Multibondic Cluster Algorithm for Monte Carlo Simulations of First-Order Phase Transitions

Wolfhard Janke and Stefan Kappler

Institut für Physik, Johannes Gutenberg-Universität Mainz, Staudinger Weg 7, 55099 Mainz, Germany

(Received 6 July 1994)

Inspired by the multicanonical approach to simulations of first-order phase transitions we propose for q-state Potts models a combination of cluster updates with reweighting of the bond configurations in the Fortuin-Kastelein-Swendsen-Wang representation of this model. Numerical tests for the two-dimensional models with $q = 7$, 10, and 20 show that the autocorrelation times of this algorithm grow with the system size V as $\tau \propto V^{\alpha}$, where the exponent takes the optimal random walk value of $\alpha \approx 1$.

PACS numbers: 05.50.+q, 11.15.Ha, 64.60.Cn, 75.10.Hk

Conventional Monte Carlo simulations of phase transitions suffer from strong correlations between successive configurations. For continuous phase transitions the reason for the slowing down can be traced back to the divergence of spatial correlations at criticality. At first-order phase transitions one encounters a different and even worse type of slowing down problem which is caused by exponentially small tunneling rates between coexisting phases separated by high energy barriers [1]. Since the inverse rate is proportional to the autocorrelation time τ of the Monte Carlo process, thus being a measure for the average time between tunnelings, this implies for a D-dimensional system of size $V = L^D$ an exponentially fast growth of $\tau \propto \exp(cL^{D-1})$ with increasing system size [2]. To overcome this problem Berg and Neuhaus [3] have recently introduced multicanonical simulations which are based on reweighting ideas and can, in principle, be combined with any legitimate update algorithm. Using *local* update algorithms (Metropolis or heat-bath) it has been demonstrated in several applications [4] that the growth of autocorrelation times with system size is reduced to a power law, $\tau \propto V^{\alpha}$ with $\alpha \geq 1$. For the two-dimensional q-state Potts model values of $\alpha \approx 1.3$ have been reported for $q = 7$ [5] and $q = 10$ [6].

Since by construction the multicanonical energy distribution is constant over the interesting energy range, invoking a random walk argument, one would expect an exponent $\alpha = 1$ for an optimally designed update algorithm. Promising candidates are *nonlocal* cluster algorithms [7,8] which proved very successful at continuous phase transitions. A direct implementation, however, turned out to be difficult because the multicanonical energy is effectively long ranged. The purpose of this Letter is to present for Potts models a variant of the multicanonical approach which does take advantage of cluster updates and turns out to be optimal in the above sense. Basically the idea is to treat the cluster flips in the first place and to reweight the bond degrees of freedom instead of the energy.

The basis of cluster update algorithms is the equivalence of the Potts model

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

with the Fortuin-Kastelein (FK) and random cluster (RC) representations [9],

$$Z_{\text{Potts}} = Z_{\text{FK}} = Z_{\text{RC}}, \tag{2}$$

where

$$Z_{\text{FK}} = \sum_{\{\sigma_i\}} \sum_{\{b_{ij}\}} \prod_{\langle ij \rangle} [p \delta_{\sigma_i \sigma_j} \delta_{b_{ij},1} + \delta_{b_{ij},0}], \tag{3}$$

$$Z_{\text{RC}} = \sum_{\{b_{ij}\}} p^{\sum_{\langle ij \rangle} b_{ij}} q^{N_c(\{b_{ij}\})}, \tag{4}$$

with $p = \exp(\beta) - 1$. Here $b_{ij} = 0$ or 1 are bond occupation numbers and $N_c(\{b_{ij}\})$ denotes the number of connected clusters (including one-site clusters.) According to (3) a Swendsen-Wang cluster update sweep then consists of (1) setting $b_{ij} = 0$ if $\sigma_i \neq \sigma_j$, or assigning values 0 and 1 with relative probability 1: p if $\sigma_i = \sigma_j$, (2) identifying clusters of spins that are connected by "active" bonds ($b_{ij} = 1$), and (3) choosing a new random value $1, \ldots, q$ independently for each cluster.

