JOURNAL OF COMPUTATIONAL PHYSICS 95, 167–174 (1991)

# Accelerated Monte Carlo by Embedded Cluster Dynamics*

R. C. BROWER

*Department of Electrical, Computer and Systems Engineering, and Department of Physics, Boston University, Boston, Massachusetts 02215*

N. A. GROSS

*Department of Physics, Boston University, Boston, Massachusetts 02215*

AND

K. J. M. MORIARTY

*Institute for Computational Studies, Department of Mathematics, Statistics and Computing Science, Dalhousie University, Halifax, Nova Scotia, Canada B3H 3J5*

Received October 23, 1989; revised April 2, 1990

We present an overview of the new methods for embedding Ising spins in continuous fields to achieve accelerated cluster Monte Carlo algorithms. The methods of Brower and Tamayo and Wolff are summarized and variations are suggested for the $O(N)$ models based on multiple embedded $Z_2$ spin components and/or correlated projections. Topological features are discussed for the $XY$ model and numerical simulations presented for $d=2$, $d=3$ and mean field theory lattices. © 1991 Academic Press, Inc.

## 1. INTRODUCTION

The cluster algorithms introduced by Swendsen and Wang [1] for the Potts model have recently been generalized to statistical systems and quantum field theories with continuous fields. For example, algorithms, which either eliminate or drastically reduce critical slowing down, have been proposed and tested by Brower and Tamayo [2] for the single component $\lambda\phi^4$-theory and by Wolff [3] for the $N$-vector spin models. Similar suggestions have been made by Niedermayer [4]. It has now become evident that these new embedded Swendsen–Wang algorithms can be applied to a wide range of field theories encountered in the Higgs sector of particle physics and to continuous spin models of statistical mechanics. Moreover,

* Talk presented at the "Fourth Copper Mountain Conference on Multigrid Methods, Copper Mountain, Colorado, April 9–14, 1989."



0021-9991/91 $3.00

efforts are also underway to extend these embedding methods to continuous gauge theories [5].

The purpose of this paper is to give a brief overview of the central concept of embedding that makes possible these more powerful techniques. We will also present some new results for the $XY$ model in $d=2$ and $d=3$ dimensions and in mean field theory (or infinite dimensions) limit. The mean field theory results are particularly interesting because they are often an indication of the dynamics expected for the upper critical dimension ($d=4$) encountered in applications to particle physics. In fact, based on the apparent absence of critical slowing down in our mean field simulations presented below, it is natural to conjecture that there is also *no critical slowing down* with the Wolff improved Swendsen–Wang algorithm for all the $N$-vector $\phi^4$-theories at the upper critical dimension $d=4$.

## 2. Embedding Schemes for Continuous Fields

The Swendsen–Wang dynamics was first introduced for Potts models that have the special property of a single energy level per bond. This feature is essential to the application of the Fortuin–Kasteleyn map [6] and the physical properties of the Coniglio–Klein percolation clusters [7]. Until recently this discrete two-level feature appeared to militate against useful extensions to field theories with continuous degrees of freedom. To apply these methods more widely, one is faced with the task of choosing a restricted set of states represented by discrete variables and augmenting the Swendsen–Wang dynamics with some other random process to ensure ergodicity.

Heuristically, we can understand the recent success of embedding schemes of the Swendsen–Wang dynamics as follows. In the ideal algorithm, one would like to update the system from any equilibrium field configuration $C' = \{\boldsymbol{\phi}'(x)\}$ to another $C'' = \{\boldsymbol{\phi}''(x)\}$ in one cycle.[1] How could a set of $Z_2$ variables ever approximate this ideal? We can of course introduce a set of $Z_2$ spin variables, $s_x = \pm 1$, to connect these two configurations by defining a linear relation,

$$\boldsymbol{\phi}(x) = \mathbf{a}(x) + s_x \mathbf{b}(x), \tag{1}$$

such that $\phi'(x) \to \phi''(x)$ as $s_x \to -s_x$. For a general polynomial field theory,

$$A = \sum_{x,y} K[\boldsymbol{\phi}(x), \boldsymbol{\phi}(y)] + \sum_x V[\boldsymbol{\phi}(x)] \tag{2}$$

[1] The spatial lattice sites are labelled by lower case $x$ or $y$ whereas the spin components in the $O(N)$ model are labelled by superscripts $1, 2, ..., N$ or, in the case of the two-component $XY$ model, by upper case $X$ and $Y$.

