# Wang-Landau Multibondic Cluster Simulations for Second-Order Phase Transitions

Bernd A. Berg[1,2] and Wolfhard Janke[3]

[1]*Department of Physics, Florida State University, Tallahassee, Florida 32306, USA*
[2]*School of Computational Science, Florida State University, Tallahassee, Florida 32306, USA*
[3]*Institut für Theoretische Physik and Centre for Theoretical Sciences (NTZ), Universität Leipzig, Augustusplatz 10/11, D-04109 Leipzig, Germany*

(Received 14 October 2006; published 22 January 2007)

For a second-order phase transition the critical energy range of interest is larger than the energy range covered by a canonical Monte Carlo simulation at the critical temperature. 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 one loses the advantage due to cluster algorithms. A cluster version of the Wang-Landau recursion together with a subsequent multibondic simulation improves for 2D and 3D Ising models the efficiency of the conventional Wang-Landau or multicanonical approach by power laws in the lattice size. In our simulations real gains in CPU time reach 2 orders of magnitude.

DOI: 10.1103/PhysRevLett.98.040602 PACS numbers: 05.10.Ln, 02.50.Ng, 05.20.−y, 64.60.Cn

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 [2–5]; for reviews see Refs. [6–8].

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 for second-order phase transitions, although the convenience of such applications is claimed by Landau and collaborators [9]. However, they lose the crucial advantage which cluster algorithms [10,11] provide for MCMC simulations of second-order phase transitions. Here we present a generalization to cluster algorithms. To keep the Letter accessible for nonexperts, we restrict our investigations to 2D and 3D Ising models, while the points made are generally valid for cluster algorithms.

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 simulation at the (infinite volume) critical temperature $T_c = 1/\beta_c$?

For simplicity, our lattices are of shape $L^D$ and periodic boundary conditions are assumed. We denote the probability density of the energy from a canonical MCMC simulation by $P(E)$. Finite-size scaling (FSS) arguments [12] imply $C \sim L^{\alpha/\nu}$ for the specific heat at $\beta_c$, where $\alpha$ and $\nu$ are, respectively, the critical exponents of the specific heat $C$ and the correlation length $\xi$. A second-order phase transition requires $\nu > 0$. Let us first assume $\alpha > 0$. The fluctuation-dissipation theorem gives

$$\langle (E - \hat{E})^2 \rangle \sim L^{D+\alpha/\nu} \quad \text{where } \hat{E} = \langle E \rangle; \tag{1}$$

implying for the range covered by the simulation at $\beta_c$

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

where $E_q$, $q = 0.25$ and $q = 0.75$, are fractiles of the energy distribution [7]. In the vicinity of $\beta_c$ ($A$ constant)

$$\hat{E}(\beta)/L^D = \hat{E}(\beta_c)/L^D + A(\beta - \beta_c)^{1-\alpha}, \tag{3}$$

and using the hyperscaling relation [12] $\alpha = 2 - D\nu$, Eq. (2) translates into a reweighting range

$$\triangle \beta \sim L^{-1/\nu}. \tag{4}$$

The desired reweighting range is determined by the need to cover the maxima of all divergent observables measured. Let the maximum value of such an observable $\hat{S}_L(\beta)$ be $\hat{S}_L^{\max} = \hat{S}_L(\beta_L^{\max})$ and denote the critical exponent of $S$ by $\sigma$. Then FSS theory implies

$$\hat{S}_L^{\max} \sim L^{\sigma/\nu}. \tag{5}$$

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

$$\hat{S}_L(\beta) = r\hat{S}_L^{\max}, \qquad 0 < r < 1, \tag{6}$$

which becomes large for $r$ small. 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 $\beta_c$ than the other and assume

$$\triangle \beta_L^r = |\beta_L^r - \beta_c| = a^r L^{-\kappa}, \tag{7}$$

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

 

$$\hat{S}_L(\beta_L^r) = S^{\mathrm{reg}} + A(\triangle\beta_L^r)^{-\sigma} \tag{8}$$

holds, where $S^{\mathrm{reg}}$ is a regular background term. Combining Eqs. (5), (7), and (8) we conclude

$$\kappa = 1/\nu; \tag{9}$$

i.e., the desired range (7) scales in the same way as the canonical range (4). However, the proportionality factor $a^r$ 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 specific heat $C(\beta)$ and for $S(\beta)$ the first structure factor (see, e.g., Ref. [13]), whose maximum scales $\sim L^{\gamma/\nu}$. The desired reweighting range is more than 17 times larger than the canonical reweighting range from a simulation at $\beta_L^{\max}$ of the specific heat (in realistic applications one does not know $\beta_c$ *a priori* and $\beta_L^{\max}$ of a suitable observable is a good substitute).

