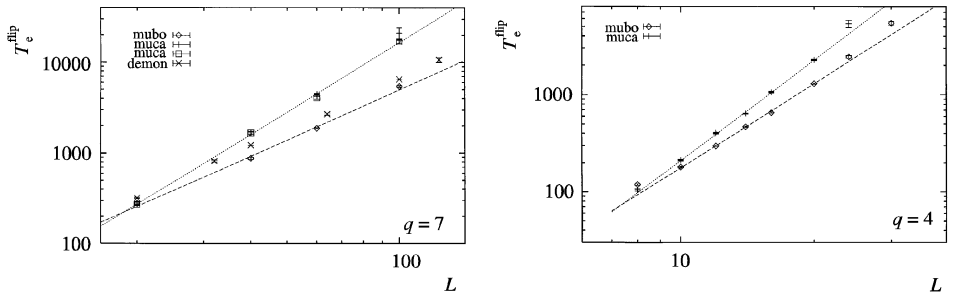

Fig. 5. Autocorrelation times in multibondic (mubo) cluster and multicanonical (muca) heat-bath simulations at the first-order transition point of the 2D 7-state (left plot) and 3D 4-state (right plot) Potts model. In 2D also some previous results are shown for comparison: "muca □" (Ref. [28]), "demon ×" (Ref. [33]).

5. Multicanonical multigrid sampling

Another non-local update scheme which can be quite easily combined with multicanonical reweighting is the multigrid method [3,58,59]. The basic idea of multigrid methods is to perform updates on different length scales. In the *uni*grid formulation one works with the original lattice and proposes collective moves for blocks of 1, $2^d, 4^d, 8^d, \ldots, 2^{nd} = V$ neighboring field variables which, similar to the Metropolis algorithm, are accepted or rejected according to the energy change. The simplest choice are piecewise constant collective excitations (blocks), but in principle also other shapes could be considered (e.g., pyramids). Another important ingredient is the sequence in which the various block sizes are selected in a cycle, the analog of a sweep.

In the mathematically equivalent, recursive *multi*grid formulation the blocks of size $2^{(n-k)d}$ are represented by auxiliary variables $\phi_i^{(k)}$ on coarse-grained lattices – the multigrids $\Xi^{(k)}$ of size 2^{kd}, $k = n, \ldots, 0$. The shape of the collective moves is controlled by an operator $\mathscr{P}$ which interpolates $\phi_i^{(k)}$ back to the next finer grid $\Xi^{(k+1)}$, and the acceptance of the proposed moves is governed by a coarse-grid Hamiltonian, $H^{(k)}(\phi_i^{(k)}) = H^{(k+1)}(\phi_i^{(k+1)} + \mathscr{P}\phi_i^{(k)})$, $H^{(n)} = H$, which is defined recursively by freezing the field variables $\phi_i^{(k+1)}$. The update on level $\Xi^{(k)}$ thus consist of:

(1) n_1 presweeps using any local update scheme with Hamiltonian $H^{(k)}$.

(2) Calculating the Hamiltonian for the next coarser grid $\Xi^{(k-1)}$ (whose parameters depend on the current configuration on grid $\Xi^{(k)}$) and initializing the variables on grid $\Xi^{(k-1)}$ to zero.

(3) Updating the variables $\phi_i^{(k-1)}$ with the multigrid scheme γ_{k-1} times.

(4) Interpolating the variables of grid $\Xi^{(k-1)}$ back to grid $\Xi^{(k)}$.

(5) n_2 postsweeps using the local update algorithm with Hamiltonian $H^{(k)}$.

On the coarsest grid $\Xi^{(0)}$, of course, one only performs steps 1 and 5. The recursive multigrid formulation leads automatically to a more efficient implementation if the functional form of the Hamiltonian remains invariant under the coarsening prescription (similar to FFT) and, due to its recursive structure, makes the definition of level sequences determined by the parameters γ_k more transparent. For $\gamma_k \equiv 1$ (V-cycle)

every grid is given the same weight, while for $\gamma_k \equiv 2$ (W-cycle) the coarser grids are updated more frequently. The cycle names derive from their graphical representation [29] which resemble the letters V and W.

