

Cluster Algorithms: Beyond Suppression of Critical Slowing Down

Erik Luijten¹ and Jiwen Liu

*Department of Materials Science and Engineering,
University of Illinois at Urbana-Champaign, 1304 West Green Street, Urbana, Illinois 61801*

Abstract. The cluster algorithm pioneered by Swendsen and Wang is widely acclaimed for its ability to suppress dynamic slowing down near a critical point. However, the cluster approach permits the formulation of Monte Carlo algorithms that yield important *additional* efficiency gains. For systems with long-range interactions, Luijten and Blöte have introduced a method in which the number of operations per spin flip is independent of the number of interactions between a spin and the other spins in the system. Thus, the computational effort for the simulation of an N -particle system is reduced from $\mathcal{O}(N^2)$ to $\mathcal{O}(N)$, which has helped to resolve several open questions concerning critical behavior in systems with long-range interactions. As a second example of what can be achieved with cluster methods, we discuss some illustrative properties of a newly-developed geometric cluster algorithm for interacting fluids.

INTRODUCTION

The first cluster Monte Carlo algorithm was introduced by Swendsen and Wang over 15 years ago [1] and has had a large impact on the study of critical phenomena in lattice spin models. Conventional, Metropolis-type algorithms suffer from dynamic slowing down near the critical point: the autocorrelation time diverges as a power-law with increasing system size. Thus, the computing time required to generate an *independent* configuration increases superlinearly with the system volume and it becomes prohibitively difficult to obtain accurate data for large system sizes. Since numerical results over an appreciable range of system sizes are required for an accurate finite-size scaling analysis of critical phenomena, this behavior has proven to be a limiting factor (cf. Ref. [2] for a more detailed discussion). The Swendsen–Wang algorithm features *nonlocal* spin updates that lead to a rapid decorrelation of spin configurations and consequently to a strong suppression of critical slowing down. Wolff’s single-cluster implementation [3], which is now the most widely used variant because of its particular simplicity, improves the situation even further. It is important to stress that these cluster algorithms go beyond a mere collective update of a group of spins: such updates will typically lead to exponentially small acceptance rates. By contrast, the methods discussed here rely on the Fortuin–Kasteleyn mapping of the Potts model on the random cluster model [4, 5], which relates the Potts Hamiltonian to a sum over *independent* clusters of spins. Thus, upon decomposition of a spin configuration into appropriate clusters, an arbitrary spin

¹ Corresponding author. E-mail: luijten@uiuc.edu

state can be assigned to each cluster in a *rejection-free* scheme.

In recent years, it has transpired that cluster Monte Carlo algorithms offer specific advantages that go beyond the elimination of the problem of critical slowing down. In particular, it has been shown [6] that, for ferromagnetic interactions, the cluster construction process can be formulated such that it becomes *independent* of the number of interactions per spin, making it particularly efficient for systems with long-range interactions. Indeed, for systems with power-law interactions, where each of the N spins in a system interacts with all other spins, a conventional algorithm requires an $\mathcal{O}(N)$ effort to update a single spin. The long-range cluster algorithm reduces this effort to $\mathcal{O}(1)$, making it as efficient as algorithms for short-range interactions. In this sense, the performance gain is comparable to what has been achieved by particle mesh Ewald [7] and fast-multipole methods [8, 9] in the case of electrostatic interactions. Since the suppression of critical slowing leads to an additional efficiency improvement $\mathcal{O}(L^z)$, where z is the dynamic critical exponent, the total speed-up compared to Metropolis-type algorithms amounts to a factor $N \cdot L^z = L^{d+z}$ at criticality, which is as large as 10^8 for the largest systems studied. In this paper, we briefly discuss the basic ideas underlying the long-range cluster algorithm.

