ELSEVIER

Computer Physics Communications 147 (2002) 58–63

Computer Physics Communications

www.elsevier.com/locate/cpc

Cluster Monte Carlo: Extending the range

H.W.J. Blöte [a,b,*], J.R. Heringa [a], E. Luijten [c]

[a] *Laboratory of Applied Physics, Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands*
[b] *Instituut Lorentz, Leiden University, Niels Bohrweg 2, P.O. Box 9506, 2300 RA Leiden, The Netherlands*
[c] *Institute for Physical Science & Technology, University of Maryland, College Park, MD 20742-2431, USA*

Abstract

Monte Carlo simulations with local updates tend to become time-consuming when large-scale correlations exist, such as in critical systems. For a limited, but increasing number of model systems, nonlocal 'cluster' algorithms are available that are orders of magnitude more efficient than algorithms with local updates. Cluster algorithms can be defined on the basis of the symmetry properties of the Hamiltonian; different symmetries can thus lead to different cluster algorithms. We review a number of existing cluster algorithms, and describe new ones for an Ising-like model with two- and three-spin interactions, and for the chiral Potts model. New simulation data for the Ising-like model allow an accurate determination of its specific-heat exponent; this result confirms existing ideas that the model belongs to the 4-state Potts universality class.

PACS: 02.50.Ng; 02.70.Lq; 05.50.+q

Keywords: Monte Carlo methods; Critical phenomena

1. Introduction

Monte Carlo simulations of many-particle systems frequently use a Metropolis-type update mechanism [1]. Such algorithms produce local changes: each configuration update is restricted to a small neighborhood. Near criticality, such algorithms suffer from the critical-slowing-down phenomenon. This effect is described by the dynamical exponent z; the autocorrelation time τ_L (expressed in updates *per particle*) of a critical system of linear size L satisfies the following scaling equation

$$\tau_L \propto L^z. \tag{1}$$

Models with short-range interactions in d dimensions require roughly L^d operations to update every particle in the system. The generation of a statistically independent state thus involves a number of the order of L^{d+z} operations. The value of z still depends on the universality class of the model and the simulation dynamics, but one often finds $z \approx 2$; for instance, z is close to $13/6$ for the Ising model in two dimensions [2]. As a consequence, the required simulation time increases rapidly with the system size L. The critical-slowing-down phenomenon thus poses a barrier that restricts the computation of statistically accurate data to small systems.

A remarkable breakthrough was achieved by Swendsen and Wang [3], who formulated an algorithm for the Potts model (which includes the Ising model as a special case) that produces 'nonlocal' updates: in-

* Corresponding author.
E-mail address: bloete@tnw.tudelft.nl (H.W.J. Blöte).

0010-4655/02/$ – see front matter
PII: S0010-4655(02)00204-7

stead of single spins, whole clusters of spins are simultaneously flipped. The sizes of these clusters are essentially random, and can be large. Thus large steps in the configuration space can now be made, and as a consequence, the dynamic exponent z decreases considerably [3].

The validity of this cluster algorithm can be explained on the basis of the Kasteleyn–Fortuin random-cluster decomposition [4] of the Potts model. This is a probabilistic process that divides the spins into groups called 'random clusters'. While all spins in a random cluster have the same sign, correlations between spins in different clusters are zero. Thus, one has the freedom to randomly and independently assign a new spin value to each cluster. This is precisely what happens in the Swendsen–Wang algorithm. Even more efficient varieties of this cluster algorithm have since then been developed [5,6]. The Wolff single-cluster algorithm [6] stands out because of its simplicity: only one cluster is formed and flipped at a time.

Several generalizations of these algorithms have been derived. The Creutz formulation [7] of cluster algorithms in terms of demons may be helpful, for instance for the purpose of handling antiferromagnetic interactions. In models which use vector spins instead of discrete spins, such as the Heisenberg and XY models, one can still single out one Cartesian spin component and treat it as an Ising variable [8]. It has also been demonstrated that multispin interactions can, at least in some cases, be handled by a cluster algorithm [9].