the effective action for the discrete spins, $s_x$, is

$$A_{\text{eff}} = \sum_{x,y} \beta_{xy} s_x s_y + \sum_x h_x s_x, \tag{3}$$

an Ising model in a magnetic field.[2]

Of course this construction begs the question of how to pick the coefficients $a(x)$ and $b(x)$ so as not to upset detailed balance, and the question of whether any particular Swendsen–Wang dynamics is in fact efficient. However, if we are close to a second-order phase transition, we expect that the coefficients $a(x)$ and $b(x)$ are generally smooth functions of $x$, inside the correlation length. Thus we can try to approximate the mapping by simple choices of $a(x)$ and $b(x)$ for which detailed balance is easy to impose and for which the Swendsen–Wang dynamics is easy to implement.

For the single component embedding of Ref. [1], the choice was

$$\phi''(x) = a(x) + s_x\,|\phi'(x) - a(x)| \tag{4}$$

with $a(x)$ taken to zero to simplify the action. Ergodicity was then enforced by alternating the Swendsen–Wang dynamics on the embedded $s_x$ variables with a local heat bath cycle on the $\phi(x)$ variables. We also note that since we have defined the spins so that all the $\beta_{xy}$'s are positive, the resulting effective theory has no frustrations. Also it should be noticed that $\beta_{xy} = |\phi(x)\,\phi(y)|$, so that the percolation rate $p_{xy} = 1 - \exp(-2\beta_{xy})$ is turned off adiabatically as we approach small values of the fields. We feel this softening of the domain boundaries is essential to the success of the embedding. Moreover, the numerical result for the dynamical critical exponent indicates that this embedding maps $\phi^4$-theory into the same dynamics universality class as the Swendsen–Wang Ising dynamics.

For the $XY$ model considered by Wolff, we can embed $Z_2$ into the angular coordinate,

$$\theta''(x) = a(x) + s_x\,|\theta'(x) - a(x)|. \tag{5}$$

Again the effective action has the same form, $A_{\text{eff}} = \beta_{xy} s_x s_y + h_x s_x$, where Wolff chose to eliminate the magnetic field by setting $a(x) = \theta_0$ with $\theta_0$ chosen to be a random variable on the interval $(0, \pi)$.

Let us reconsider Wolff's embedding in a slightly different form suggestive of other generalizations. The constant, $\theta_0$, is a global rotation, so we can enforce ergodicity by literally performing such a random rotation between each update. If we now consider the complex $\phi^4$-theory (as a variant on the $XY$ model), we see at once how closely this is related to the real $\phi^4$-theory example of Brower and Tamayo. Each $XY$ component, real or imaginary, is treated exactly like the one real component in the $\phi^4$-theory. It is natural also to extend the $O(N)$ embedding to $N$ copies of $Z_2$ variables, one for each component in the $N$-vector $\phi^4$-theory. The

[2] $\beta_{xy} = 1/4 \sum_{s_x, s_y} s_x s_y K[\boldsymbol{\phi}(x), \boldsymbol{\phi}(y)]$ and $h_x = 1/2 \sum_{s_x} s_x \{\sum_y K[\boldsymbol{\phi}(x), \boldsymbol{\phi}(y)] + V[\boldsymbol{\phi}(x)]\}$.

*multiple embedding* for general $O(N)$ introduces $N$ types of spins, $s_x^\alpha$ for each $\phi^\alpha$ component,

$$\boldsymbol{\phi}(x) = (s_x^1 \, |\phi_1(x)|, s_x^2 \, |\phi_2(x)|, ..., s_x^N \, |\phi_N(x)|). \tag{6}$$

Ergodicity is again enforced by a random rotation.

Similar embeddings can be introduced for chiral models as well. If the action is given by $\mathrm{tr}(U_x^\dagger U_y)$, one can introduce a discrete reflection $U_x \to -U_x^\dagger$ corresponding to $s_x \to -s_x$ and a random uniform "rotation" of the axes by $U_x \to A U_x B$, where $A$ and $B$ lie in the same unitary group as $U_x$ [8].

Now let us briefly recall the cluster dynamics that one can use on the effective Ising theory. It consists of three steps:

1. Pick an embedding scheme and find the effective Ising action with percolation probabilities for the links $p_{xy} = 1 - \exp[-\beta_{xy}(1 + s_x s_y)]$.

2. For Swendsen–Wang dynamics percolate the entire lattice and flip each cluster with 50% probability, or for the single cluster Wolff dynamics grow a single cluster from a random seed and flip it with 100% probability.