In a rather different development, cluster methods have also been able to overcome long-standing computational hurdles in off-lattice fluids. We have generalized the *geometric* cluster algorithm of Dress and Krauth [10] to interacting fluids and demonstrated that it is capable of achieving very significant performance increases for various classes of systems. Some illustrative examples are provided below.

CLUSTER METHOD FOR LONG-RANGE INTERACTIONS

Consider a regular short-range Ising model, featuring a d -dimensional lattice structure with a spin $s_i = \pm 1$ on each lattice site i . Nearest neighbors interact via a ferromagnetic coupling K . Schematically, Wolff's version of the cluster algorithm works as follows:

1. Randomly choose a spin s_i from the lattice. This spin becomes the first member of a cluster.
2. Consider each spin that interacts with this spin s_i . It is added to the cluster with a probability $p = 1 - \exp(-2K)$, provided that it has the same sign as s_i .
3. Repeat step 2 in turn for each spin that is newly added to the cluster, where one now considers all spins that interact with this spin, rather than with s_i . This is iterated until all neighbors of all spins in the cluster have been considered for inclusion.

Upon completion of this process, all spins that are part of the cluster are inverted, and the next cluster is constructed. This mechanism can be directly generalized to systems with an arbitrary number of different interaction types per spin, cf. Ref. [11]. In this case, step 2 above must be repeated for *each* spin s_j that interacts with s_i , and p depends on the coupling strength K_{ij} . This allows application of a cluster algorithm to systems with ferromagnetic long-range interactions. However, in such systems K will typically be very small, and hence $p \approx 2K$. The magnitude of p can be estimated through a mean-field approximation, in which the critical coupling K_c satisfies $zK_c = 1$, where z is the

coordination number. For long-range interactions in a system containing N spins, this becomes $NK_c = 1$ or $p \approx 2/N$. Thus, $\mathcal{O}(N)$ operations are required to add a single spin to the cluster, just as in a Metropolis-type scheme.

Now, the important observation is made that the cluster construction process outlined here can be reformulated in such a way that each operation leads to a spin that is actually added to the cluster. This will be illustrated here for a one-dimensional spin chain in which all spins interact via a distance-dependent coupling $K_{ij} = K(|i - j|)$. Upon random selection of a starting spin s_i , all other spins in the system are added to the cluster with a probability $p(s_i, s_j) = \delta_{s_i s_j} p_{ij}$, where $p_{ij} = 1 - \exp[-2K(|i - j|)]$ and the Kronecker delta asserts that the spins have the same sign. For each spin that is actually added to the cluster, its address is also placed on the *stack*. When all spins interacting with the first one have been considered, a new spin is read from the stack and the process is reiterated until the stack is empty. The spin from which we are currently adding spins is called the *current spin*. In order to avoid testing each single spin for inclusion in the cluster, we first consider the *provisional* probability p_{ij} appearing in $p(s_i, s_j)$ and introduce the concept of the *cumulative probability* $C(j)$,

$$C(j) \equiv \sum_{n=1}^j P(n) \quad (1)$$

with

$$P(n) = \left[\prod_{m=1}^{n-1} (1 - p_m) \right] p_n . \quad (2)$$

$p_j \equiv 1 - \exp(-2K_j)$ is an abbreviation for p_{0j} (and $K_j \equiv K_{0j}$), i.e., we define the origin at the position of the current spin. $P(n)$ is the probability that, starting from the current spin, $n - 1$ spins are skipped and the n th spin is added, provided that it has the same sign as the current spin. Thus, the next spin j that is provisionally added can be determined from the cumulative probability by means of a single random number $g \in [0, 1)$: $j - 1$ spins are skipped if $C(j - 1) \leq g < C(j)$. If the j th spin indeed has the same sign as the current spin then s_j is added to the cluster. Subsequently, again a number of spins is skipped before the spin at a distance $k > j$ is provisionally added. Owing to the condition $k > j$, the function P must be shifted,

