CLUSTER ALGORITHMS FOR NON-LINEAR SIGMA MODELS

Uli Wolff
CERN, Theory Division
CH-1211 Genève 23, Switzerland

Abstract

Percolation cluster Monte Carlo algorithms for nonlinear $\sigma$-models on the lattice are reviewed with special emphasis on their possible generalizations. While they have been found to practically eliminate critical slowing down for the standard $O(n)$ invariant vector models, their extension to other physically similar models — like $RP_{n-1}$ and $SU(n) \times SU(n)$ chiral models — is less straight forward than one might have thought. I outline the present situation in this area of research. In the second part of my talk I described a numerical calculation of a physical running coupling constant in the $O(3)$ model. This represents an application of the cluster technique in a preparatory study for a later lattice gauge theory calculation. This material can be found in Ref. 11.

Talk presented at the workshop on Fermion Algorithms at HLRZ, Jülich, Germany, April 10–12, 1991
In the following I discuss the method of Monte Carlo simulation—a both by local and by cluster techniques—as far as possible for a general nearest neighbor coupled abstract spin system given by a lattice partition function

\[ Z = \int \prod_x d\mu(s_x) \exp \left\{ -\beta \sum_{<xy>} E(s_x, s_y) \right\} . \]  

(1)

The spins may live in some manifold \((Z_2, R, S_n, SU(n), \ldots)\) with \(d\mu(.)\) denoting the appropriate invariant measure. \(E(s, s')\) is the nearest neighbor energy assumed to be invariant under a global symmetry like for instance the scalar product \(s \cdot s'\) in the \(O(n)\) \(\sigma\)-model. For such a system, updates by local, Multigrid or Fourier accelerated algorithms all involve a proposal of changing a given configuration \(\{s_x\}\) by

\[ s_x \rightarrow s'_x = A^t_x s_x , \]  

(2)

where \(A^t_x\) is an operator acting on a spin at \(x\) with one (or a few) “strength” parameter \(t\) characterizing the “size” of the potential move. It is reasonable to define \(t\) such that

\[ A^0_x = id, \quad A^{-t}_x = (A^t_x)^{-1} \]  

(3)

holds. The strength \(t\) is now chosen according to the induced effective Hamiltonian

\[ H_{eff}(t) = H[A^t_x s_x] = \sum_{<xy>} E(A^t_x s_x, A^t_y s_y) \]  

(4)

by a Metropolis or heatbath procedure, and the substitution (2) is carried out with this value of \(t\). With this presentation I emphasize the fact that one may consider these updates as proceeding through a sequence of embedded systems with one (or a few) degrees of freedom, for which a procedure close to direct independent sampling is feasible. I now elucidate the above by explaining what \(A^t_x\) is in some well known cases:

- for discrete spins, \(t\) can only assume discrete values; for the Ising case, for example, \(t = 0, 1\), and \(A^1_x\) flips \(s_x\)
- for any local update step at \(x_0\), \(A^t_x = id\) whenever \(x \neq x_0\)
- multigrid at block \(B\): \(A^t_x = id\) whenever \(x \notin B\); in the version of Ref. 2, for instance, one has \(A^t_x = \exp(it\lambda f(x)t)\) with \(\lambda\) a generator of the global

\(^a\)Reviews on the subject can be found in Ref. 1
symmetry group, and $f(x)$ characterizes the shape of the collective move performed

- Fourier acceleration: $A^i_x \phi_x = \phi_x + t \cos(k \cdot x)$ for a scalar field $\phi$ (e.g. in a $\phi^4$ model)

I wish to emphasize the following common features of these cases:

- the type and geometric shape of the moves "offered" to the system are fixed
- the "strength" $t$ of a move is chosen by the dynamics and energetics of the system

It is clear that the success and the efficiency of these procedures depend on our ability to guess the kind of moves that fit with the physics of the model. Cluster algorithms can be viewed as following a slightly different strategy as a first step: instead of one degree of freedom one embeds local Ising variables into the given spin model. This embedded Ising model is then updated collectively by the Swendsen Wang algorithm or its single cluster variant. The embedding proceeds analogously to (2),

$$s_x \rightarrow s'_x = (R_x)^{\tau_x} s_x,$$

where the Ising variables are $\tau_x = 0, 1$ (connected by $\tau_x = \frac{1}{2} (1 - \sigma_x)$ with standard spins $\sigma$), and for consistency with Ising spins $R_x \circ R_x = id$ has to hold. The embedded spins are updated according to the induced effective Hamiltonian

$$H_{eff}[\sigma_x] = H[(R_x)^{\tau_x} s_x] = - \sum_{<xy>} J_{xy} \sigma_x \sigma_y - \sum_x h_x \sigma_x.\tag{6}$$

