Available online at www.sciencedirect.com
ScienceDirect
Physics Procedia 7 (2010) 19–28
Physics Procedia
www.elsevier.com/locate/procedia

Wang-Landau Multibondic Cluster Approach to Simulations of Second-Order Transitions

Bernd A. Berg[a,b], Wolfhard Janke[c,d]

[a]*Department of Physics, Florida State University, Tallahassee, FL 32306, USA*
[b]*School of Computational Science, Florida State University, Tallahassee, FL 32306, USA*
[c]*Institut für Theoretische Physik, Universität Leipzig, Augustusplatz 10/11, 04109 Leipzig, Germany*
[d]*Centre for Theoretical Sciences (NTZ), Universität Leipzig, Emil-Fuchs-Straße 1, 04105 Leipzig, Germany*

Abstract

At second-order phase transitions the critical energy range covered by a canonical Monte Carlo simulation close to the critical temperature is often smaller than the energy range needed for reliable reweighting analyses of certain observables. Such an extended energy range can be covered by performing a Wang-Landau recursion for the spectral density followed by a multicanonical simulation with fixed weights. But in the conventional approach based on local update rules one loses the advantage due to non-local cluster algorithms which are well known to drastically reduce critical slowing down. We develop a cluster version of the Wang-Landau recursion together with a subsequent multibondic simulation and show for 2D and 3D Ising models that the efficiency of the conventional Wang-Landau multicanonical approach can be improved by power laws in the lattice size. In our simulations real gains in CPU time reach two orders of magnitude.

1. Introduction

Equilibrium properties of statistical physics systems are often estimated by Markov chain Monte Carlo (MCMC) simulations [1]. In many cases one is interested in calculating expectation values for a range of temperatures with respect to the Gibbs canonical ensemble. It has turned out that instead of performing simulations of the canonical ensemble at distinct temperatures it is often advantageous to combine them into one simulation of a "generalized" ensemble [2, 3, 4, 5]; for reviews see [6, 7, 8].

While the power of generalized ensembles is well documented for first-order phase transitions and complex systems such as spin glasses and peptides (small proteins), this is not the case

Email addresses: berg@scs.fsu.edu (Bernd A. Berg), wolfhard.janke@itp.uni-leipzig.de (Wolfhard Janke)

1875-3892
doi:10.1016/j.phpro.2010.09.040

for second-order phase transitions. Here we give an extended version of our recent work [9] on this subject.

In MCMC simulations of second-order phase transitions one wants to cover the scaling region in which many physical observables diverge with increasing lattice size. So we ask the question: How large is the energy range of this region on a finite lattice and is it eventually already covered by a single canonical simulations at the (infinite volume) critical temperature $T_c = 1/\beta_c$? This question will be addressed and answered in section 2.

Previously the convenience of applications of generalized ensembles to second-order phase transitions was claimed by Landau and collaborators [10]. However, in their approach they lost the crucial advantage which cluster algorithms [11, 12] provide. In section 3 we present a generalization to cluster algorithms, starting from the multibondic (MUBO) [13] cluster version of the multicanonical (MUCA) [3] ensemble.

To keep these lecture notes simple, we restrict our investigations to 2D and 3D Ising models, while the points made are generally valid for cluster algorithms. These numerical illustrations are compiled in section 4. A short summary and some conclusions are given in the final section 5.

2. Finite-Size Scaling at Second-Order Transitions

We consider a second-order phase transition and an observable S, so that its expectation value $\widehat{S} = \langle S \rangle$ diverges near the critical point on an infinite lattice according to (for simplicity we consider only symmetric behavior in $\pm t$)

$$\widehat{S}(\beta) \sim |t|^{-\sigma} \quad \text{or} \quad \widehat{S}(\beta) \sim -\ln|t|\,, \tag{1}$$

where the reduced temperature t is defined by

$$t = \frac{T - T_c}{T_c} = \frac{\beta_c - \beta}{\beta} \tag{2}$$

and σ is the critical exponent of S. We first consider the case $\sigma > 0$ and afterwards the case of a logarithmic divergence.