However, it is clear that these highly efficient non-local algorithms are not as easy to generalize as local (Metropolis-type) algorithms, and are thus restricted to a limited range of applicability. The efficiency of a cluster algorithm is obviously related to cluster-size distribution. A possible pitfall of newly devised cluster algorithms is the situation that the clusters tend to occupy practically the whole system, resulting only in trivial changes of the spin configuration, and in a limited efficiency. This situation naturally arises when competing interactions are present. For optimal efficiency, the percolation threshold of the cluster-formation process should coincide with the critical point. In a number of cases, such as the Swendsen–Wang algorithm and the geometric cluster algorithm (see Section 4) applied to the Potts model, this coincidence can be proven. In other cases it is clear that the critical point lies well within the region where the cluster-formation process achieves percolation, away from the percolation threshold. Thus the clusters tend to be too large for maximum efficiency. As demonstrated below (Section 3), even then a cluster algorithm may prove to be quite useful.

After this brief (and necessarily incomplete) introductory review of cluster algorithms, we use the following sections to summarize some of our own contributions in this field. We include new algorithms and simulation results.

2. Long-range interactions

In systems where each particle interacts with every other particle, the evaluation of the probability of a local configuration change normally requires of the order of L^d operations, so that the time needed per independent critical configuration is of the order of L^{2d+z}. This steep L-dependence, which is even worse than in the case of short-range interactions, has greatly restricted the simulation of critical systems with long-range forces.

A few years ago, a cluster algorithm became available for spin systems with long-range interactions [10]. Remarkably, it uses only roughly of the order of L^d operations per independent configuration. Its applications [11–14] include the one-dimensional Ising chain with interactions decaying algebraically as a function of the distance r_{ij}:

$$\mathcal{H}/k_{\mathrm{B}}T = -\sum_{ij} J r_{ij}^{-(d+\sigma)} s_i s_j, \tag{2}$$

where i and j label the lattice sites of spins s_i and s_j, and the sum counts each spin pair once. The spins can assume the values ± 1. This model functions as a prototype for the study of the dependence of critical properties on the range of the interactions, and several of its properties are known analytically. But most of our knowledge relies on a renormalization-group analysis [15], and actual numerical results are scarce. Simulations have been performed for systems of up to 150,000 spins, which is two to three orders of magnitude larger than in other studies of this model. From a finite-size scaling analysis, a number of critical properties were determined for several values of $0 < \sigma < 1$. Among these properties is the

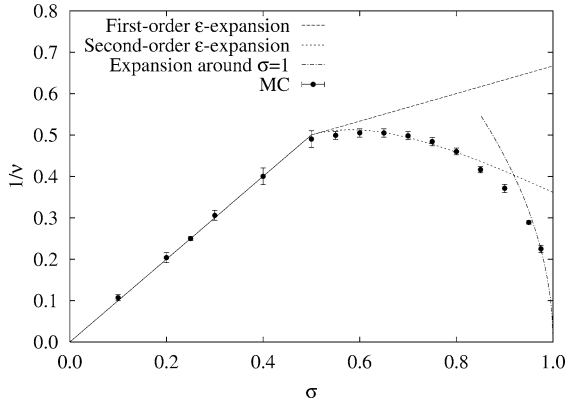

Fig. 1. The inverse correlation-length exponent $1/\nu$ for the one-dimensional Ising chain with interactions decaying as $r^{-(d+\sigma)}$. For a description see the text.

correlation-length exponent ν (upon approach of $T \to T_c$, the spin–spin correlation length ξ diverges as $|T - T_c|^{-\nu}$), which is depicted in Fig. 1. The accuracy of the data allows a detailed comparison to the analytical results: for $\sigma \leqslant \frac{1}{2}$, the system is predicted to exhibit classical critical behavior, albeit with a σ-dependent correlation exponent $\nu = 1/\sigma$. For $\sigma > \frac{1}{2}$, our knowledge relies on an expansion in $\varepsilon \equiv 2\sigma - 1$. The interesting situation arises that (unlike for the usual short-range interactions) a detailed verification of the ε-expansion becomes possible for $\varepsilon \ll 1$. Indeed, our numerical results compare favorably to the second-order expansion up to $\sigma \approx 0.8$ ($\varepsilon \approx 0.6$), beyond which the crossover to the Kosterlitz–Thouless transition at $\sigma = 1$ [16] takes place, as indicated by our results for σ close to unity, which nicely approach the lowest-order expansion $\nu = 1/\sqrt{2(1-\sigma)}$ [17].

The long-range cluster algorithm is well applicable to systems in $d > 1$ dimensions [10]. In particular, it has been used to investigate the universal behavior in the range of σ where deviations from short-range criticality first appear [14], and the crossover between classical and short-range critical behavior [18].

