Cluster method for the Ashkin-Teller model

Shai Wiseman* and Eytan Domany†

Department of Electronics, Weizmann Institute of Science, Rehovot 76100, Israel

(Received 10 May 1993)

A cluster Monte Carlo algorithm for the Ashkin-Teller (AT) model is constructed according to the guidelines of a general scheme for such algorithms. Its dynamical behavior is tested for the square lattice AT model. We perform simulations on the line of critical points along which the exponents vary continuously, and find that critical slowing down is significantly reduced. We find continuous variation of the dynamical exponent z along the line, following the variation of the ratio of specific-heat and correlation-length exponents α/ν, in a manner which satisfies the Li-Sokal bound $z_{\text{cluster}} \geq \alpha/\nu$, that was so far proved only for Potts models.

I. INTRODUCTION

The Ashkin-Teller (AT) model [1] has been studied extensively, ever since its introduction, by a variety of methods. In two dimensions, in particular, much is known about the phase diagram and critical behavior of the model. Nevertheless there are problems that were not addressed extensively before; the critical behavior of random AT models is one such issue about which not much is known, hope for analytic treatment is slim, and therefore one expects numerical simulations to be the main tool of investigation. With this aim in mind, we set out to develop an efficient Monte Carlo (MC) cluster algorithm for the AT model.

A convenient representation of the *generalized* AT model (GAT) is in terms of two Ising spin variables, σ_i and τ_i, placed on every site of a lattice. Denoting by $\langle ij \rangle$ a pair of nearest-neighbor sites, the Hamiltonian is given by

$$\mathcal{H} = -\sum_{\langle ij \rangle} [K_\sigma \sigma_i \sigma_j + K_\tau \tau_i \tau_j + L \sigma_i \sigma_j \tau_i \tau_j] \ .$$

Here K_σ (K_τ) are the strengths of the interactions between neighboring σ (τ) spins, and L is a four-spin coupling. The phase diagram of the ferromagnetic general AT model is known in two dimensions from duality transformations and renormalization-group studies [2,3]. The three-dimensional model has been studied as well [4]. In this paper we are concerned with the $Z(4)$ subspace of the general model, in which $K_\sigma = K_\tau = K$. The phase diagram in this subspace is reviewed in Sec. II. The critical properties of the model in this subspace are of special interest, since it has a line of critical points, along which the exponents vary continuously, and have been determined analytically [5], interpolating between Ising and four-state Potts exponents. For instance, the value of the ratio α/ν varies from 0 at the Ising ($L = 0$) critical point to $\alpha/\nu = 1$ at the four-state Potts point $K = L$. This exponent is of special interest to us since it has been proved that for Potts models it serves as lower bound to

the dynamic exponent z of cluster algorithms [6]. Furthermore, the bound seems to be reached in the case of the Ising and Potts models [6].

Cluster algorithms [7] as introduced by Swendsen and Wang (SW) [8], and extended by Wolff [9], are reviewed briefly in Sec. III. These algorithms give rise to dynamics whose relaxation time, τ_{SW}, is significantly shorter than that of standard, single-spin flip MC methods. This is most important at a critical point, where the relaxation time of a finite system grows with its linear size L according to

$$\tau \sim L^z \ .$$

Cluster algorithms have a significantly lower dynamic exponent z_{SW} than that of standard MC. Hence if one is interested in performing extensive simulations of a model such as AT, it is well worth spending time on developing an appropriate cluster algorithm.

Creating an efficient algorithm can be a challenge. Naive application of the original SW scheme does not work (as explained in Sec. III and demonstrated in Sec. IV). We set out to construct an efficient cluster algorithm using the guidelines and methods that were presented by Kandel and Domany [10] . This general scheme is guaranteed to yield an algorithm that satisfies detailed balance, and once the important excitations of the model have been identified and incorporated, we are guaranteed to get a good algorithm. This general formalism is briefly reviewed, and the resulting algorithm is presented in Sec. III. Interestingly, the cluster algorithm found this way is identical to what would have been obtained had we used Wolff's Ising embedding method [9,11], as shown in the Appendix. Numerical results are given in Sec. IV; in particular, efficiency of our algorithm is tested by measuring the dynamic exponent along the critical line, which is indeed significantly lower than that of standard MC.