We place the system on finite lattices of shape L^D with periodic boundary conditions, so that $\widehat{S}_L(\beta)$ can be calculated by MCMC simulations. Finite-size scaling (FSS) theory, for a review see [14], implies

$$\widehat{S}_L(\beta_c) \sim L^{\sigma/\nu}\,. \tag{3}$$

The standard argument [15] is that in the scaling region $t \to 0$

$$\widehat{S}_L(\beta) \sim |t|^{-\sigma} f(L/\xi) \tag{4}$$

holds, where $\xi \sim |t|^{-\nu}$ is the relevant correlation length of the system, and f a scaling function, which depends only on the ratio of the two relevant length scales. A second-order phase transition requires $\nu > 0$. As $\widehat{S}_L(\beta_c)$ is finite, $f(L/\xi)$ has to eliminate the singularity of $|t|^{-\sigma}$, implying $f(|t|^\nu L) \sim (|t|^\nu L)^{\sigma/\nu}$ for $t \to 0$ and, therefore, Eq. (3).

We denote the probability density of the energy from a canonical MCMC simulation by $P(E)$. The expectation value of the specific heat per site C scales at the critical inverse temperature as $\widehat{C}_L \sim L^{\alpha/\nu}$, so that the fluctuation-dissipation theorem gives

$$\left\langle (E - \widehat{E})^2 \right\rangle \sim L^{D+\alpha/\nu}\,. \tag{5}$$

This implies for the range covered by the simulation at β_c,

$$\triangle E = |E_{0.75} - E_{0.25}| \sim L^{D/2+\alpha/(2\nu)}\,, \tag{6}$$

where E_q, $q = 0.25$ and $q = 0.75$, are fractiles (see, e.g., [7]) of the energy distribution, i.e., if $P(E)$ is normalized to unity, then $\sum_{E=E_{\min}}^{E_q} P(E) = q$, where $E_{\min}$ is the smallest energy of the system. The choice $|E_{0.75} - E_{0.25}|$ thus means that 50% of all energy measurements fall into this range. In the vicinity of β_c a Taylor expansion gives

$$\widehat{E}(\beta)/L^D = \widehat{E}(\beta_c)/L^D + A\,|\beta - \beta_c|^{1-\alpha} = \widehat{E}(\beta_c)/L^D + A\,(\triangle\beta)^{1-\alpha}\,, \tag{7}$$

where A is a constant. Using the hyperscaling relation [14] $\alpha = 2 - D\nu$, Eqs. (6) and (7) translate into

$$\triangle E/L^D = L^{-D/2+\alpha/(2\nu)} = L^{-D+1/\nu} \;\sim\; (\triangle\beta)^{1-\alpha} = (\triangle\beta)^{-1+D\nu}\,, \tag{8}$$

which implies the *canonical* reweighting range

$$\triangle\beta \sim L^{-1/\nu}\,. \tag{9}$$

The *desired* reweighting range, on the other hand, is determined by the need to cover the maxima of all divergent observables measured. Let the maximum value of our observable $\widehat{S}_L(\beta)$ be $\widehat{S}_L^{\max} = \widehat{S}_L(\beta_L^{\max})$. In the infinite-volume limit the maximum is located at β_c, so it cannot run away from $\widehat{S}_L(\beta_c)$ and FSS theory implies also

$$\widehat{S}_L^{\max} \sim L^{\sigma/\nu}\,. \tag{10}$$

Reweighting has to cover a reasonable range about the maximum, say from β_L^{r-} to $\beta_L^{r+} > \beta_L^{r-}$ defined as solutions of

$$\widehat{S}_L(\beta) = r\,\widehat{S}_L^{\max},\;\; 0 < r < 1\,. \tag{11}$$

We define $\beta_L^r \in \{\beta_L^{r-}, \beta_L^{r+}\}$ to be the $\beta_L^{r^\pm}$ value which is further away from β_c than the other and assume

$$\triangle\beta_L^r = \left|\beta_L^r - \beta_c\right| = a^r\,L^{-\kappa}\,, \tag{12}$$