3. If necessary for ergodicity, insert random rotation of axes or a local heat bath update on the full theory and return to the first step.

For the Wolff single cluster scheme the effective "time" is rescaled by the fraction of spins in a cluster relative to the entire lattice. Hence in both algorithms an autocorrelation time $\tau$ refers to the number of passes through the entire lattice to decorrelate the system.

## 3. Topological Features

The work of Wolff [3] and Edwards and Sokal [9] on the $d=2$ $XY$ model has demonstrated a truly remarkable speedup of the critical dynamics associated with the Kosterlitz–Thouless transition. Additional evidence is presented in this paper for $d=2$, $d=3$, and mean field theory. To understand intuitively how this speedup might come about for $d=2$, it is necessary to understand how the vortices can be rapidly reconfigured by the cluster update scheme. Consider the domain structure in both the spin variables corresponding to the real ($X$) and imaginary ($Y$) components ($s_x^1, s_x^2$). A vortex is represented by the intersection of the type 1 (or $X$) and type 2 (or $Y$) domain walls. This picture can easily be understood by noting that the winding number, given by calculating the phase change in a circle around the intersection of the domain walls, must be $\pm 2\pi$. Thus, the intersection of two domains will give a vortex and antivortex pair as seen in Fig. 1. A single cluster update for one of these variables ($s_x^1$ or $s_x^2$) is capable of destroying any or all of the vortices and creating a vastly different arrangement of vortices.

A similar accelerated dynamics occurs for the cluster algorithms for the $XY$ model in $d=3$ and mean field theory. In $d=3$, the intersections of the contours

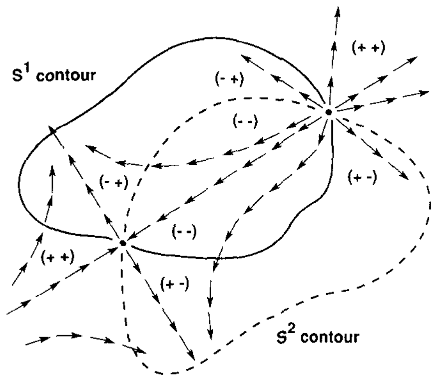


FIG. 1. Embedded spins, $s^1$ and $s^2$, for the $XY$ model projected against the $X$ and $Y$ axes, respectively. The intersection between the $s^1$ domain wall (solid line) and the $s^2$ domain wall (dashed line) locates a vortex–antivortex pair.

represent the vortex tubes in space. Similarly if we look at the $O(3)$ model in $d=3$, we see that the mutual intersection of the three types of domain walls for $(s_x^1, s_x^2, s_x^3)$ correspond to the "Hedgehog" solution. All of these intersections give position coordinates for the point defects in the continuum limit. On the lattice the choice of axis for the embedded spin projections only slightly alters the local crossing points for the domain walls.

## 4. ERGODICITY

In Refs. [2, 3], the overall ergodicity of the scheme was ensured by supplementing the Swendsen–Wang process with another random process. Since the Swendsen–Wang process satisfies detailed balance by itself, it is sufficient if this new random process also satisfies detailed balance. For instance, in the single cluster update scheme, Wolff chose to project the spins against a random axis $\hat{e}$. More general choices are possible. For the $O(N)$ model, we tried the following correlated axis approach:

- Let one of the projection axes $\hat{e}$ be correlated to the vector $\phi$ at the seed by a symmetric probability distribution $\rho(\alpha)$, where $\alpha$ is the angle between $\phi(x)$ and $\hat{e}$.

If the axis is generally forced to align with the seed spin, one grows larger clusters on average, so one might expect even faster decorrelations. However, as one can see in Fig. 2b the fully aligned axis $(\alpha=0)$ algorithm gives almost the same renormalized decorrelation time $\tau$ in 3D as Wolff's random axis approach.