By differentiating $\ln Z$ with respect to β it is easy to see that the average of the energy $E = -\sum_{\langle ij \rangle} \delta_{\sigma_i \sigma_j}$ can be expressed in terms of the average of the number of active bonds $B = \sum_{\langle ij \rangle} b_{ij}$,

$$\frac{\partial \ln Z}{\partial \beta} = -\langle E \rangle = \frac{p+1}{p} \langle B \rangle, \tag{5}$$

and for the specific heat per site C one finds

$$\frac{CV}{\beta^2} = -\frac{\partial \langle E \rangle}{\partial \beta} = -\frac{p+1}{p^2} \langle B \rangle + \left(\frac{p+1}{p} \right)^2 (\langle B^2 \rangle - \langle B \rangle^2). \tag{6}$$

Equation (5) suggests that the bond histogram $P_b(B)$ should develop for $\beta = \beta_t \pm \delta\beta$ a pronounced peak around $B_{o,d} = -[p/(p+1)]E_{o,d}$ and for $\beta \approx \beta_t$ a double-peak structure similar to $P_e(E)$. In fact, as is illustrated in Fig. 1 for $q = 7$ and $L = 60$, a plot of P_b vs $[(p+1)/p]B$ is hardly distinguishable from $P_e(E)$. For other values of q and L the comparison looks very similar.

 0031-9007/95/74(2)/212(4)$06.00

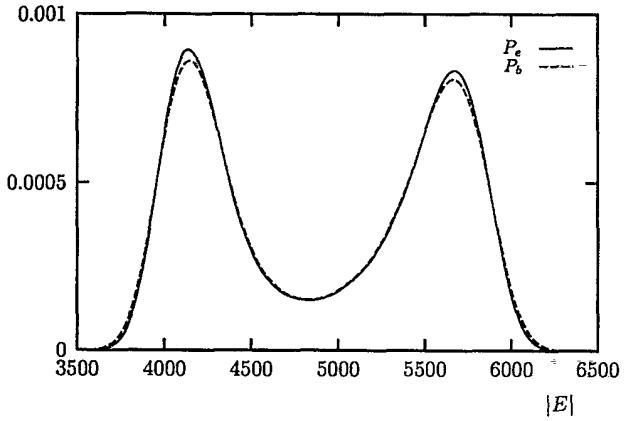

FIG. 1. Canonical energy and bond histograms for $q = 7$, $L = 60$, and $\beta = 1.292\,283$. The bond histogram is plotted vs $[(p + 1)/p]B$, where $p = \exp(\beta) - 1$.

In terms of P_b the slowing down of canonical simulations is thus caused by the strongly suppressed configurations near the minimum between the two peaks, analogous to the well known argument for P_e. To enhance these probabilities we therefore introduce in analogy to multicanonical simulations a "multibondic" partition function

$$Z_{\mathrm{mubo}} = \sum_{\{\sigma_i\}} \sum_{\{b_{ij}\}} \prod_{\langle ij \rangle} [p\,\delta_{\sigma_i\sigma_j}\delta_{b_{ij},1} + \delta_{b_{ij},0}] \exp[-f_b(B)], \tag{7}$$

where $f_b(B) = \ln P_b(B)$ between the two peaks and $f_b(B) = 0$ otherwise. Of course, as in multicanonical simulations, any reasonable approximation of $P_b(B)$ can be used in practice. Canonical expectation values can always be recovered exactly by applying the inverse of the reweighting factor $\exp[-f_b(B)]$.