Here I simply wrote the most general nearest neighbor Ising Hamiltonian that may result; the bond strengths $J_{xy}$ and the local magnetic fields $h_x$ follow from the parameters and the configuration of the original model (see below). The points essential in the present context are now:

- moves are locally restricted to the involutive kind, $(R_x)^2 = 1$, hence the strength cannot be made arbitrarily small, but only yes or no
- the geometric shape of the move in phase space parametrized by a whole local field is nonlocal and still arbitrary; it is determined dynamically by the Ising update (e.g. Swendsen Wang)

A comment is in order here. It may be objected that the Swendsen Wang algorithm has originally been formulated for arbitrary $q$-state Potts models and that the view just presented, which always goes through Ising embedding,
seems too narrow. It is however possible to replace the original algorithm for
$q > 2$ by one with embedded Ising spins. In this case one takes a randomly
chosen transposition\(^a\) of 2 of the \(q\) spin values as flip operation \(R_x\) to which
the 2-valued \(\tau_x\) are attached. The resulting algorithm differs in detail from
the original proposal, but it captures essentially the same collective moves.
In connection with the single cluster method it may actually be superior in
certain physical situations. It was precisely this point of view, which enabled
the treatment of the interesting antiferromagnetic Potts models with amazing
efficiency\(^6\). As far as I am aware, all successful cluster simulations carried out
to date can essentially be viewed as a reduction to randomly embedded one-bit
Ising fields in a similar way.

It is not difficult to determine the parameters in the effective action (6) for
the generic spin model (1). The answer reads

\[
-J_{xy} = \frac{1}{4} \sum_{\tau_x, \tau_y = 0, 1} E(R_x^{\tau_x} s_x, R_y^{\tau_y} s_y)(-1)^{\tau_x + \tau_y}
\]

(7)

and

\[
-h_x = \sum_{y = n.n.(x)} \frac{1}{4} \sum_{\tau_x, \tau_y = 0, 1} E(R_x^{\tau_x} s_x, R_y^{\tau_y} s_y)(-1)^{\tau_x},
\]

(8)

where the \(y\) sum is over all nearest neighbors and for simplicity we assumed the
measure in (1) to be invariant, \(d\mu(R_x s_x) = d\mu(s_x)\). The practical realization
of an embedded update starts from all \(\sigma_x = +1\) (corresponding to the "old"
configuration \(s_x\)), then updates the Ising field with one or several cluster steps
using the Hamiltonian (6), and finally transfers the Ising values, at which
one arrives, back to the \(s_x\) spins by (5). Then a new embedding \(R_x\) may be
chosen and the process is iterated. Although it is not hard to show detailed
balance and one has a legal algorithm\(^b\), it may be very inefficient for one of
the following reasons:
(i) \(H_{eff}[\sigma_x = 1]\) is dominated by large magnetic fields and the Ising spins
cannot move
(ii) \(|J_{xy}| = O(1)\), and \(\prod_{\text{plaqu}} J_{xy} < 0\) for many plaquettes and typical \(s_x\); then
the embedded system is of the spin glass type, for which Swendsen Wang
and all other presently known algorithms are slow

\(^a\)Note that transpositions are the smallest elements of the global symmetry
group, which here is the permutation group of \(q\) elements

\(^b\)To arrange for ergodicity, if necessary by blending in other algorithms, is
usually not a problem.
(iii) even if the $\sigma_x$ move a lot it may be that not “enough” $s_x$ are sampled

An important special case arises if $R_x$ leaves the local energy function
invariant

$$E(R_x s_x, R_y s_y) = E(s_x, s_y).$$  \hfill (9)

Then the magnetic fields $h_z$ in (8) vanish, and in (7) only two different terms
appear under the sum. The successful $O(n)$ model reflection algorithm\(^3\) belongs to this class. With spins being $n$-component real unit vectors a spin
direction $r$ ($|r| = 1$) is chosen randomly and the flip operator is globally taken
to be

$$R_x = (1 - 2 P_r),$$  \hfill (10)

where $P_r$ is the projector onto the $r$-direction in spin space. The corresponding
bonds appear in the factorized form

$$J_{xy} = s_x \cdot r s_y \cdot r$$  \hfill (11)

which implies the absence of frustration. The sign of $J_{xy}$ is a pure gauge in
$Z_2$ in the language of gauge theory.