3. Multispin interactions

In the presence of multispin interactions, the formulation of a cluster algorithm is less straightforward than in the case of pair interactions, and the feasibility depends on the precise form of the Hamiltonian. The Baxter–Wu model [19] contains three-spin interactions in all elementary faces of the triangular lattice. One may thus single out one of the three sublattices and consider its spins as fixed; then only pair interactions remain between spins in the other two sublattices. Each of these interactions may be ferromagnetic, zero or antiferromagnetic, depending on the signs of two spins on the third sublattice. In spite of the presence of antiferromagnetic interactions, there is no 'frustration': whenever non-zero interactions form a loop, the loop contains an even number of antiferromagnetic interactions. Under these conditions an efficient cluster algorithm is well possible [9].

The Baxter–Wu model can be generalized in at least two ways such that a cluster algorithm remains possible. First, one can assign different couplings K_1 and K_2 to the up and down triangles respectively. The model is self-dual [20]; the self-dual line is given by $\sinh(2K_1)\sinh(2K_2) = 1$. Simulations of this model [21] show that the self-dual line is critical, and that the leading scaling behavior is that of the 4-state Potts model. However, the logarithmic correction factors (which typically appear in this class of models) seem to depend on the ratio K_1/K_2. While for $K_1/K_2 = 1$ logarithmic factors are absent [19], slow finite-size convergence suggests that they reappear at other ratios [21]. Second, the model can be generalized [22] to q-state Potts spins with $q > 2$ while its self-duality is preserved. We have simulated this model using a suitably generalized cluster algorithm [21], and found that for $q = 3$ and $q = 4$ it undergoes a first-order phase transition at the self-dual point. For $q > 2$ the cluster algorithm appears to be less effective than for $q = 2$; in the absence of a critical state the cluster formation process resides no longer at the percolation threshold.

In addition, we have constructed a cluster algorithm for an Ising-like model with two- and three-spin interactions described by

$$\mathcal{H}/k_B T = -K \sum_{x,y} (s_{x,y}s_{x,y+1} + s_{x,y}s_{x+1,y}s_{x+2,y}), \tag{3}$$

where the spin $s_{x,y}$ can assume the values ± 1, and x, y are the coordinates of its lattice site. This model is critical at $K = \log(1+\sqrt{2})/2$ [20,23]. It has four equivalent ordered phases, and may be expected to belong to the four-state Potts universality class. Long

Table 1
Monte Carlo results for the dimensionless specific heat C/k_B of $L \times L$ Ising systems with two- and three-spin interactions. For comparison we include two old results obtained by a special-purpose computer [23]. The numbers of steps of the new simulations are given as the numbers of flipped clusters (c) and those of the old simulations as the numbers of Metropolis sweeps (s). The number of visits per spin in both types of simulations is roughly the same (the average fractional cluster size is about 1/3 for the largest systems), but the statistical accuracy (given in parentheses) of the cluster method is better for similar system sizes

L	# steps	C/k_B		L	# steps	C/k_B	
6	5.5×10^8 c	1.1043	(2)	96	5.5×10^8 c	9.689	(6)
9	5.5×10^8 c	1.5906	(3)	144	5.5×10^8 c	13.16	(1)
12	5.5×10^8 c	2.0165	(4)	192	5.5×10^8 c	16.40	(1)
18	5.5×10^8 c	2.7686	(7)	288	5.5×10^8 c	22.41	(2)
24	5.5×10^8 c	3.442	(1)	384	5.5×10^8 c	27.98	(3)
36	5.5×10^8 c	4.665	(2)	576	2.0×10^8 c	38.6	(1)
48	5.5×10^8 c	5.776	(2)	64	2.0×10^8 s	9.016	(22)
72	5.5×10^8 c	7.810	(4)	128	3.0×10^8 s	15.23	(7)

simulations using a local algorithm on a special-purpose computer [23] indeed were consistent with the four-state Potts universality class, but the exponents could not be well determined. The accuracy was not only limited by the critical slowing-down, but also by an anomalously slow finite-size convergence.

By 'freezing' one randomly chosen sublattice according to $x \bmod 3 = 0$, 1 or 2, the three-spin interactions are reduced to pair interactions, so that just as for the Baxter–Wu model, a Wolff-like cluster simulation becomes possible. However, unlike the case of the Baxter–Wu model, frustrated loops may occur, and indeed we find that the percolation threshold of the clusters does not coincide with the critical point. The useful changes produced by a cluster move are thus essentially local in character. However, the scale of these changes is much larger than a single lattice unit, so that the algorithm is still considerably more efficient than the Metropolis method. New results for the specific heat C are shown in Table 1.