Obviously, once the b_{ij} are given, we can update the σ_i exactly as in the original Swendsen-Wang cluster algorithm. To update the b_{ij} we proceed as follows. If $\sigma_i \neq \sigma_j$ then the bond b_{ij} is never active, and we always set $b_{ij} = 0$. If $\sigma_i = \sigma_j$ then we define $B' = B - b_{ij}$ and choose new values $b_{ij} = 0$ or 1 with probabilities $\mathcal{P}(b_{ij} = 0) = \mathcal{N} \exp[-f_b(B')]$ and $\mathcal{P}(b_{ij} = 1) = \mathcal{N} p \exp[-f_b(B' + 1)]$, where $\mathcal{N} = 1/\{\exp[-f_b(B')] + p \exp[-f_b(B' + 1)]\}$ is a trivial normalization factor. Since this is nothing but a local heat-bath algorithm for the b_{ij}, the whole procedure is obviously a valid update algorithm.

To evaluate the performance of the multibondic (mubo) cluster algorithm we performed simulations of the Potts model (1) in two dimensions with $q = 7$, 10, and 20. The investigated lattice sizes and simulation temperatures are compiled in Table I. For comparison we run standard multicanonical (muca) simulations using the heat-bath update algorithm with the same parameters. In each run we recorded $N = 100\,000$ measurements of E, B, and two definitions of the magnetization in a time-series file. (The only exception is the multicanonical simulation for $q = 7$, $L = 100$ with $N = 30\,000$.) Between the measurements we performed M lattice sweeps, with M adjusted in such a way that the autocorrelation times in units of measurements and thus the effective statistics of practically uncorrelated data was roughly the same in all simulations.

TABLE I. Simulation parameters: β_0 is the inverse simulation temperature, M_{muca} and M_{mubo} denote the number of lattice sweeps between measurements, and $|E|_{\mathrm{min,max}}$ and $B_{\mathrm{min,max}}$ are the cuts used in the definition of the autocorrelation time τ^{flip}.

q	L	β_0	M_{muca}	$\lvert E\rvert_{\mathrm{min}}$	$\lvert E\rvert_{\mathrm{max}}$	M_{mubo}	B_{min}	B_{max}
	20	1.284 690	10	426	644	10	310	462
7	40	1.291 050	50	1801	2542	30	1310	1840
	60	1.292 283	100	4139	5675	70	3003	4118
	100	1.293 089	100	11 692	15 672	100	8476	11 367
	12	1.407 380	10	120	247	20	91	186
	16	1.415 340	20	222	439	40	166	326
10	20	1.418 864	30	353	676	60	270	512
	26	1.421 642	70	614	1137	140	467	867
	34	1.423 380	200	1065	1938	400	813	1474
	50	1.424 752	200	2349	4185	400	1797	3182
	4	1.577 747	5	6	32	10	5	24
	6	1.639 809	6	17	72	12	14	56
	8	1.665 033	12	33	124	25	28	96
	10	1.676 647	17	56	192	35	45	151
20	12	1.683 517	32	79	280	65	66	215
	14	1.688 195	67	112	357	135	92	295
	16	1.690 278	87	154	470	175	122	384
	18	1.692 013	125	194	593	250	155	485
	20	1.693 698	175	231	728	350	201	595

To make sure that the new multibondic algorithm was implemented correctly, we have analyzed some of the usually considered canonical quantities such as the specific heat C and the Binder parameter $V = 1 - \langle E^4\rangle/3\langle E^2\rangle^2$. In Table II we compare results for the specific-heat maximum and Binder-parameter minimum for $q = 10$ obtained from our multicanonical and multibondic simulations. The error bars are estimated by the jackknife method [10]. The data are in good agreement with each other and also with results from independent canonical Metropolis [11] or single-cluster [12] high-statistics simulations.

As discussed in Refs. [13,14] in the context of a multicanonical multigrid implementation, it is not completely obvious which definition of the autocorrelation time should be used to characterize the dynamics of multicanonical or multibondic simulations. One could, e.g., analyze the (multicanonical or multibondic) autocorrelation function of E, $E\exp[f_e(E)]$, B, or $B\exp[f_b(B)]$, where $f_e(E)$ is the multicanonical analog of $f_b(B)$. More relevant from a practical point of view is the effective autocorrelation time [13,14] for canonical observables which can be defined from the ratio of proper and naive error estimates ($\epsilon/\epsilon_{\mathrm{naiv}} = \sqrt{2\tau^{\mathrm{eff}}}$). The third possibility, which allows a direct comparison with previous work, is to define flipping (or more properly diffusion)

TABLE II. Two-dimensional 10-state Potts model: Comparison of results for specific-heat maxima and Binder-parameter minima from multicanonical (muca) and multibondic (mubo) simulations.