The authors of Ref. 7 analyze the success of the $O(n)$ algorithm and argue
for which ingredients may be decisive. If one imagines a classically smooth spin
field $s_x$, it is clear from (11) that the bonds $J_{xy}$ vanish on surfaces $s_x \cdot r = 0$.
As this is one real condition, these surfaces are expected to be generically
of dimension $d - 1$ if the spins live on a $d$ dimensional space. Equivalently,
the fixed point manifold of (10) has codimension one\(^7\). As a consequence,
the embedded Ising model decouples along these surfaces, and space typically
breaks up into a corresponding number of Swendsen Wang clusters. Had $R_x$
a codimension larger than one, then, according to this classical argument, the
resulting fixed point surfaces are insufficient to divide up space, and typically
most spins may end up in one percolating cluster. If, on the other hand, the
correlation length is finite, these are not the dynamically appropriate moves.
The above scenario gains even more plausibility if one takes into account that
by simple group theory

$$\langle s_x \cdot s_y \rangle = n(s_x \cdot r s_y \cdot r)$$  \hfill (12)

holds with $r$ being the randomly chosen embedding direction. As an exponential
decay of the right hand side is expected to arise from sign oscillations, the
sign changes and therefore zeroes of $s_x \cdot r$ control both the breakup and the
correlation length.
One may object that the argument is still only heuristic. Even for large correlation length, spin fields in 2 dimensions, for instance, do not look very smooth. Also percolation clusters do not look like smooth cells into which space decomposes; rather they are fuzzy interpenetrating objects that almost certainly have fractional dimension. With the stochastic cluster formation we do not really need vanishing bonds for clusters to break, as one sees by comparison with pure uncorrelated bond percolation. Nevertheless a practical test conducted for the $O(4)$ model with a flip analogous to (10), but using a two dimensional projector (codimension of the flip = 2), supports the picture. The embedded Ising model still has no magnetic field, but it can be frustrated. For this case in Ref. 7 a perfect algorithm was emulated by performing a large number of Ising sweeps to be counted as one iteration. It was found that no efficient algorithm for the $O(4)$ model results. The authors conclude that codimension one is necessary for embedding algorithms, and possible generalizations to other models should take this into account.

As one generalization of the method, simulations of a version of the $RP_{n-1}$ model have been attempted\(^8,9\). In this case the spins are still $n$ component real vectors with an energy only depending on the rays given by the spins,

$$E(s_x, s_y) = \frac{1}{2} (s_x \cdot s_y)^2. \tag{13}$$

With $R_x$ as in (10) codimension one is fulfilled, $h_x$ vanishes, but the bonds

$$J_{xy} = s_x \cdot P_x s_y s_x \cdot (1 - P_x) s_y \tag{14}$$

now allow for frustration if $n \geq 3$. My findings in an exploratory simulation of this model in 2 dimensions were mixed. Using the single cluster algorithm I found autocorrelation times $\tau \leq 10$ steps/spin for correlation lengths up to 20 on a $256^2$ lattice, which is much better than any local method but definitely worse than for the $O(n)$ vector model. Another observation is that the average cluster size grows much faster than the susceptibility and correlation length seen in the fundamental tensor field $s_x^a s_x^b$ which is in contrast to the $O(n)$ case. It looks possible that the update clusters percolate at a $\beta$ whose correlation length is finite. Then one would expect a deterioration of the algorithm. In Ref. 8 “too big” clusters were also noticed. Their average size is made smaller by correlating $r$ with the seed spin of the single clusters. This is legitimate if certain symmetry restrictions are obeyed to preserve detailed balance. However no details are given on the effect of this cure on correlation
times and the resulting cluster distribution. It should also be noted that this correlation to my knowledge deprives one of the possibility to employ reduced variance estimators. They are a second big advantage of cluster algorithms, which I do not discuss here in further detail\(^4\). On the physics of the RP\(_{n-1}\) models there are controversial claims in the literature\(^10\), and a careful and powerful numerical study is clearly still of interest\(^5\).