where a^r and $\kappa > 0$ are constants (κ independent of r and a^r becomes large for r small). For sufficiently large L we suppose that

$$\widehat{S}_L(\beta_L^r) = S^{\rm reg} + A\;(\triangle\beta_L^r)^{-\sigma} \tag{13}$$

holds, where $S^{\rm reg} = S^{\rm reg}(\beta)$ is a regular background term. Combining Eqs. (10) to (13) we have $L^{\sigma/\nu} \sim L^{\sigma\kappa}$ and conclude

$$\kappa = 1/\nu\,, \tag{14}$$

i.e., the desired range (12) scales with the same exponent as the canonical range (9). However, the proportionality factor a^r in (12) can be much larger than the one encountered for the canonical range. With the modest value $r = 2/3$ this point is made in Fig. 1 for the 3D Ising model on an 80^3 lattice. We plot the expectation values of the specific heat $C(\beta)$ and of the first structure factor $S(\beta)$ whose maximum scales $\sim L^{\gamma/\nu}$ (see, e.g., Ref. [16] for a definition of structure factors and [17] for their recent MCMC calculation in some Potts models). The desired reweighting range is then from the smallest β_L^{r-} to the largest β_L^{r+} over all considered observables. In Fig. 1 it is seen to be more than 17 times larger than the canonical reweighting range from a simulation at $\beta_L^{\max}$

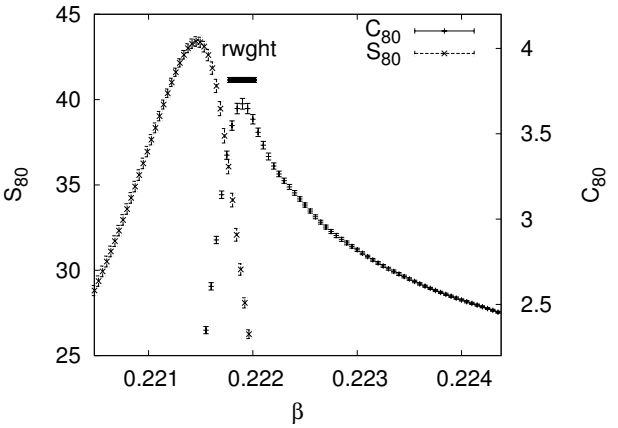

Figure 1: Canonical (indicated by "rwght") versus desired (entire β axis) reweighting range on an 80^3 lattice.

of the specific heat (in realistic applications one does not know β_c a-priori and $\beta_L^{\max}$ of a suitable observable is a good substitute).

In case of a logarithmic singularity the $t \to 0$ multiplicative cancellation (4) of the singularity is no longer possible, but becomes additive:

$$\widehat{S}_L(\beta) = S^{\rm reg} - A \ln|t| + f(L/\xi) = S^{\rm reg} - A \ln|t| + f(|t|^\nu L) \tag{15}$$

and

$$f(|t|^\nu L) \;=\; \frac{A}{\nu} \ln(|t|^\nu L) \;=\; A \ln|t| + \frac{A}{\nu} \ln(L)\,, \tag{16}$$

$$\widehat{S}_L(\beta_c) \;=\; S^{\rm reg} + \frac{A}{\nu} \ln(L)\,, \quad \widehat{S}_L^{\max} \;=\; S^{\rm reg}_{\max} + \frac{A_m}{\nu} \ln(L) \tag{17}$$

follows. Further

$$\frac{r\,A_m}{\nu} \ln(L) \sim r\,\widehat{S}_L^{\max} = \widehat{S}_L(\beta_L^r) = r\,S^{\rm reg}_{\max} - A_m \ln|t_L^r| \sim A_m\,\kappa \ln(L) \tag{18}$$

and one finds that the exponent κ in Eq. (12) is no longer independent of r, but

$$\kappa \;=\; r/\nu\,. \tag{19}$$

While the canonical reweighting range scales still $\sim L^{-1/\nu}$, the desired reweighting range becomes $\sim L^{-r/\nu}$, so that the ratio desired/canonical diverges $\sim L^{(1-r)/\nu}$. With $S = C$ the 2D Ising model provides an example.