$$P_j(k) = \left[\prod_{m=j+1}^{k-1} (1 - p_m) \right] p_k , \quad (3)$$

and Eq. (2) is simply a special case of Eq. (3). The corresponding cumulative probability is given by a generalization of Eq. (1),

$$C_j(k) = \sum_{n=j+1}^k P_j(n) . \quad (4)$$

By using the specific form of the probability p_{ij} one finds that this reduces to

$$C_j(k) = 1 - \exp\left(-2 \sum_{n=j+1}^k K_n\right). \quad (5)$$

Thus, the probability that the next spin that will be added lies at a distance in the range $[j+1, k]$ is given by an expression that has the same form as the original probability, in which the coupling constant is replaced by the sum of all the couplings with the spins in this range! There are various ways to exploit this property [6]. In essence, $C_j(k)$ is equated to a random number and Eq. (5) is solved for k . Thus, each random number leads to a spin that is actually added to the cluster (provided it has the correct sign). See also Ref. [12] for further technical details.

The algorithm described here has been applied to address a variety of questions pertaining to the critical behavior of systems with long-range interactions, which were hitherto essentially inaccessible to Monte Carlo methods. These include systems with algebraically decaying interactions [13], where the upper critical dimension, separating classical from non-classical critical behavior, is a function of the decay rate of the interactions. This made it possible to resolve a long-standing controversy regarding the nature of finite-size scaling above the upper critical dimension [14]. Another example concerns the study of crossover phenomena, which are relevant in critical fluids and in the demixing behavior of polymer blends. Crossover scaling functions have been obtained through application of the long-range cluster algorithm [15, 16, 17], permitting a reanalysis of experimental data [18] and a stringent test of analytical theories [19]. In addition, the occurrence of a Kosterlitz–Thouless transition in one-dimensional systems has been demonstrated [20] and an old controversy regarding the boundary between long-range and short-range criticality has been resolved [21]. Further works have applied and extended the algorithm to q -state Potts chains [22] and spin layers with dipolar interactions [23].

GEOMETRIC CLUSTER ALGORITHMS

A radically different cluster approach is based upon the identification of clusters via a geometric operation, as proposed by Dress and Krauth for hard-sphere fluids [10]. In this method, a particle configuration is rotated over an angle π around an arbitrary pivot and then overlaid with the original configuration. Overlapping spheres lead to clusters of particles, which are exchanged with their counterparts at the opposite side of the pivotal point. The non-local character of the particle moves helps in overcoming so-called jamming problems that plague simulations of liquids containing particles with different sizes. If the size asymmetry becomes large, the intrinsic time scales of the different constituents start to differ widely and the larger species move prohibitively slowly compared to the smaller species. This profound problem affects Monte Carlo and molecular dynamics simulations alike and has essentially prevented the study of collective phenomena in systems with a size asymmetry (measured in terms of particle diameter ratio) larger than 10. Although not hindered by large size asymmetries, the method of Dress

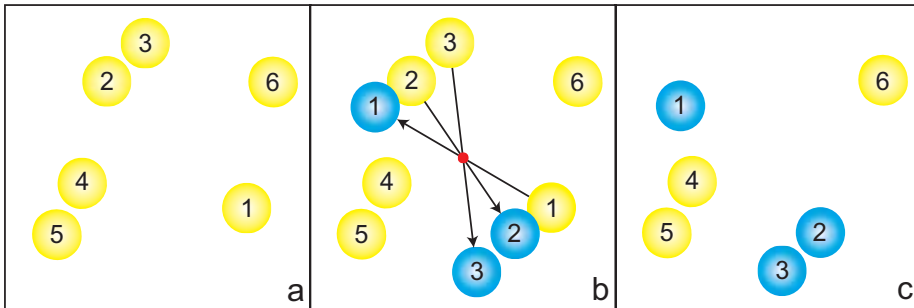


FIGURE 1. Two-dimensional illustration of the interacting geometric cluster algorithm. Light and dark colors label the particles before and after the geometrical operation, respectively. The small circle denotes the pivot. a) Initial configuration; b) construction of a new cluster and move of particles 1–3 to new positions through point reflection with respect to the pivot; c) final configuration.