Using the same line of arguments for a logarithmic singularity

$$S(\beta) = S^{\mathrm{reg}} - A\ln|\beta - \beta_c|, \tag{10}$$

one finds that the exponent $\kappa$ in Eq. (7) is no longer independent of $r$, but

$$\kappa = r/\nu. \tag{11}$$

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 than one canonical simulation 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 multihistogram approach [14]. Besides 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. This requires two steps: 1. Obtain a working estimate of the weight factors. 2. Perform a MCMC simulation with fixed weights.

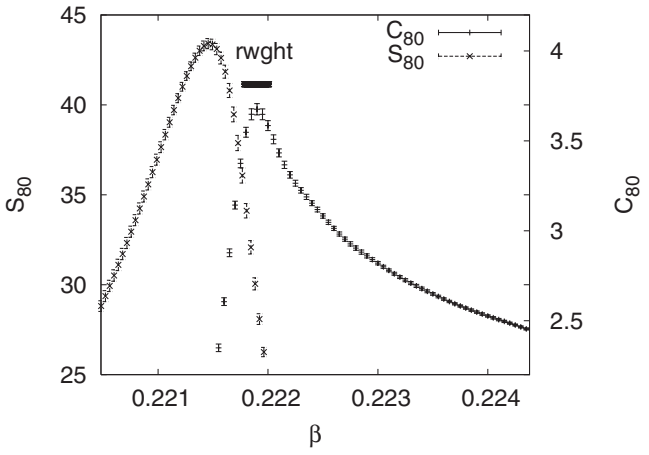


FIG. 1. Canonical (indicated by "rwght") versus desired (entire $\beta$ axis) reweighting range on an $80^3$ lattice.

To be definite we confine our discussion to the multicanonical (MUCA) simulations [3]. Extensions to cluster algorithms are known [15,16]. We will rely on multibondic (MUBO) simulations [15]. 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" [17]. Most convenient is to have an efficient general purpose recursion for step 1 at hand. While designs were reported in a number of papers [18], see also Refs. [7,8,16], the winning approach appears to be the one of Wang and Landau (WL) [5] (although somewhat surprisingly we found only one comparative study [19]).

The WL recursion differs fundamentally from the earlier approaches by iterating the weight at energy $E$ *multiplicatively* with a factor $f_{\mathrm{WL}} > 1$ rather than additively. At a first glance the WL approach is counterintuitive, 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 $H(E)$ and not to $1/f_{\mathrm{WL}}^{H(E)}$. 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 it is sufficiently covered, the iteration factor is refined by $f_{\mathrm{WL}} \to \sqrt{f_{\mathrm{WL}}}$, so that $f_{\mathrm{WL}}$ 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 [20]. We now generalize this approach to cluster algorithms.

We use the energy function of the $q$-state Potts models,

$$E = -2\sum_{\langle ij\rangle}\delta_{q_i q_j}, \tag{12}$$

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 2 has been introduced so that $q = 2$ matches with the energy and $\beta$ conventions of the Ising model literature.

In the Fortuin-Kasteleyn (FK) cluster language [21] the Potts model partition is written as

$$Z_{\mathrm{FK}} = \sum_{\{q_i\}} \sum_{\{b_{ij}\}} Z(\{q_i\}, \{b_{ij}\}) \quad \text{with}$$
$$Z(\{q_i\}, \{b_{ij}\}) = \prod_{\langle ij \rangle} [a \delta_{q_i q_j} \delta_{b_{ij} 1} + \delta_{b_{ij} 0}], \tag{13}$$

where $a = e^{2\beta} - 1$. For a fixed configuration $\{q_i\}$ of Potts states the Swendsen-Wang updating procedure [10] is to generate bond variables $b_{ij}$ (simply called bonds in the following) on links with $\delta_{q_i q_j} = 1$: Bonds $b_{ij} = 1$ are generated with probability $p$ and bonds $b_{ij} = 0$ with probability $q$ so that $p/q = a$ and $p + q = 1$ holds. This gives $p = 1 - e^{-2\beta}$ for $b_{ij} = 1$ and $q = 1 - p = e^{-2\beta}$ for $b_{ij} = 0$. On $\delta_{q_i q_j} = 0$ links we have the new bonds $b'_{ij} = 0$ with probability one. We call bonds with $b_{ij} = 1$ active or set. A cluster of spins is defined as a set of spins connected by active bonds and an update is to flip entire clusters of spins, $\{q_i\} \to \{q'_i\}$.