In conclusion many more simulations than one canonical are typically needed to cover a relevant part of the scaling region of a second-order phase transition. In principle this can be done by patching canonical simulations from several temperatures together, relying on a multi-histogram approach [18]. Besides that dealing with many simulations is tedious, weaknesses of

these approaches are that the histograms fluctuate independently and that their patching has to be done in regions where the statistics is reduced due to the decline of the number of histograms entries. More stable estimates are obtained by constructing a generalized ensemble, which allows the random walker to cover the entire region of interest.

3. Multibondic Cluster Updating and Wang-Landau Recursion

Generalized ensemble simulations require two steps:

1. Obtain a working estimate of the weight factors.
2. Perform a MCMC simulation with fixed weights.

To be definite we confine our discussion to MUCA simulations [3]. Extension to cluster algorithms are known [13, 19] and we will rely on the MUBO version [13]. This defines step 2, but leaves still many options open to deal with step 1. "Working estimate" means that the approximation of the weights of the generalized ensemble is good enough so that the energy range in question is covered in step 2. Historically step 1 has been one of the stumbling blocks of umbrella sampling: "The difficulty of finding such weighting factors has prevented wide applications of the umbrella sampling method to many physical systems" [20]. Most convenient is to have an efficient general purpose recursion for step 1 at hand. While designs were reported in a number of papers [21], see also Refs. [7, 8, 19], the winning approach appears to be the one of Wang and Landau (WL) [5] (although somewhat surprisingly we found only one comparative study [22]).

The WL recursion differs fundamentally from the earlier approaches by iterating the weight at energy E *multiplicatively* with a factor $f_{\rm WL} > 1$ rather than additively. At a first glance the WL approach is counter-intuitive, because the correct iteration of the weight factor close to the desired fixed point is obviously proportional to one over the number of histogram entries, $1/H(E)$, and not to $1/f_{\rm WL}^{H(E)}$. However, what matters is a rapid approach to a working estimate. The advantage of the WL over the other recursions is that it moves right away rapidly through the targeted energy range. When this range is sufficiently covered, the iteration factor is refined by $f_{\rm WL} \to \sqrt{f_{\rm WL}}$, so that it approaches 1 rapidly. Once the system cycles with frozen weights through the desired energy range the goal of a working estimate has been reached and the WL recursion is no longer needed [23]. We now generalize this approach to cluster algorithms.

We consider q-state Potts models with energy function

$$E = -2 \sum_{\langle ij \rangle} \delta_{q_i q_j} \,, \tag{20}$$

where the sum is over the nearest-neighbor sites of a D-dimensional cubic lattice of $N = L^D$ Potts spins, which take the values $q_i = 1, \ldots, q$. The factor of two has been introduced so that the special case $q = 2$ matches with the energy and β conventions of the Ising model literature.

In the Fortuin-Kasteleyn (FK) cluster language [24] the Potts model partition is written as

$$\mathcal{Z}_{\rm FK} = \sum_{\{q_i\}} \sum_{\{b_{ij}\}} Z(\{q_i\}, \{b_{ij}\}) \tag{21}$$

with Boltzmann weight

$$Z(\{q_i\}, \{b_{ij}\}) = \prod_{\langle ij \rangle} \left[a\, \delta_{q_i q_j}\, \delta_{b_{ij}1} + \delta_{b_{ij}0} \right] \tag{22}$$

where $a = e^{2\beta} - 1$. The bond variables b_{ij} (simply called bonds in the following) take the values $b_{ij} = 0$ and 1, interpreted as "broken" and "active" or "set" bonds, respectively. Carrying out the sum over the $\{b_{ij}\}$-configurations yields the canonical Potts Boltzmann factor:

$$Z(\{q_i\}) = \prod_{\langle ij \rangle} \left[a\, \delta_{q_i q_j} + 1 \right] = \prod_{\langle ij \rangle} \left[e^{2\beta}\, \delta_{q_i q_j} - \delta_{q_i q_j} + 1 \right] = \prod_{\langle ij \rangle} e^{2\beta \delta_{q_i q_j}} \tag{23}$$