From the unigrid viewpoint it is immediately clear that multigrid methods can also be used to update the multicanonical distribution [60–64]. Let us first consider the multimagnetical variant. Since on level k we propose a move of $2^{(n-k)d}$ variables in conjunction, this update proposal would change the magnetization M by an amount of $2^{(n-k)d}\Delta\phi_{i_0}^{(k)}$. The only modification for the update on level k will therefore be to compute the energy difference according to

$$\beta\Delta E_{\mathrm{muca}}^{(k)} = \beta\Delta E^{(k)} + f(M + 2^{(n-k)d}\Delta\phi_{i_0}^{(k)}) - f(M)\,, \tag{14}$$

where $\Delta E^{(k)}$ is the energy difference computed with the coarse-grid Hamiltonian $H^{(k)}$ as in the usual canonical multigrid formulation. It should be emphasized that the modifications in a recursive multigrid implementation are precisely the same. For multicanonical reweighting in the energy, M would have to be replaced by E and $2^{(n-k)d}\Delta\phi_{i_0}^{(k)}$ by $\Delta E^{(k)}$, which is also straightforward to implement in both the unigrid and multigrid formulation. For a discussion of more general situations, see Ref. [63].

In order to evaluate the performance of the multicanonical multigrid algorithm we studied the field-driven first-order phase transitions in the two-dimensional scalar ϕ^4 lattice model with Hamiltonian

$$H = \sum_{i=1}^{V}\left[\frac{1}{2}\sum_{\kappa=1}^{2}(\phi_{i_\kappa} - \phi_i)^2 - \frac{\mu^2}{2}\phi_i^2 + \frac{1}{4}\phi_i^4\right], \tag{15}$$

and compared the autocorrelation times of the magnetization in multicanonical simulations with Metropolis and multigrid W-cycle updates [63]. Least-squares fits to the power-law ansatz $\tau_m^{\mathrm{eff}} \propto V^{\alpha}$ gave for both update algorithms exponents of $\alpha \approx 1.2$, 1.4, and 1.5 for $\mu^2 = 1.30$, 1.35, and 1.40, respectively, i.e., the multigrid update does not improve the asymptotic behavior. The autocorrelation times of the W-cycle, however, were found to be about 20 times smaller than those of the Metropolis algorithm. If one finally takes into account that a multigrid W-cycle requires more elementary operations than a Metropolis sweep, one obtains a *real time* improvement factor of about 10. Further applications of this algorithm to quantum mechanical tunneling problems are discussed in Refs. [60,62], and more technical details can be found in Ref. [64].

6. Conclusions

In the past few years multicanonical simulations have proven to be a promising tool for dealing with the numerical problems inherent in simulations of complex physical systems. Due to its very general formulation it has been applied by now to many problems in statistical physics covering such diverse fields as first-order phase transitions in condensed matter and high-energy physics, the low-temperature behavior of disordered systems like spin glasses, and the protein folding problem.

Combinations of multicanonical reweighting with non-local update schemes such as cluster or multigrid algorithms are feasible and lead to a further improvement of the performance. The combination with cluster updates is optimal in the sense that for q-state Potts models the exponent α in the power law, $\tau = aV^{\alpha}$, is consistent with the optimal random walk estimate $\alpha = 1$. Asymptotically for large system sizes the multibondic algorithm is therefore always superior over standard multicanonical simulations using local update algorithms where $\alpha > 1$. The prefactor a, however, grows rapidly with the number of Potts states q, such that for reasonable lattice sizes multibondic simulations turned out to be more efficient only for small q ($\leqslant 10 \ldots 12$ in 2D, and $\leqslant 5 \ldots 6$ in 3D). For the combination with multigrid update techniques a real-time improvement of about one order of magnitude was observed for all considered lattice sizes in simulations of the two-dimensional ϕ^4 lattice model.

Acknowledgements

Most of the material discussed in this talk is based on joint work with Bernd Berg, Malcolm Carroll, Stefan Kappler, Mohammad Katoot, and Tilman Sauer. I wish to thank all of them for fruitful and enjoyable collaborations, and the Deutsche Forschungsgemeinschaft for a Heisenberg Fellowship.