An interesting question we set out to resolve concerns comparison of z_{SW} with α/ν along the AT critical line. We found that the Li-Sokal bound [6]

$$z_{\text{cluster}} \geq \frac{\alpha}{\nu}$$

is satisfied, variation of the dynamic exponent follows that of α/ν, and within our numerical accuracy and limitations due to finite-size effects, our results indicate that the two are equal.

II. PHASE DIAGRAM OF THE ASHKIN-TELLER MODEL

We now review the phase diagram of the square-lattice AT model [4] and some of its critical properties. Figure 1 gives the phase diagram of the model plotted as a function of the parameters X and Z where

$$Z = \exp(-4K) , \quad X = \exp[-2(K+L)] . \tag{1}$$

At times we use the notation $\mathbf{X} = (X, Z)$ to denote a point in phase space. Full lines of phase transitions separate three phases.

(a) Paramagnetic, labeled as P. The couplings are sufficiently weak so that the system is in a paramagnetic phase in which neither σ nor τ (nor $\sigma\tau$) are ordered.

(b) Ferromagnetic phase, labeled as F. The couplings are sufficiently strong so that σ and τ independently order in a ferromagnetic fashion so that $\langle\sigma\rangle = \pm\langle\tau\rangle$. In this phase $\langle\sigma\tau\rangle$ is also different from zero and has the same sign as $\langle\sigma\rangle\langle\tau\rangle$.

(c) A phase labeled $\langle\sigma\tau\rangle$, in which $\sigma\tau$ is ordered ferromagnetically but $\langle\sigma\rangle = \langle\tau\rangle = 0$. This phase arises only for $L > K$.

On the dashed line $Z = X^2$ we have $L = 0$; and thus it is obviously a subspace of two decoupled Ising models, having an Ising transition at the point X_0. The dashed line $Z = X$ has $L = K$, in which case the AT model becomes the four-state Potts model. The point X_4 is a four-state Potts multicritical point.

A marginal operator generates a continuous variation of critical exponents along the line X_0X_4, isomorphic to the known critical line of the eight-vertex model [12]. Through an exact duality type transformation this line is mapped onto the critical line of a staggered eight-vertex model and through a relation with the Coulomb gas its critical exponents are known exactly [5]:
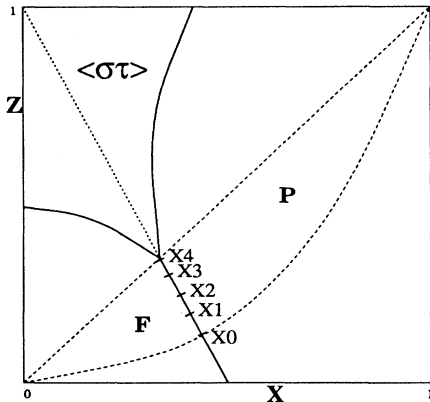
$$y_t = 2 - (2/g_R) , \tag{2}$$

where

$$g_R = \frac{8}{\pi} \sin^{-1}\left(\frac{1}{2}\coth(2K)\right) , \tag{3}$$

and $\frac{\gamma}{\nu} = \frac{7}{4}$ all along the line. Lastly the lines X_4B and X_4C flow under renormalization to Ising type fixed points.

The exact location of the transition line X_0X_4 can be found through the duality transformation of the AT model [3]. It is given by the self-dual line $Z = 1 - 2X$.

III. CLUSTER METHOD FOR THE ASHKIN-TELLER MODEL

A. SW cluster method

Cluster algorithms have proved to be a useful method of reducing critical slowing down in MC simulations. For completeness we review here the pioneering cluster algorithm of Swendsen and Wang [8] for the Ising model [13] with the Hamiltonian

$$\mathcal{H} = -J \sum_{<k,j>} \sigma_k\sigma_j , \quad \sigma = \pm 1 . \tag{4}$$

The SW procedure stochastically identifies clusters of aligned spins, and then flips whole clusters simultaneously. Starting from a given configuration u, SW go over all the bonds, and either "freeze" or "delete" them. A bond connecting two neighboring sites k and j is deleted with probability P_d, and frozen with probability $P_f = 1 - P_d$, where

$$P_d = e^{-\beta(J\sigma_j\sigma_k+J)} . \tag{5}$$

Having gone over all the bonds, all spins which have a path of frozen bonds connecting them are identified as being in the same cluster. Now the new configuration is generated by flipping every cluster with probability $\frac{1}{2}$. Note that according to (5), only spins of the same sign can be frozen in the same cluster. SW identify correctly the elementary large scale excitations as clusters of aligned spins. This correct identification is essential to the success of a cluster method for any more general model.