and Krauth faces a fundamental limitation: it only pertains to particles with hard-core interactions. Accordingly, every (non-overlapping) configuration has the same Boltzmann factor and the formulation of a valid MC scheme is straightforward. The inclusion of additional pair interactions has been attempted by imposing a Metropolis-type acceptance criterion [24], in which, upon construction, a cluster is only moved with a certain probability. It is *a priori* clear, however, that such an approach faces severe consequences: (i) Smooth interparticle potentials (e.g., Lennard-Jones) cannot be simulated, as the cluster-building process imposes a repulsive core of infinite strength. (ii) The Metropolis criterion requires the computationally expensive calculation of *all* interactions between particles that constitute the cluster and the remainder of the system; for strong interactions, the large number of “broken” pair interactions will lead to a very low acceptance probability. (iii) The absence of a relation between actual interactions and the cluster construction process implies that the percolation threshold will not coincide with the critical point, a flaw that has been proven to be fatal in other situations, such as frustrated systems.

The generalization of the hard-sphere algorithm to fluids of interacting particles (schematically illustrated in Fig. 1) addresses these issues via a cluster-construction procedure that takes into account all interactions [25]. The resulting, rejection-free algorithm exhibits several features that make it particularly suitable for the study of colloid-nanoparticle solutions, binary mixtures, and other fluids in which the constituents have a large size asymmetry. The efficiency gain that can be reached is illustrated through the simulation of a mixture of large and small hard spheres, in which the large particles also have a Yukawa repulsion. Both particle types occur at identical packing fraction 0.1 and have a diameter ratio α . Figure 2(a) shows the energy autocorrelation time as a function of α , both for Metropolis-type updates and for the interacting geometric cluster algorithm. Already for a modest size ratio of 7, a performance increase by more than three orders of magnitude is achieved. Figure 2(b) illustrates that also in the absence of a size asymmetry the new algorithm yields an improvement. The divergence of the energy autocorrelation time for a critical Lennard-Jones fluid, which scales as a power law

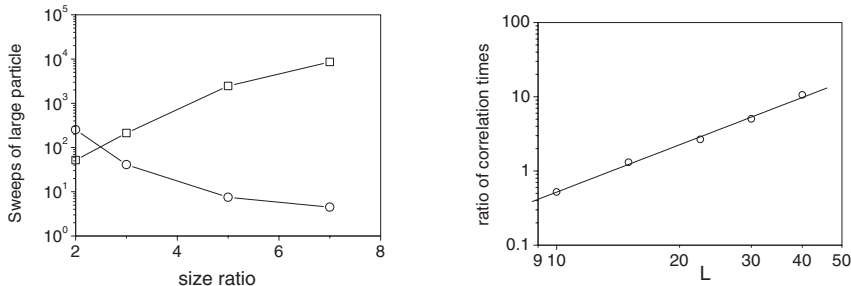


FIGURE 2. (a) Energy autocorrelation time in unit of sweeps of large particles as a function of size ratio. The open squares represent a conventional Metropolis-type method; the circles pertain to the geometric cluster algorithm. For a further discussion see the text. (b) Ratio of energy autocorrelation times for Metropolis-type updates and the geometric cluster method in a critical Lennard-Jones fluid, as a function of linear system size. The efficiency improvement amounts to approximately a factor L^2 .

of the system size, is suppressed in the geometric cluster algorithm by approximately a factor L^2 . This makes this method particularly appropriate for the study of critical fluids.