References

[1] J.M. Hammersley, D.C. Handscomb, Monte Carlo Methods, Chapman and Hall, London, 1964.
[2] K. Binder, in: C. Domb, M.S. Green (Eds.), Phase Transitions and Critical Phenomena, vol. 5b, Academic Press, New York, 1976, p.1.
[3] A.D. Sokal, Monte Carlo Methods in Statistical Mechanics: Foundations and New Algorithms, Cours de Troisième Cycle de la Physique en Suisse Romande, Lausanne, 1989; and Bosonic Algorithms, in: M. Creutz (Ed.), Quantum Fields on the Computer, World Scientific, Singapore, 1992, p. 211.
[4] W. Janke, Monte Carlo simulations of spin systems, in: K.H. Hoffmann, M. Schreiber (Eds.), Computational Physics: Selected Methods – Simple Exercises – Serious Applications, Springer, Berlin, 1996, p. 10.
[5] W. Janke, Nonlocal Monte Carlo Algorithms for Statistical Physics Applications, Mainz preprint (April 1997), to appear in: Monte Carlo Methods, Proc. IMACS Workshop, Brussels, April 1997.
[6] G.M. Torrie, J.P. Valleau, Chem. Phys. Lett. 28 (1974) 578; J. Comp. Phys. 23 (1977) 187; J. Chem. Phys. 66 (1977) 1402; I.S. Graham, J.P. Valleau, J. Phys. Chem. 94 (1990) 7894; J.P. Valleau, J. Comp. Phys. 96 (1991) 193.
[7] D. Chandler, Introduction to Modern Statistical Mechanis, Oxford University Press, Oxford, 1987, pp. 168–175.
[8] B.A. Berg, T. Neuhaus, Phys. Lett. B 267 (1991) 249.
[9] B.A. Berg, T. Neuhaus, Phys. Rev. Lett. 68 (1992) 9.
[10] K. Binder, Rep. Prog. Phys. 50 (1987) 783.
[11] H.J. Herrmann, W. Janke, F. Karsch (Eds.), Dynamics of First Order Phase Transitions, World Scientific, Singapore, 1992.
[12] W. Janke, in: Ref. [11], p. 365 [reprinted in Int. J. Mod. Phys. C 3 (1992) 1137].
[13] A.P. Lyubartsev, A.A. Martsinovski, S.V. Shevkunov, P.N. Vorontsov-Velyaminov, J. Chem. Phys. 96 (1992) 1776.
[14] E. Marinari, G. Parisi, Europhys. Lett. 19 (1992) 451.
[15] K. Hukusima, K. Nemoto, J. Phys. Soc. Japan 65 (1996) 1604.