as for $e^{2\beta x} = e^{2\beta}\, x - x + 1$ holds for $x = \delta_{q_i q_j} = 0,\, 1$. Therefore,

$$\mathcal{Z}_{\rm Potts} = \sum_{\{q_i\}} Z(\{q_i\}) = \sum_{\{q_i\}} \prod_{\langle ij \rangle} e^{2\beta \delta_{q_i q_j}} = \sum_{\{q_i\}} e^{-\beta E(\{q_i\})}\,. \tag{24}$$

The Swendsen-Wang cluster updating procedure [11] starts from the FK partition function (21) and generates new bonds b'_{ij} on links with $\delta_{q_i q_j} = 1$: Bonds $b'_{ij} = 1$ with probability p_1 and bonds $b'_{ij} = 0$ with probability p_0 so that $p_1/p_0 = a$ and $p_0 + p_1 = 1$ holds. This gives $p_1 = 1 - e^{-2\beta}$ for $b'_{ij} = 1$ and $p_0 = 1 - p_1 = e^{-2\beta}$ for $b'_{ij} = 0$. On $\delta_{q_i q_j} = 0$ links we have $b'_{ij} = 0$ with probability one. A cluster of spins is defined as a set of spins connected by active bonds and in the subsequent spin update one assigns randomly a new value $q'_i = 1, \ldots, q$ to an entire cluster of spins, $\{q_i\} \to \{q'_i\}$.

Let us denote the number of active bonds by $B = \sum_{\langle ij \rangle} b_{ij}$. The MUBO partition function [13] is defined by

$$\mathcal{Z}_{\rm MUBO} = \sum_{\{q_i\}} \sum_{\{b_{ij}\}} Z(\{q_i\}, \{b_{ij}\})\, W(B) \tag{25}$$

where a bond weight factor $W(B)$ has been introduced. A valid updating procedure for the configurations of this partition function is formulated in the following:

A. For $q_i \neq q_j$ a bond is never set. This applies to the initial as well as to the updated bond on this link, so that B does not change.

B. For $q_i = q_j$ there are two possibilities:

1. The initial bond is not set, $b_{ij} = 0$. Then $B' = B$ for $b'_{ij} = 0$ and $B' = B + 1$ for $b'_{ij} = 1$. The updating probabilities are ($a = p_1/p_0 = e^{2\beta} - 1$)

$$P_1(0 \to 0) \;=\; \frac{W(B)}{W(B) + a\, W(B+1)} \tag{26}$$

and $P_1(0 \to 1) = 1 - P_1(0 \to 0)$.

2. The initial bond is set, $b_{ij} = 1$. Then $B' = B - 1$ for $b'_{ij} = 0$ and $B' = B$ for $b'_{ij} = 1$. The updating probabilities are

$$P_2(1 \to 0) \;=\; \frac{W(B-1)}{W(B-1) + a\, W(B)} \tag{27}$$

and $P_2(1 \to 1) = 1 - P_2(1 \to 0)$.

After the configuration is partitioned into clusters [25], the update is completed as in a canoncial cluster algorithm by assigning with uniform probability a spin in the range $1, \ldots, q$ to each cluster.

In its generalization to cluster algorithms the WL recursion updates then $\ln W(B)$ according to

$$\ln W(B) \;\rightarrow\; \ln W(B) - a_{\rm WL}\,, \quad a_{\rm WL} = \ln(f_{\rm WL})\,, \tag{28}$$

whenever a configuration with B bonds is visited. After we derived this equation, we learned that it had been previously obtained by Yamaguchi and Kawashima [26]. All recursions are started with $a_{\rm WL} = 1$ and we iterate $a_{\rm WL} \rightarrow a_{\rm WL}/2$ according to the following criteria:

1. The Markov chain just cycled from $\overline{B}_L^{r-}$ to $\overline{B}_L^{r+}$ and back. Here $\overline{B}_L^{r-}$ and $\overline{B}_L^{r+}$ are bond estimates corresponding to β_L^{r-} and β_L^{r+}, respectively, determined by short canonical simulations.
2. The bond histogram $h(B)$, measured since the last iteration, fulfilled a flatness criterion $h_{\rm min}/h_{\rm max} > cut$, where cut was equal to $1/3$ in most of our runs.
3. We freeze the weights after a last iteration is performed with the desired minimum value $a_{\rm WL}^{\rm min}$.