CONCLUSION

In summary, we have demonstrated that cluster methods not only suppress critical slowing down—a feature for which they rightfully have become famous—but also allow a rethinking of existing concepts in Monte Carlo simulations. The application of the algorithms described is an area full of opportunities and unsolved issues. The long-range cluster algorithm can be generalized to vector-spin models (XY , Heisenberg) with algebraically decaying interactions, a topic that has received only scant attention. Also the dynamic critical properties of systems with such interactions, have not been investigated numerically. We anticipate that the geometric cluster algorithm for interacting fluids will find widespread application in the simulation of complex fluids.

ACKNOWLEDGMENTS

This work is supported by the U.S. Department of Energy, Division of Materials Sciences under Award No. DEFG02-91ER45439 through the Frederick Seitz Materials Research Laboratory at the University of Illinois at Urbana-Champaign and by the STC Program of the National Science Foundation under Agreement No. CTS-0120978.

REFERENCES

1. Swendsen, R. H., and Wang, J.-S., *Phys. Rev. Lett.*, **58**, 86–88 (1987).
2. Binder, K., and Luijten, E., *Physics Reports*, **344**, 179–253 (2001).
3. Wolff, U., *Phys. Rev. Lett.*, **62**, 361–364 (1989).
4. Kasteleyn, P. W., and Fortuin, C. M., *J. Phys. Soc. Jpn. Suppl.*, **26s**, 11–14 (1969).
5. Fortuin, C. M., and Kasteleyn, P. W., *Physica*, **57**, 536–564 (1972).
6. Luijten, E., and Blöte, H. W. J., *Int. J. Mod. Phys. C*, **6**, 359–370 (1995).
7. Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G., *J. Chem. Phys.*, **103**, 8577–8593 (1995).
8. Greengard, L., and Rokhlin, V., *Acta Numerica*, pp. 229–269 (1997).
9. Hrycak, T., and Rokhlin, V., *Siam J. Sci. Comput.*, **19**, 1804–1826 (1998).
10. Dress, C., and Krauth, W., *J. Phys. A*, **28**, L597–L601 (1995).
11. Baxter, R. J., Kelland, S. B., and Wu, F. Y., *J. Phys. A*, **9**, 397–406 (1976).
12. Luijten, E., “Monte Carlo simulation of spin models with long-range interactions,” in *Computer Simulation Studies in Condensed-Matter Physics*, edited by D. P. Landau, S. P. Lewis, and H. B. Schüttler, Springer, Heidelberg, 2000, vol. XII, pp. 86–99.
13. Luijten, E., and Blöte, H. W. J., *Phys. Rev. B*, **56**, 8945–8958 (1997).
14. Luijten, E., and Blöte, H. W. J., *Phys. Rev. Lett.*, **76**, 1557–1561, 3662(E) (1996).
15. Luijten, E., Blöte, H. W. J., and Binder, K., *Phys. Rev. Lett.*, **79**, 561–564 (1997).
16. Luijten, E., and Binder, K., *Phys. Rev. E*, **58**, R4060–R4063 (1998), also **59**, 7254(E) (1999).
17. Luijten, E., and Binder, K., *Europhys. Lett.*, **47**, 311–317 (1999).
18. Luijten, E., and Meyer, H., *Phys. Rev. E*, **62**, 3257–3261 (2000).
19. Kim, Y. C., Anisimov, M. A., Sengers, J. V., and Luijten, E., *J. Stat. Phys.*, **110**, 591–609 (2003).
20. Luijten, E., and Meßingfeld, H., *Phys. Rev. Lett.*, **86**, 5305–5308 (2001).
21. Luijten, E., and Blöte, H. W. J., *Phys. Rev. Lett.*, **89**, 025703 (2002).
22. Uzelac, K., and Glumac, Z., *Physica*, **314**, 448–453 (2002).
23. Hucht, A., *J. Phys. A*, **35**, L481–L487 (2002).
24. Malherbe, J. G., and Amokrane, S., *Mol. Phys.*, **97**, 677–683 (1999).
25. Liu, J., and Luijten, E., preprint (2003).