L	alg.	$\beta_{C_{\max}}$	$C_{\max}$	$\beta_{V_{\min}}$	$V_{\min}$
12	muca	1.406 21(28)	44.89(18)	1.392 56(29)	0.503 79(75)
12	mubo	1.407 33(29)	44.92(18)	1.393 73(29)	0.505 21(79)
16	muca	1.414 80(17)	73.68(25)	1.407 57(17)	0.525 03(53)
16	mubo	1.414 51(15)	73.87(27)	1.407 29(15)	0.524 47(61)
20	muca	1.418 37(11)	109.92(32)	1.413 92(11)	0.535 41(44)
20	mubo	1.418 602(95)	109.33(33)	1.414 150(97)	0.536 49(44)
26	muca	1.421 325(71)	177.61(44)	1.418 786(72)	0.544 51(34)
26	mubo	1.421 506(54)	177.91(43)	1.418 969(54)	0.544 53(34)
34	muca	1.423 313(36)	296.13(52)	1.421 873(36)	0.550 01(24)
34	mubo	1.423 312(26)	297.14(50)	1.421 872(26)	0.549 41(23)
50	muca	1.424 801(28)	627.79(97)	1.424 155(34)	0.554 37(21)
50	mubo	1.424 834(21)	627.6(1.1)	1.424 188(21)	0.554 39(22)

times $4\tau_E^{\text{flip}}$ by counting the number of update sweeps that are needed to travel from $E < E_{\min}$ to $E > E_{\max}$ and back. Here $E_{\min,\max}$ (or $B_{\min,\max}$ for τ_B^{flip}) are cuts which are usually chosen as the peak locations $E_{o,d}(L)$ of the canonical probability distribution. Alternatively, one could also use the infinite volume limits $\hat{E}_{o,d}$ of $E_{o,d}(L)$ for all lattice sizes. For 2D Potts models this is straightforward since $\hat{E}_{o,d}$ are known exactly. In our simulations we have tested if E (or B) has passed the cuts after each sweep. We observed significantly larger τ^{flip} when performing this test only every Mth sweeps, since then any cut crossings during the $M-1$ sweeps between the measurements cannot be detected.

Our results for τ_E^{flip} obtained in multicanonical and multibondic simulations for $q = 7$, 10, and 20 are shown in the log-log plots of Figs. 2–4. Here we have used the canonical peak locations for the energy cuts. Let us first concentrate on the results for $q = 7$ in Fig. 2 where we have included for comparison the data from previous multicanonical simulations [5] and also Rummukainen's results for his hybridlike two-step algorithm which combines microcanonical cluster updates with a multicanonical demon refresh [15]. Both cluster update versions show qualitatively the same behavior and, for $L > 20$, perform much better than the standard multicanonical algorithm. From least-squares fits by

$$\tau_E^{\text{flip}} = aV^{\alpha}, \tag{8}$$

we estimate $\alpha \approx 1.3$ for multicanonical heat-bath and $\alpha \approx 1$ for multibondic cluster simulations; see Table III where we also give results for fixed energy cuts. Our results for τ_B^{flip} are almost indistinguishable from τ_E^{flip} which, recalling Fig. 1, is no surprise. Furthermore, we have also measured the effective autocorrelation times and find that they are systematically smaller for both algorithms.