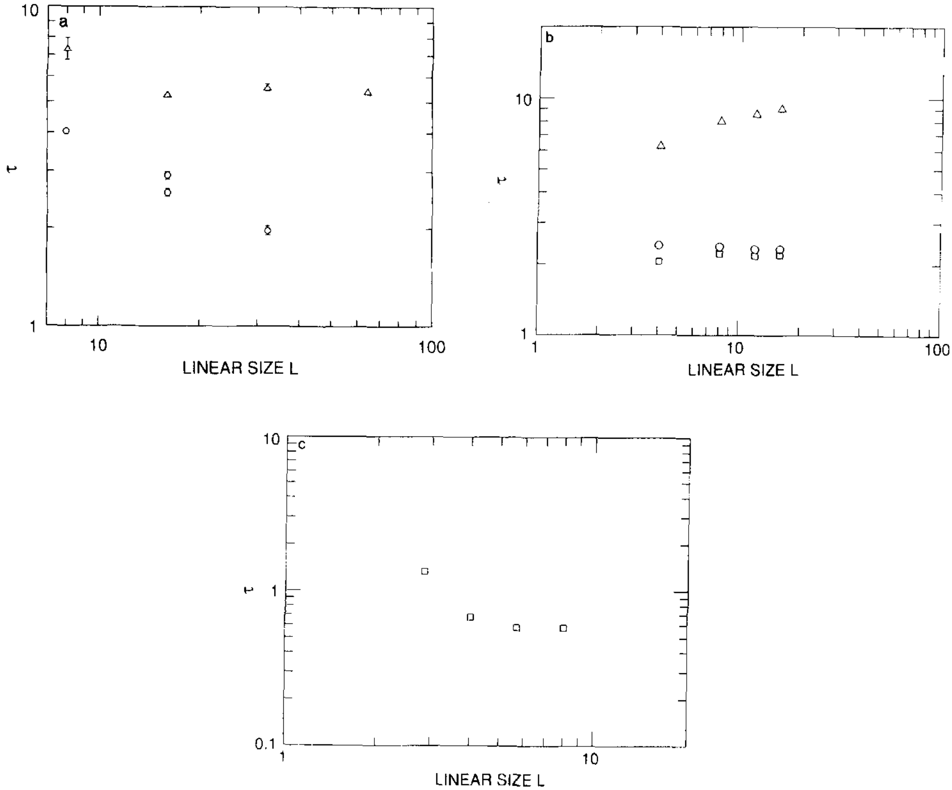


FIG. 2. (a) The relaxation time ($\tau$) for susceptibility autocorrelation vs lattice size for the 2D $XY$ model. Triangles (△): Swendsen–Wang; circles (○): aligned Wolff. (b) Susceptibility autocorrelation $\tau$ for the 3D $XY$ model. Triangles (△): Swendsen–Wang; circles (○): aligned Wolff; squares (□): random Wolff. (c) Susceptibility $\tau$ for the mean field $XY$ model for the random Wolff update, where the linear size $L$ is defined by $N^{1/4}$.

## 5. Simulations and Conclusions

We have performed extensive simulations on the $XY$ model in $d=2$, $d=3$, and mean field theory lattices. We have measured the autocorrelations and estimated the dynamical critical exponent $z$ from finite size effects, i.e.,

$$\tau = \text{const} \times L^{z}, \qquad (7)$$

where $L$ is the linear size of the lattice. Figures 2a, b, and c summarize our present measurements of the susceptibility autocorrelations for $d=2, 3$, and mean field theory, respectively. These simulations were run at $\beta = 1.02$, 0.474, and $2/(N-1)$, respectively, $N$ being the volume of the mean field system. These temperatures are in the scaling region, near the critical $\beta$, of their respective systems.

In two dimensions, the number of iterations were from 40 k for the largest lattice to 200 k for the smallest. For 3D, the number of iterations for both types of Wolff updates were from 500 to 1000 k, while the Swendsen–Wang algorithm varies from 150 to 400 k iterations. The mean field simulations for the random Wolff algorithm were all 100 k iterations long. The $\tau$'s were extracted by fitting to a single exponential decay mode of the correlation function.

For two dimensions (Fig. 2a) the Swendsen–Wang algorithm shows a slope that is within the error bars of zero. The aligned Wolff algorithm has a *negative* slope with $z = -0.53 \pm 0.02$. This is a little surprising although not logically impossible. It is also possible that we are seeing finite size effects, and that this curve may flatten out as we go to higher lattice sizes; although there is no evidence for this so far. For three dimensions (Fig. 2b) $z_{sw}$ is small but nonzero. A fit puts the slope at $0.18 \pm 0.021$. The slope of both Wolff algorithms appears to be zero within the error bars. For mean field theory (Fig. 2c) there is data for the random Wolff algorithm only. The slope appears to approach zero asymptotically.