Let us denote the number of active bonds by $B$. The MUBO partition function [15] is defined by

$$Z_{\mathrm{MUBO}} = \sum_{\{q_i\}} \sum_{\{b_{ij}\}} Z(\{q_i\}, \{b_{ij}\}) W(B), \tag{14}$$

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

$$P_1(0 \to 0) = \frac{qW(B)}{qW(B) + pW(B+1)} \tag{15}$$

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

$$P_2(1 \to 0) = \frac{qW(B-1)}{qW(B-1) + pW(B)} \tag{16}$$

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

After the configuration is partitioned into clusters [22], the update is completed 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

$$\ln W(B) \to \ln W(B) - a_{\mathrm{WL}}, \qquad a_{\mathrm{WL}} = \ln(f_{\mathrm{WL}}), \tag{17}$$

whenever a configuration with $B$ bonds is visited. All recursions are started with $a_{\mathrm{WL}} = 1$ and we iterate $a_{\mathrm{WL}} \to a_{\mathrm{WL}}/2$ according to the following criteria: 1. The Markov chain just cycled from $\bar{B}_L^{r-}$ to $\bar{B}_L^{r+}$ and back. Here $\bar{B}_L^{r-}$ and $\bar{B}_L^{r+}$ are bond estimates corresponding to $\beta_L^{r-}$ and $\beta_L^{r+}$, respectively, determined by short canonical simulations with starting configurations in the corresponding phases. 2. The bond histogram $h(B)$, measured since the last iteration, fulfilled a flatness criterion $h_{\min}/h_{\max} > \mathrm{cut}$, 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 $a_{\mathrm{WL}}^{\min}$.

TABLE I. 3D Ising model simulations on $L^3$ lattices.

| $L$ | $\beta_L^{r-}$ | $\beta_L^{r+}$ | $a_{\mathrm{WL}}^{\min}$ | 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 | $36 \times 131\,072$ |

After a short equilibration run, measurements are performed during the subsequent simulation with fixed weights, each tuned to approximately 1000 cycling events. Canonical expectation values at inverse temperature $\beta$, $\beta_L^{r-} \leq \beta \leq \beta_L^{r+}$ are obtained by reweighting (14). Table I 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 $\bar{B}_L^{r-}$ and $\bar{B}_L^{r+}$, take less than 3%.

From the production statistics we calculate integrated autocorrelation times $\tau_{\mathrm{int}}$ and compare them in Fig. 2 with those of a MUCA simulation. From the MUBO time series we calculated $\tau_{\mathrm{int}}$ for (a) energies and (b) bonds and found the results almost 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 3 times faster than one MUBO sweep.

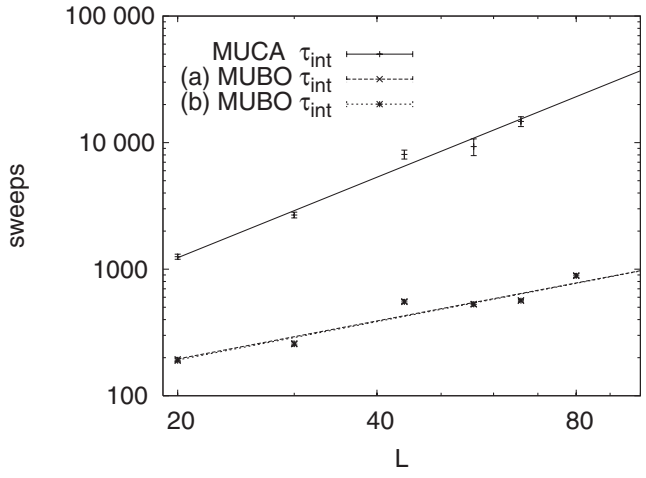


FIG. 2. $\tau_{\mathrm{int}}(L)$ for the 3D Ising model (see text).

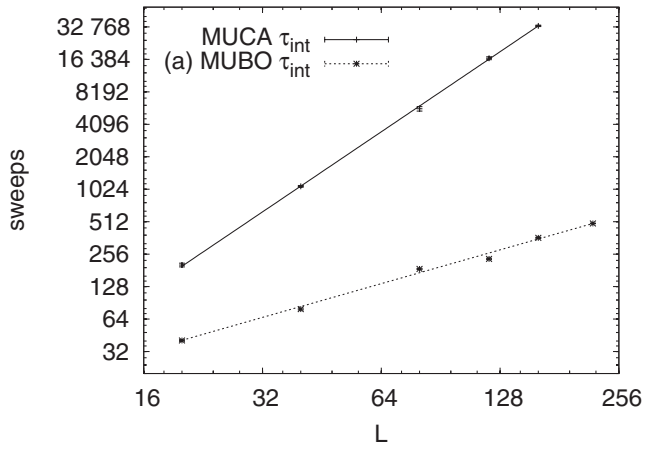


FIG. 3. $\tau_{\mathrm{int}}(L)$ for the 2D Ising model.