Unfortunately, for $q = 10$ and 20 the situation is less favorable for the multibondic algorithm. While we still find an exponent of $\alpha \approx 1$, the prefactor in (8) turns out to be so large that we can take advantage of this asymptotic

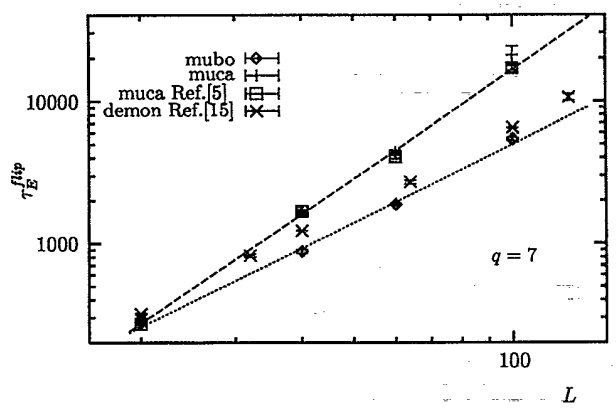

FIG. 2. Log-log plot of autocorrelation times τ_E^{flip} of the energy vs lattice size for $q = 7$, using L-dependent energy cuts defined by the peak locations of the canonical energy distribution.

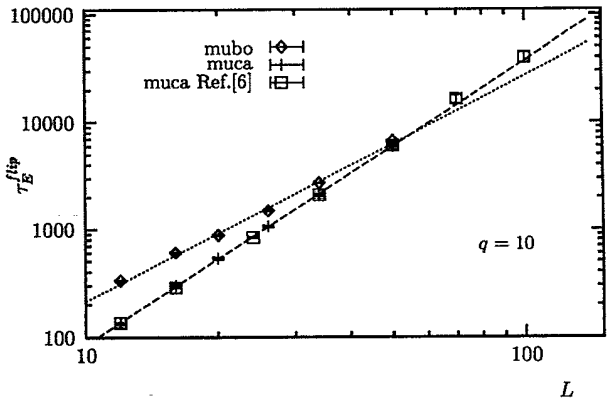

FIG. 3. Same as Fig. 2 for $q = 10$.

TABLE III. Results for the dynamical exponent α in multicanonical (muca) and multibondic (mubo) simulations from fits to $\tau_E^{\text{flip}} = aV^{\alpha}$, using L-dependent cuts defined by the canonical peak locations (α_{max}) and fixed cuts at the infinite volume limits of $E_{d,o}$ and $B_{d,o}$ (α_{fix}).

	α_{max}		α_{fix}	
q	muca	mubo	muca	mubo
7	1.27(2)	0.92(2)	1.53(2)	1.02(2)
10	1.32(2)	1.05(1)	1.43(1)	1.12(1)
20	1.26(1)	1.09(1)	1.46(1)	1.18(1)

improvement only for very large lattice sizes. As can be seen in Fig. 3, for $q = 10$ the crossover happens around $L = 50$. Extrapolating to $L = 100$, we estimate that the multibondic algorithm would perform for this lattice size about 1.5 times faster than the standard multicanonical heat bath. For $q = 20$ the same comparison clearly favors the standard algorithm for all reasonable lattice sizes—and we certainly cannot recommend the new algorithm for large q.

In summary, we have proposed for Potts models a combination of cluster update techniques with reweighting in the random bond representation and have shown that this approach is feasible in practice. In fact, it is technically not more involved than the standard multicanonical approach, and one lattice sweep takes about the same CPU time. Numerical tests for the two-dimensional q-state Potts model with $q = 7$, 10, and 20 show that the multibondic cluster algorithm is optimal in the sense that the exponent α in the power law, $\tau^{\text{flip}} = aV^{\alpha}$, is consistent with $\alpha = 1$, the value one would expect in an idealized random walk picture. For $q = 7$ the multibondic algorithm clearly outperforms the standard multicanonical heat-bath algorithm. Compared with Rummukainen's hybridlike two-step cluster variant, the multibondic autocorrelation times are smaller for all lattice sizes by a roughly constant factor of 1.5. For larger values of q, however, the prefactor a turns out to be relatively large, rendering the new algorithm for reasonable lattice sizes more efficient than multicanonical simulations only for $q < q_0$ with q_0 somewhat over 10.