4. Numerical Results

With fixed weights, after a short equilibration run, measurements are performed during the subsequent simulation which was tuned to cover approximately 1 000 cycling events. Canonical expectation values at inverse temperature β, $\beta_L^{r-} \le \beta \le \beta_L^{r+}$ are obtained by reweighting (25). Table 1 gives an overview of our 3D Ising model statistics. The effectiveness of the recursion is seen from the fact that it takes never more than 3% percent of the statistics used for production (these numbers are in sweeps). Similarly the initial simulations, which determine $\overline{B}_L^{r-}$ and $\overline{B}_L^{r+}$, take less than 3%. If this is done by canonical, say Metropolis or heatbath simulations, it is important to start the simulation for $\overline{B}_L^{r-}$ in the disordered and for $\overline{B}_L^{r+}$ in the ordered phase, because crossing the phase transition may lead to a divergence of the needed computer time with lattice size.

From the production run we calculate integrated autocorrelation times $\tau_{\rm int}$ and compare them in Fig. 2 with those of a MUCA simulation of similar statistics (jackknife error bars with respect to the number of bins given in Table 1 are shown, see [7] for computational details). From the MUBO time series we calculated $\tau_{\rm int}$ for (a) energies and (b) bonds and found the results almost

Table 1: Parameters of the 3D Ising model simulations on L^3 lattices.

L	β_L^{r-}	β_L^{r+}	$a_{\rm WL}^{\rm min}$	recursion	production
20	0.210 649	0.233 690	2^{-18}	19 962	$32 \times 32\,768$
30	0.216 443	0.229 336	2^{-18}	27 344	$32 \times 32\,768$
44	0.218 545	0.227 013	2^{-19}	33 266	$32 \times 65\,536$
56	0.219 755	0.225 914	2^{-19}	56 323	$32 \times 65\,536$
66	0.220 063	0.224 709	2^{-21}	62 884	$32 \times 131\,072$
80	0.220 482	0.224 377	2^{-21}	108 618	$36 \times 131\,072$

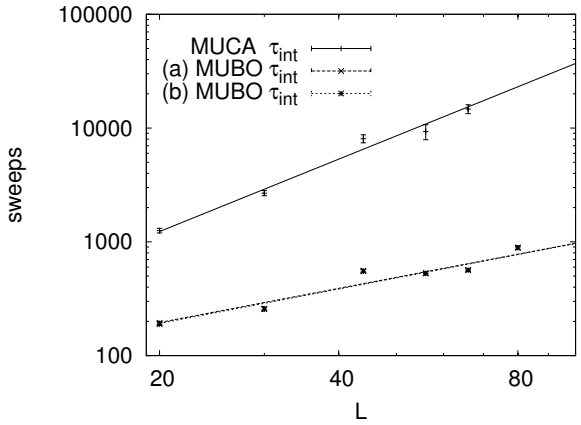

Figure 2: Integrated autocorrelation time $\tau_{\rm int}(L)$ for the 3D Ising model (see text).

identical (slightly higher for the energies, but indistinguishable on the scale of the figure). For MUCA the estimates are from energies. Up to a constant factor practically identical results are obtained from cycling times. In our code one MUCA sweep was about three times faster than one MUBO sweep.