The critical slowing down is $\sim L^z$. 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 $\beta_L^{r-}$ and $\beta_L^{r+}$ 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 $\beta_c$, $z = 0.50(3)$, for the Swendsen-Wang algorithm [23]. 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 real application 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 $\beta > \beta_L^{\max}$.

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. (11). The dynamical critical exponent takes then 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 [7,24].

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. [25].

This work started while B. B. was at Leipzig University. In part it was supported by the U. S. Department of Energy under Contract No. DE-FG02-97ER41022 and by the Deutsche Forschungsgemeinschaft (DFG) under Contract No. JA 483/23-1.

*Note added in proof.*—After submitting this Letter we learned from Prof. Y. Okabe that Eq. (17) was previously derived in Ref. [26].

[1] *The Monte Carlo Method in Physical Sciences*, edited by J. Gubernatis, AIP Conf. Proc. No. 690 (AIP, New York, 2003).
[2] G. M. Torrie and J. P. Valleau, J. Comput. Phys. **23**, 187 (1977).
[3] B. A. Berg and T. Neuhaus, Phys. Rev. Lett. **68**, 9 (1992).
[4] A. P. Lyubartsev *et al.*, J. Chem. Phys. **96**, 1776 (1992); E. Marinari and G. Parisi, Europhys. Lett. **19**, 451 (1992).
[5] F. Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001).
[6] U. H. E. Hansmann and Y. Okamoto, Annu. Rev. Comput. Phys. **6**, 129 (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, edited by B. Dünweg, D. P. Landau, and A. I. Milchev (Kluwer, Dordrecht, 2003), Vol. 114, p. 137.
[9] D. P. Landau, S.-H. Tsai, and M. Exler, Am. J. Phys. **72**, 1294 (2004).
[10] R. H. Swendsen and J.-S. Wang, Phys. Rev. Lett. **58**, 86 (1987).
[11] U. Wolff, Phys. Rev. Lett. **62**, 361 (1989).
[12] A. Pelissetto and E. Vicari, Phys. Rep. **368**, 549 (2002).
[13] H. E. Stanley, *Introduction to Phase Transitions and Critical Phenomena* (Clarendon, Oxford, 1971), p. 98.
[14] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **63**, 1195 (1989); N. A. Alves, B. A. Berg, and R. Villanova, Phys. Rev. B **41**, 383 (1990).
[15] W. Janke and S. Kappler, Phys. Rev. Lett. **74**, 212 (1995).
[16] J. Viana Lopes *et al.*, Phys. Rev. E **74**, 046702 (2006).
[17] Z. Li and H. A. Scheraga, J. Mol. Struct. (Theochem) **179**, 333 (1988).
[18] B. A. Berg, J. Stat. Phys. **82**, 323 (1996); B. A. Berg, A. Billoire, and W. Janke, Phys. Rev. B **61**, 12 143 (2000); Y. Sugita and Y. Okamoto, Chem. Phys. Lett. **329**, 261 (2000).
[19] Y. Okamoto, in [1]. Also available as cond-mat/0308119.
[20] For $f_{\mathrm{WL}}$ very close to 1, the difference to a simulation with fixed weights becomes negligible and the original suggestion of WL is to keep on iterating $f_{\mathrm{WL}} \to \sqrt{f_{\mathrm{WL}}}$. However, that prevents early equilibration and the important consistency check that the *equilibrium* Markov process covers the targeted energy range.
[21] C. M. Fortuin and P. W. Kasteleyn, Physica (Amsterdam) **57**, 536 (1972).
[22] This is the Swendsen-Wang approach [10]. To work out the Wolff version [11] is left as an exercise to the reader.
[23] U. Wolff, Phys. Lett. B **228**, 379 (1989).
[24] D. P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, Cambridge, England, 2000).
[25] S. Trebst, D. A. Huse, and M. Troyer, Phys. Rev. E **70**, 046701 (2004).
[26] C. Yamaguchi and N. Kawashima, Phys. Rev. E **65**, 056710 (2002).