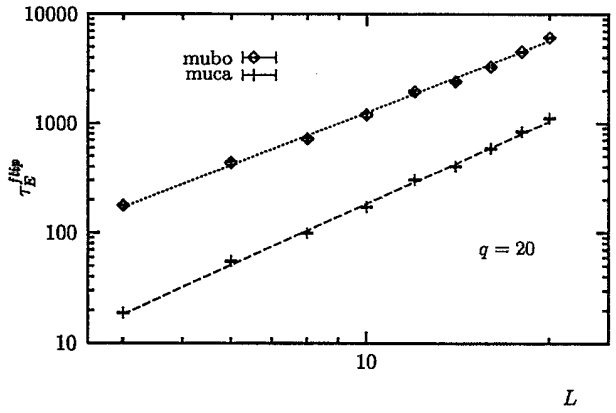

FIG. 4. Same as Fig. 2 for $q = 20$.

The multibondic cluster algorithm may be of value for a wide range of investigations, since it can be applied to any systems where conventional cluster update techniques are applicable.

W. J. would like to thank the DFG for a Heisenberg fellowship, and S. K. gratefully acknowledges a fellowship by the Graduiertenkolleg "Physik und Chemie supramolekularer Systeme." The Monte Carlo simulations were performed on the CRAY Y-MP of the Höchstleistungsrechenzentrum Jülich, the CRAY Y-MP's of the Norddeutscher Rechnerverbund in Kiel and Berlin under Grant bvpf03, and on the Mainz cluster of fast RISC workstations.

[1] J. D. Gunton, M. S. Miguel, and P. S. Sahni, in *Phase Transitions and Critical Phenomena,* edited by C. Domb and J. L. Lebowitz (Academic Press, New York, 1983), Vol. 8; K. Binder, Rep. Prog. Phys. **50**, 783 (1987); *Dynamics of First Order Phase Transitions,* edited by H. J. Herrmann, W. Janke, and F. Karsch (World Scientific, Singapore, 1992).

[2] A. Billoire, R. Lacaze, A. Morel, S. Gupta, A. Irbäck, and B. Petersson, Nucl. Phys. **B358**, 231 (1991).

[3] B. A. Berg and T. Neuhaus, Phys. Lett. B **267**, 249 (1991). For earlier related ideas, see G. M. Torrie and J. P. Valleau, Chem. Phys. Lett. **28**, 578 (1974); J. Comp. Phys. **23**, 187 (1977).

[4] For a list of references see, e.g., W. Janke, *Computer Simulations in Condensed Matter Physics VII,* edited by D. P. Landau, K. K. Mon, and H. B. Schüttler (Springer Verlag, Heidelberg, Berlin, 1994).

[5] W. Janke, B. A. Berg, and M. Katoot, Nucl. Phys. **B382**, 649 (1992).

[6] B. A. Berg and T. Neuhaus, Phys. Rev. Lett. **68**, 9 (1992).

[7] R. H. Swendsen and J. S. Wang, Phys. Rev. Lett. **58**, 86 (1987).

[8] U. Wolff, Phys. Rev. Lett. **62**, 361 (1989).

[9] F. Y. Wu, Rev. Mod. Phys. **54**, 235 (1982); **55**, 315(E) (1983).

[10] R. G. Miller, Biometrika **61**, 1 (1974); B. Efron, *The Jackknife, the Bootstrap and other Resampling Plans* (SIAM, Philadelphia, PA, 1982).

[11] A. Billoire, R. Lacaze, and A. Morel, Nucl. Phys. **B370**, 773 (1992).

[12] W. Janke (unpublished).

[13] W. Janke and T. Sauer, Phys. Rev. E **49**, 3475 (1994); and in *Proceedings of the Fourth International Conference Path Integrals from meV to MeV: Tutzing, 1992,* edited by H. Grabert, A. Inomata, L. S. Schulman, and U. Weiss (World Scientific, Singapore, 1993), p. 17.

[14] W. Janke and T. Sauer, FU Berlin Report No. FUB-HEP 15/93 [J. Stat. Phys. (to be published)].

[15] K. Rummukainen, Nucl. Phys. **B390**, 621 (1993).