The critical slowing down is described by $\tau_{\rm int} \sim L^z$. For the dynamical critical exponent we find $z = 2.22\,(11)$ for MUCA and $z = 1.05\,(5)$ for MUBO. So the performance gain is a bit better than linear with the lattice size L. The data in Fig. 2 scatter more than one might have expected about the fits because our β_L^{r-} and β_L^{r+} values are based on MCMC estimates, which are by themselves noisy. Our exponent for cluster updating is significantly higher than the one estimated from simulations at β_c, $z = 0.50\,(3)$, for the Swendsen-Wang algorithm [27]. The reason is that the efficiency of the cluster algorithm deteriorates off the critical point, even when one is still in the scaling region. Therefore, we think that our exponent of $z \approx 1$ reflects the slowing down in realistic applications more accurately than the small value of the literature. In particular the cluster algorithm becomes rather inefficient for calculating the long tail of the specific heat for $\beta > \beta_L^{\rm max}$.

In Fig. 3 we show integrated autocorrelation times from simulations of the 2D Ising model for which we adjusted our simulation parameter to cover the full width at half-maximum of the specific heat. This corresponds to $r = 1/2$ in Eq. (19). The dynamical critical exponent takes than the values $z = 2.50\,(4)$ for MUCA and $z = 1.04\,(2)$ for MUBO. The MUCA value reflects that the number of canonical simulations needed to cover the desired energy range grows now $\sim L^{1/2}$, while the canonical critical value is approximately two [7, 28] (a slightly higher value, $z = 2.17$ has been reported in the literature [28], but the method which we employ tends to give lower results, see p. 217 of Ref. [7]).

5. Summary and Conclusions

At second-order phase transitions, the energy range explored by canonical MCMC simulations is often too small to allow accurate reweighting analyses around the peak of a given quan-

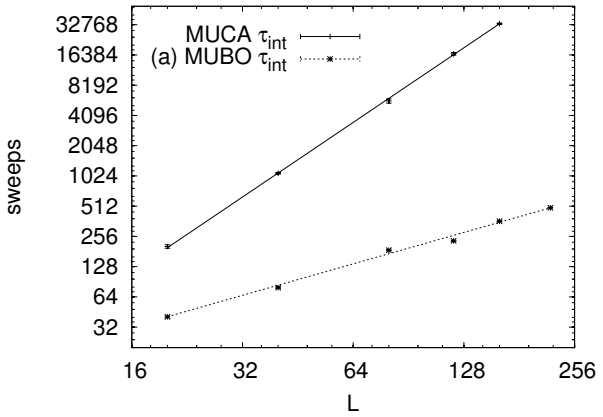

Figure 3: Integrated autocorrelation time $\tau_{\rm int}(L)$ for the 2D Ising model.

tity. Generalized ensembles such as those generated with the multicanonical or Wang-Landau methods can provide specifically tailored broader ranges but suffer in their conventional formulations from large autocorrelation times due to the use of local update algorithms. We first recall that much more efficient non-local cluster algorithms can be employed in the multibondic formulation and then propose a generalization of this method to weight recursions of Wang-Landau type. When the desired broad reweighting range is adjusted to the lattice size L according to finite-size scaling theory, we obtain for the Ising model with this cluster approach a gain in efficiency by power laws in L ($\simeq L^{1.45}$ in 2D and $\simeq L^{1.15}$ in 3D) compared to the conventional Wang-Landau multicanonical approach with local updates.

Finally we remark that the efficiency of simulations of second-order phase transitions can presumably be further improved by optimizing the weights with respect to cycling along the lines introduced in Ref. [29].

Acknowledgments

This work started while BB was the Leibniz Professor at Leipzig University. In part it was supported by the US Department of Energy under contract DE-FG02-97ER41022 and by the Deutsche Forschungsgemeinschaft (DFG) under contract JA 483/23-1.

References

[1] J. Gubernatis (editor):, *The Monte Carlo Method in Physical Sciences: Celebrating the 50th Anniversary of the Metropolis Algorithm*, AIP Conference Proceedings, Vol. **690** (Melville, NY, 2003).
[2] G.M. Torrie, J.P. Valleau: J. Comp. Phys. **23**, 187 (1977).
[3] B.A. Berg, T. Neuhaus: Phys. Rev. Lett. **68**, 9 (1992); B.A. Berg, T. Celik: Phys. Rev. Lett. **69**, 2292 (1992).
[4] A.P. Lyubartsev, A.A. Martsinovski, S.V. Shevkanov, P.N. Vorontsov-Velyaminov: J. Chem. Phys. **96**, 1776 (1992); E. Marinari, G. Parisi: Europhys. Lett. **19**, 451 (1992).
[5] F. Wang, D.P. Landau: Phys. Rev. Lett **86**, 2050 (2001).
[6] U.H. Hansmann, Y. Okamoto: Ann. Rev. Comp. Phys. **6**, 120 (1999).