The first thing to note is that for all simulations below the upper critical dimension, the embedded dynamics for the $XY$ model is *faster* than the corresponding pure Ising model. In fact, the Wolff and aligned algorithms are so fast in $d = 2$ and $d = 3$ that there is no apparent critical slowing down (in times rescaled to a full lattice equivalent of updated spins). This is surprising in the sense that there is no theoretical analysis such as that provided by Fortuin and Kasteleyn and Coniglio and Klein for the Ising system to relate the percolation clusters to the $XY$ model. We are investigating more closely the structure of the clusters in the $XY$ model to try to establish such a connection.

In conclusion, we have attempted to present an overview and a heuristic motivation for the $Z_2$ embedding schemes for continuous field theories. Also we have pointed out (after the fact) how the topological solutions to the field theories are related to the multiple embedded Ising systems. This step along with other efforts to arrive at a "taxonomy" of cluster dynamics may help to inspire a real dynamical theory.

We have also presented some additional simulations for the $XY$ model in the mean field limit that exhibit zero critical slowing down for the Wolff single cluster update scheme. Since for the Ising model, the Wolff dynamics is consistent with zero critical slowing down for both mean field theory and 4D Wolff [10], it is natural to conjecture that in general Wolff's approach will give no critical slowing down for the upper critical dimension. It is an important goal for future simulations to give convincing numerical evidence for (or against) this conjecture for the entire class of 4D $O(N)$ Higgs theories. Moreover, if critical slowing down is indeed absent, a major theoretical goal should be to find the underlying mechanism by which it is eliminated. We feel that ultimately, any "perfect" dynamics, with $z = 0$, can, like the fast Fourier transform for linear systems, be a source of deeper theoretical understanding of the physics of Higgs models.

## ACKNOWLEDGMENTS

We thank William Klein, Eric Myers, Alan Sokal, and Pablo Tamayo for useful discussions throughout this work. We thank John Buchanan, Terence R. B. Donahoe, and Don Cameron of the Government of Nova Scotia for their continued interest, support, and encouragement and grant support; James R. Berrett, Samuel W. Adams, and Dana Hoffman for their continued interest, support, and encouragement and access to the NEC SX-2 in the Houston Advanced Research Center in The Woodlands, Texas, where some of these calculations were performed; HNSX Supercomputers, Inc. for an HNSX Supercomputer Fellowship Award; the National Allocation Committee for the John von Neumann National Supercomputer Center for access to the two CDC CYBER 205's and the two ETA-10's at JVNC (Grants 110128, 171805, E171805, 171812, E171812, 171813, 551701–551706), where the rest of these calculations were performed. We acknowledge the Natural Sciences and Engineering Research Council of Canada (Grants NSERC A8420 and NSERC A9030) for financial support and the Canada/Nova Scotia Technology Transfer and Industrial Innovation Agreement (Grants 87TTII01, 88TTII01, 89TTII01, and 90TTII01) for further financial support. We thank Steve F. McCormick for the invitiation to present a talk at the "Fourth Copper Mountain Conference on Multigrid Methods," while one of the authors (R.C.B.) would like to thank Control Data Corporation for a grant which made possible his visit to Copper Mountain, Colorado.

## REFERENCES

1. R. H. SWENDSEN AND J.-S. WANG, *Phys. Rev. Lett.* **58**, 86 (1987).
2. R. C. BROWER AND P. TAMAYO, *Phys. Rev. Lett.* **62**, 1087 (1989).
3. U. WOLFF, *Phys. Rev. Lett.* **62**, 361 (1989).
4. F. NIEDERMAYER, *Phys. Rev. Lett.* **61**, 2026 (1988).
5. R. C. BROWER AND S. HUANG, "Random Surface Dynamics of $Z_2$ Gauge Theory," Boston University Preprint No. BUHEP-89-17, 1989 (unpublished).
6. C. M. FORTUIN AND P. W. KASTELEYN, *Physica* **57**, 536 (1972).
7. A. CONIGLIO AND W. KLEIN, *J. Phys. A* **13**, 2775 (1980).
8. A. SOKAL, Courant Institute of Mathematical Sciences, New York University, New York, NY, private communication (1989).
9. R. G. EDWARDS AND A. D. SOKAL, "Dynamic Critical Behavior of Wolff's Collective-Mode Monte Carlo Algorithm for the Two-Dimensional $O(n)$ Nonlinear $\sigma$-Model," NYU preprint, 1989 (unpublished).
10. P. TAMAYO, R. C. BROWER, AND W. KLEIN, *J. Statist. Phys.* **58**, 1083 (1990).