[16] W. Kerler, A. Weber, Phys. Rev. B 47 (1993) R11563; W. Kerler, C. Rebbi, A. Weber, Phys. Rev. D 50 (1994) 6984; Nucl. Phys. B 450 (1995) 452.
[17] B.A. Berg, in: Ref. [11], p. 311 [reprinted in Int. J. Mod. Phys. C 3 (1992) 1083]; and in: F. Karsch, B. Monien, H. Satz (Eds.), Multiscale Phenomena and Their Simulation, World Scientific, Singapore, 1997, p. 137.
[18] B.A. Berg, U. Hansmann, T. Neuhaus, Phys. Rev. B 47 (1993) 497.
[19] B.A. Berg, U. Hansmann, T. Neuhaus, Z. Phys. B 90 (1993) 229.
[20] B. Hesselbo, R. Stinchcombe, Phys. Rev. Lett. 74 (1995) 2151.
[21] B.A. Berg, T. Celik, Phys. Rev. Lett. 69 (1992) 2292.
[22] J. Lee, Phys. Rev. Lett. 71 (1993) 211, 2353 (erratum).
[23] B.A. Berg, U. Hansmann, Y. Okamoto, J. Phys. Chem. 99 (1995) 2236.
[24] B.A. Berg, J. Stat. Phys. 82 (1996) 323.
[25] G.R. Smith, A.D. Bruce, Phys. Rev. E 53 (1996) 6530; J. Phys. A 28 (1995) 6623.
[26] M.J. Thill, preprints cond-mat/9703234 and cond-mat/9703233.
[27] B.A. Berg, W. Janke, in preparation.
[28] W. Janke, B.A. Berg, M. Katoot, Nucl. Phys. B 382 (1992) 649.
[29] W. Janke, in: D.P. Landau, K.K. Mon, H.B. Schüttler (Eds.), Computer Simulations in Condensed Matter Physics VII, Springer, Heidelberg, 1994, p. 29.
[30] K. Binder, Phys. Rev. A 25 (1982) 1699.
[31] W. Janke, in: R.A. de Groot, J. Nadrchal (Eds.), Physics Computing '92, World Scientific, Singapore, 1993, p. 351.
[32] C. Borgs, W. Janke, J. Phys. I France 2 (1992) 2011.
[33] K. Rummukainen, Nucl. Phys. B 390 (1993) 621.
[34] B. Grossmann, S. Gupta, Phys. Lett. B 319 (1993) 215.
[35] A. Billoire, T. Neuhaus, B. Berg, Nucl. Phys. B 413 (1994) 795.
[36] W. Janke, Nucl. Phys. (Proc. Suppl.) B 63 (1998) 631.
[37] N.B. Wilding, Phys. Rev. E 52 (1995) 602.
[38] N.B. Wilding, M. Müller, J. Chem. Phys. 102 (1995) 2562.
[39] B. Grossmann, M.L. Laursen, T. Trappenberg, U.-J. Wiese, Phys. Lett. B 293 (1992) 175; B. Grossmann, M.L. Laursen, Nucl. Phys. B 408 (1993) 637.
[40] Y. Iwasaki, K. Kanaya, L. Kärkkäinen, K. Rummukainen, T. Yoshié, Phys. Rev. D 47 (1993) 3079.
[41] F. Csikor, Z. Fodor, J. Hein, K. Jansen, A. Jaster, I. Montvay, Nucl. Phys. B 439 (1995) 147.
[42] F. Karsch, T. Neuhaus, A. Patkós, Nucl. Phys. B 441 (1995) 629.
[43] C. Borgs, S. Kappler, Phys. Lett. A 171 (1992) 37.
[44] A. Billoire, T. Neuhaus, B. Berg, Nucl. Phys. B 396 (1993) 779.
[45] M.S. Carroll, S. Kappler, W. Janke, K. Binder, in preparation.
[46] B.A. Berg, T. Celik, Int. J. Mod. Phys. C 3 (1992) 1251.
[47] B.A. Berg, U. Hansmann, T. Celik, Phys. Rev. B 50 (1994) 1644.
[48] U.H.E. Hansmann, Y. Okamoto, J. Comp. Chem. 14 (1993) 1333.
[49] Y. Okamoto, U.H.E. Hansmann, J. Phys. Chem. 99 (1995) 11 276.
[50] C.M. Fortuin, P.W. Kasteleyn, Physica 57 (1972) 536.
[51] C.M. Fortuin, Physica 58 (1972) 393; 59 (1972) 545.
[52] W. Janke, S. Kappler, Phys. Rev. Lett. 74 (1995) 212.
[53] S. Kappler, Ph.D. Thesis, Johannes Gutenberg-Universität Mainz, 1995.
[54] W. Janke, S. Kappler, in preparation.
[55] R.H. Swendsen, J.-S. Wang, Phys. Rev. Lett. 58 (1987) 86.
[56] W. Janke, in: F. Karsch, B. Monien, H. Satz (Eds.), Multiscale Phenomena and Their Simulation, World Scientific, Singapore, 1997, p. 147.
[57] M.S. Carroll, W. Janke, S. Kappler, Mainz preprint (June 1997), J. Stat. Phys. 90 (1998), in press.
[58] H. Meyer-Ortmanns, Z. Phys. C 27 (1985) 553; J. Goodman, A.D. Sokal, Phys. Rev. Lett. 56 (1986) 1015; Phys. Rev. D 40 (1989) 2035; G. Mack, in: G. 't Hooft et al. (Eds.), Nonperturbative Quantum Field Theory, Cargèse lectures 1987, Plenum, New York, 1988, p. 309; G. Mack, S. Meyer, Nucl. Phys. B (Proc. Suppl.) 17 (1990) 293; D. Kandel, E. Domany, D. Ron, A. Brandt, E. Loh, Jr., Phys. Rev. Lett. 60 (1988) 1591; D. Kandel, E. Domany, A. Brandt, Phys. Rev. B 40 (1989) 330.
[59] W. Janke, T. Sauer, Chem. Phys. Lett. 201 (1993) 499.

[60] W. Janke, T. Sauer, in: H. Grabert, A. Inomata, L.S. Schulman, U. Weiss (Eds.), Path Integrals from meV to MeV, World Scientific, Singapore, 1993, p. 17.
[61] W. Janke, T. Sauer, Nucl. Phys. B (Proc. Suppl.) 34 (1994) 771.
[62] W. Janke, T. Sauer, Phys. Rev. E 49 (1994) 3475.
[63] W. Janke, T. Sauer, J. Stat. Phys. 78 (1995) 759.
[64] T. Sauer, Ph.D. Thesis, Freie Universität Berlin, 1994.