[7] B.A. Berg: *Markov Chain Monte Carlo Simulations and Their Statistical Analysis* (World Scientific, Singapore, 2004).
[8] W. Janke: in *Computer Simulations of Surfaces and Interfaces*, NATO Science Series, II. Math., Phys. and Chem., Vol. **114** (Kluwer, Dordrecht, 2003) p. 137.
[9] B.A. Berg, W. Janke: Phys. Rev. Lett. **98**, 040602 (2007).
[10] D.P. Landau, S.-H. Tsai, M. Exler: Am. J. Phys. **72**, 1294 (2004).
[11] R.H. Swendsen, J.-S. Wang: Phys. Rev. Lett. **58**, 86 (1987).
[12] U. Wolff: Phys. Rev. Lett. **62**, 361 (1989).
[13] W. Janke, S. Kappler: Phys. Rev. Lett. **74**, 212 (1995); M.S. Carroll, W. Janke, S. Kappler: J. Stat. Phys. **90**, 1277 (1998).
[14] A. Pelissetto E. Vicari: Phys. Rep. **368**, 549 (2002).
[15] M.E. Fisher: 'The theory of critical point singularities'. In *Critical Phenomena, Proceedings of the 1970 Enrico Fermi International School of Physics*, ed. by M.S. Green (Academic Press, New York 1971) pp. 1–99.
[16] H.E. Stanley: *Introduction to Phase Transitions and Critical Phenomena* (Clarendon Press, Oxford, 1971) p. 98.
[17] A. Bazavov, B.A. Berg, S. Velytsky: Phys. Rev. D **74**, 014501 (2006).
[18] A.M. Ferrenberg, R.H. Swendsen: Phys. Rev. Lett. **63**, 1195 (1989); N.A. Alves, B.A. Berg, R. Villanova: Phys. Rev. B **41**, 383 (1990).
[19] J.V. Lopez, M.D. Costa, J.M.B. Lopez dos Santos, R. Toral: Phys. Rev. E **74**, 045702 (2006).
[20] Z. Li, H.A. Scheraga: J. Mol. Struct. (Theochem) **179**, 333 (1988).
[21] B.A. Berg: J. Stat. Phys. **82**, 323 (1996); B.A. Berg, A. Billoire, W. Janke: Phys. Rev. B **61**, 12143 (2000); Y. Sugita, Y. Okamoto: Chem. Phys. Lett. **329**, 261 (2000).
[22] Y. Okamoto: in [1]. Also available as cond-mat/0308119.
[23] For $f_{\rm WL}$ very close to one the difference to a simulation with fixed weights becomes negligible and the original suggestion of WL is to keep on iterating $f_{\rm WL} \to \sqrt{f_{\rm WL}}$ during the entire simulation. However, that prevents early equilibration and the important consistency check that the equilibrium Markov process covers the targeted energy range. Therefore, we advise to freeze the weights whenever a working approximation has been reached.
[24] C.M. Fortuin, P.W. Kasteleyn: Physica **57**, 536 (1972).
[25] This is the Swendsen-Wang approach [11]. To work out the Wolff version [12] is left as an exercise to the reader.
[26] C. Yamaguchi, N. Kawashima: Phys. Rev. E **65**, 056710 (2002). We thank Professor Y. Okabe for informing us about this reference.
[27] U. Wolff: Phys. Lett. A **228**, 379 (1989).
[28] D.P. Landau, K. Binder: *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, Cambridge, 2000).
[29] S. Trebst, D.A. Huse, M. Troyer: Phys. Rev. E **70**, 046701 (2004).