Another class of spin systems of great interest are the principal chiral models with global SU\(_{(n)}\) × SU\(_{(n)}\) invariance. In 2 dimensions they are asymptotically free sharing for \(n = 3\) the type of variables with QCD, and in 4 dimensions they arise as effective field theories for light mesons. Spins now consist of SU\(_{(n)}\) matrices \(s_x \in SU(n)\) with energy

\[
-E(s_x, s_y) = \frac{1}{n} \text{Re} \text{tr}(s_x^* s_y).
\]

(15)

The codimension argument\(^7\) implies a “no go” theorem for the embedded Ising cluster simulation of these models without magnetic fields. A proof has been found in the mathematics literature that involutive codimension one flips exist only for spheres or discrete quotients of them. In the sequel I report a few efforts I made for the SU\(_{(3)}\) model in spite of this situation. They may perhaps help on the way to getting around the “no go” theorem (probably by violating one of its assumptions). After all, in two dimensions, the physics is not expected to be radically different from O\(_{(n)}\) models, and from that point of view one would expect to be able to use similar tools.

We start by noting that from a special flip \(\hat{R}\) one can always get a more general one by conjugating with a symmetry operation, \(R = g \hat{R} g^{-1}, g \in SU(3) \times SU(3)\). This may be realized in practice by using only the special form for the update and performing random global symmetry transformations in between. I therefore restrict myself to proposals of the special type with the extra conjugations always understood. Also, all my tests have been performed at \(\beta = 4.5\) on a \(10^2\) lattice where the correlation length is about 3, a rather modest setting for cluster algorithms. One isometric flip, that is suggested by casting the O\(_{(4)}\) algorithm into SU\(_{(2)}\) language, is taking the adjoint\(^\dagger\).

\(^a\)I cannot find the concluded topological phase transition in Ref. 8 completely compelling, although the paper certainly contains interesting information. In addition, questions on the large \(\beta\) (asymptotic freedom ??) phase remain of interest

\(^\dagger\)A -1 factor is generalized to an arbitrary center element for SU\(_{(n)}\) and can be absorbed into the conjugation for odd \(n\).
It has the largest possible codimension of 8; another possibility is to take the transpose yielding codimension 3. Being an isometry it does not produce a magnetic field. For the average coupling strength and frustration I found 
\[ |\beta J_{xy}| = 0.80, \quad \langle \text{sign} (\Pi_{\text{isotrop}} J_{xy}) \rangle = 0.48 \]
Correlation times of hundreds of sweeps result in combination with Swendsen Wang updating.

Another possibility is to give up the requirement of the absence of generated magnetic fields. Among several such options that I tried the following codimension one flip is typical. Up to exceptions of measure zero an SU(3) matrix \( u_{a1} \) is determined by giving the first column \( u_{a1} \) and the first row \( u_{1a} \), which both have to be complex unit vectors. The remaining matrix elements may be considered as functions of those. Now we define a flip by

\[
u_{11} \to v_{11}, \quad u_{1a} \to u_{1a}, \quad u_{a1} \to u_{a1}, \quad a = 2, 3. \tag{16}
\]

The lower right \( 2 \times 2 \) block is changed accordingly. All matrices for which \( u_{11} \) is real are fixed points of this operation, and thus the codimension is one. I again measured the couplings of the embedded model and found that it is unfrustrated with \( |\beta J_{xy}| = 0.25 \) and an average field \( \langle \beta h_x \rangle = 0.29 \), \( \langle \beta h_y \rangle = 0.39 \). This field is strong enough to prevent larger clusters formed in a Swendsen Wang update from being flipped. The smaller ones, which do move, are not adequate for the physics of the model, and again correlation times of hundreds of sweeps result.

It is rather unclear whether an efficient cluster algorithm for the SU(3) spin model can still be found. Clearly there is no point in simply trying more and more flip operations of the above type. Allowing magnetic fields to be able to stick to codimension one seems an interesting possibility, but there has to be a physical reason why these fields are typically small. Such an embedding remains to be found. Possible strategies are to go through intermediate embedded U(1) or SU(2) subgroups. Such embedded models have only nearest neighbor couplings, but they do not have the standard action of genuine U(1) or SU(2) spin models. They rather sit in a background gauge field inherited from and determined by the SU(3) field, which in general is not a pure gauge but has curvature (frustration). Efficient nonlocal algorithms in such a situation are precisely what is also required when gauge theory is coupled to matter and further in spin glass problems. Of course, these richer embedded models are only intermediate steps for finding Ising embeddings, if one finally wants to update with a cluster method. Some work in this direction is in progress.
Acknowledgement

I thank Burkhard Bunk for discussions on the subject of this write-up.

References

2. M. Hasenbusch, G. Mack, and S. Meyer, Kaiserslautern preprint TH-19/90,