Finite-size scaling for critical four-state Potts-like models predicts

$$C(L) \simeq C_0 + L^{2y_t-2}\left[a_1(\log L)^{-1/3} + \cdots\right], \qquad (4)$$

where $y_t = 2/(2-\alpha)$ is the temperature exponent and α describes the specific-heat divergence at the critical point by $C \propto |T - T_c|^{-\alpha}$. The new results indicate that these terms are adequate for $L \geqslant 48$ (at least for system sizes equal to multiples of 3; additional corrections to scaling occur for other sizes). A least-squares fit of this formula including y_t as a free parameter yields $y_t = 1.500 \pm 0.001$. Thus, the specific-heat exponent is very close to the expected value $\alpha = 2/3$ which applies to the 4-state Potts universality class.

4. Geometric clusters

In general a cluster algorithm can be formulated on the basis of a symmetry property of the model, i.e. the Hamiltonian should be invariant under that symmetry. Moreover, the symmetry operation should be self-inverse. These two conditions are sufficient to prove detailed balance [24]. The Swendsen–Wang algorithm [3] is thus based on the q-state permutation symmetry of the Potts model (up-down symmetry in the Ising case). A new type of cluster algorithm can be devised on the basis of a geometric symmetry, for instance, the inversion symmetry of the lattice.

Such geometric transformations were used by Dress and Krauth [25] for the simulation of hard particles in continuous space. Here the percolation threshold does not coincide with the phase transition; in the case of the simple-cubic hard-core lattice gas, it does [26], and the critical-slowing-down phenomenon is strongly suppressed.

The geometric cluster algorithm can also be applied in the presence of finite spin–spin interactions. For the case of the Potts model, it has been shown [24] that the percolation threshold of the geometric clusters coincides with the critical point. Empirical evidence [24]

indicates that this situation also occurs for the tricritical Blume–Capel model in three dimensions. Since the number of spins with a certain value is not changed by a geometric cluster step, it is possible to sample, e.g., the fixed-magnetization ensemble.

A variant of the geometric cluster algorithm is suitable for application to the q-state chiral Potts model. The three-state model is described by

$$\mathcal{H}/k_{\mathrm{B}}T = -K \sum_{\langle ij \rangle} \cos\left[\frac{2\pi}{3}(\sigma_i - \sigma_j + \vec{a}\cdot\vec{r}_{ij})\right], \quad (5)$$

where the spin σ_k can assume the values 1, 2 or 3, k labels its lattice site and $\vec{r}_{ij}$ is the relative position of spins σ_i and σ_j. The sum is on all nearest-neighbor pairs $\langle ij \rangle$. For non-zero chirality parameter $\vec{a}$ the model is not invariant under spin permutations and lattice inversions. However, a lattice inversion combined with an inversion of the q Potts states $\sigma \to s - \sigma \mod q$ (where s is an arbitrary integer) leaves the Hamiltonian invariant, and this provides the basis for a nonlocal Monte Carlo algorithm. Due to the inversion of the Potts states, the algorithm no longer conserves the magnetization. Although chirality plays no role in the Ising case $q = 2$, the algorithm is still applicable and thus provides another different way to efficiently sample the canonical distribution of the Ising model. Test calculations on the $q = 3$ chiral Potts model indicate that the new algorithm is quite effective for small chirality $\vec{a}$. For larger values it is clear that the critical point and the percolation threshold of the clusters no longer coincide.

5. Concluding remarks

For a limited group of model systems, nonlocal 'cluster' algorithms enable a much more efficient way of simulation than local algorithms. For long-range models the gain can be (see Section 2) of the order of L^{d+z}. It reaches a value of about 10^8 for the largest systems that have been simulated. The occurrence of d in the exponent arises because it can be arranged such that the algorithm executes an amount of work per spin s_i that is *not* proportional to the number of interacting neighbors, but to the integrated coupling $\sum_j K_{ij}$ in which that spin participates. As a consequence of the existence of the thermodynamic limit of the energy per spin, this amount of work remains finite in ferromagnetic models, even if the number of interacting neighbors diverges. In systems with antiferromagnetic interactions and 'frustration', such as occur in the case of dipolar interactions, and such as used in models for spin glasses and neural networks, the percolation threshold of the clusters moves away from the critical point. This reduces the efficiency by a factor of the order of L^z. However, even then a large factor (possibly of the order of L^d) may remain and thus provide a sufficient reason to apply the long-range cluster algorithm.

The geometric cluster algorithm described above uses a transformation that maps the lattice onto itself. However, it is equally well possible to exchange spins (or other sorts of particles) between disjoint systems [27]. Depending on the precise form of the Hamiltonian, there may be a considerable freedom in the choice of the symmetry operations defining the cluster formation process. In common cases the lattice model is invariant under discrete translations (if the boundary conditions are periodic) and rotations. The translations should be over half the system size, and the rotations over π: they are subject to the condition that the transformations must be self-inverse. However, arbitrary translations may be included in other symmetry operations such as a geometric inversion of the lattice (a translation followed by an inversion is another inversion; the combined operation is thus self-inverse). Moreover, invariance under geometric inversions of the lattice may apply to each Cartesian axis independently. Finally, as described above for the chiral Potts model, the symmetry properties of the lattice may be combined with those of the lattice variables. Possibilities to formulate a cluster algorithm that is especially tailored to the symmetry of the model under investigation may thus arise.

References

[1] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, E. Teller, J. Chem. Phys. 21 (1953) 1087.
[2] M.P. Nightingale, H.W.J. Blöte, Phys. Rev. B 62 (2000) 1089.
[3] R.H. Swendsen, J.-S. Wang, Phys. Rev. Lett. 58 (1987) 86.
[4] P.W. Kasteleyn, C.M. Fortuin, J. Phys. Soc. Jpn. (Suppl.) 26 (11) (1969).
[5] C.F. Baillie, P.D. Coddington, Phys. Rev. B 43 (1991) 10 617.
[6] U. Wolff, Phys. Rev. Lett. 62 (1989) 361.
[7] M. Creutz, Phys. Rev. Lett. 69 (1992) 1002.
[8] U. Wolff, Phys. Rev. Lett. 60 (1988) 1461.

[9] M.A. Novotny, H.G. Evertz, in: D.P. Landau, K.K. Mon, H.-B. Schüttler (Eds.), Computer Simulation Studies in Condensed-matter Physics, Vol. VI, Springer, Berlin, 1993, p. 188.
[10] E. Luijten, H.W.J. Blöte, Int. J. Mod. Phys. C 6 (1995) 359.
[11] E. Luijten, H.W.J. Blöte, Phys. Rev. Lett. 76 (1996) 1557.
[12] E. Luijten, H.W.J. Blöte, Phys. Rev. B 56 (1997) 8945.
[13] E. Luijten, Interaction Range, Universality and the Upper Critical Dimension, Delft University Press, Delft, 1997.
[14] E. Luijten, H.W.J. Blöte, 2001, to be published.
[15] M.E. Fisher, S.-k. Ma, B.G. Nickel, Phys. Rev. Lett. 29 (1972) 917.
[16] E. Luijten, H. Meßingfeld, Phys. Rev. Lett. 86 (2001) 5035.
[17] J.M. Kosterlitz, Phys. Rev. Lett. 37 (1976) 1577.
[18] E. Luijten, H.W.J. Blöte, K. Binder, Phys. Rev. E 54 (1996) 4626.
[19] R.J. Baxter, F.Y. Wu, Phys. Rev. Lett. 31 (1973) 1294; Aust. J. Phys. 27 (1974) 357.
[20] C. Gruber, A. Hintermann, D. Merlini, Group Analysis of Classical Lattice Systems, Springer, Berlin, 1977.
[21] J.R. Heringa, 2001, unpublished.
[22] B. Nienhuis, 2001, unpublished.
[23] H.W.J. Blöte, A. Compagner, P.A.M. Cornelissen, A. Hoogland, F. Mallezie, C. Vanderzande, Physica A 139 (1986) 395.
[24] J.R. Heringa, H.W.J. Blöte, Phys. Rev. E 57 (1998) 4976.
[25] C. Dress, W. Krauth, J. Phys. A 28 (1995) L597.
[26] J.R. Heringa, H.W.J. Blöte, Physica A 232 (1996) 369.
[27] O. Redner, J. Machta, L. Chayes, Phys. Rev. E 58 (1998) 2749; L. Chayes, J. Machta, O. Redner, J. Stat. Phys. 93 (1998) 17.