In Sec. III D we shall present a naive implementation of the original SW scheme to the AT model, and explain why it is not expected to work well. This necessitates search for a different clustering rule; the one we found is based on a general scheme which we now describe.

B. General scheme for cluster methods

A unifying view of all cluster algorithms has been given by Kandel and Domany (KD) [10], which we now review. The general scheme consists of two steps: given a spin configuration u, the first step consists of stochastically generating a new Hamiltonian $\tilde{\mathcal{H}}$. The second step con-

FIG. 1. Phase diagram of the Ashkin-Teller [Z(4)] model.

sists of simulating the model with the new Hamiltonian, thereby bringing it to a new configuration u'. To carry out the first step, KD write the Hamiltonian as

$$\mathcal{H} = \sum_l V_l. \qquad (6)$$

For example, V_l can be the energy of a single bond in the nearest-neighbor Ising Hamiltonian. Then to each V_l they assign one of N possible integers i, in a stochastic manner that depends on the starting configuration u. That is, the probability to assign i to l is written as $P_i^l = P_i^l(u)$. This probability is normalized, i.e.,

$$\sum_i P_i^l(u) = 1 \qquad (7)$$

for any term l and configuration u. Then they construct a new Hamiltonian:

$$\tilde{\mathcal{H}}_{\{i\}} = \sum_l \tilde{V}_{i(l)}^l, \qquad (8)$$

where (for any spin configuration $\tilde{u}$)

$$\tilde{V}_i^l(\tilde{u}) = V_l(\tilde{u}) - \frac{1}{\beta} \ln[P_i^l(\tilde{u})] + C_i^l. \qquad (9)$$

The free parameters C_i^l are configuration independent.

The second step consists of simulation of the model with any procedure whose transition probability $\tilde{T}_{\{i\}}(u \to u')$ satisfies the detailed balance condition with respect to the new Hamiltonian, i.e.,

$$e^{-\beta \tilde{\mathcal{H}}_{\{i\}}(u)} \tilde{T}_{\{i\}}(u \to u') = e^{-\beta \tilde{\mathcal{H}}_{\{i\}}(u')} \tilde{T}_{\{i\}}(u' \to u). \qquad (10)$$

After completing the second step, a new configuration u' is arrived at, the original Hamiltonian is restored, and the process is repeated. Equations (9) and (10) ensure that the whole procedure satisfies the detailed balance condition, with respect to the original Hamiltonian (for the proof see [10]), but ergodicity needs to be proved for each application separately.

We give now two types of modifications to the Hamiltonian. Consider a term $V_l(u)$ that can take M distinct energy values E_i, $i = 1, \ldots, M$. The first is the deletion operation, used by SW, which eliminates the interaction $V_l(\tilde{u})$ that gets replaced by

$$\tilde{V}_d^l(\tilde{u}) = 0, \qquad (11)$$

for any configuration $\tilde{u}$. To get this, we must have [see Eq. (9)]

$$P_d^l(u) = e^{\beta[V_l(u) + C_d^l]}. \qquad (12)$$

The constant C_d^l must be chosen so that $P_d^l(u) \leq 1$ for any u.

The second modification is a *"generalized" freezing* operation which we will later use in our scheme for the Ashkin-Teller model. Its probability is

$$P_g^l(u) = \begin{cases} e^{\beta[V_l(u) + C_g^l]} & \text{if } V_l(u) = E_m \text{ and } 1 \leq m \leq \mu \\ 0 & \text{if } V_l(u) = E_m \text{ and } \mu < m \leq M, \end{cases} \qquad (13)$$

and the modified interaction according to (9) is

$$\tilde{V}_g^l(\tilde{u}) = \begin{cases} 0 & \text{if } V_l(\tilde{u}) = E_m \text{ and } 1 \leq m \leq \mu \\ \infty & \text{if } V_l(\tilde{u}) = E_m \text{ and } \mu < m \leq M. \end{cases} \qquad (14)$$

Again the constant C_g^l must be chosen so that $P_g^l(u) \leq 1$ for any u. This type of modification allows free movements between some of the states u of $V_l(u)$ which did not have the same energy in the original Hamiltonian, and we will use it for the Ashkin-Teller model in Sec. III C. Perhaps it can be useful in general, in cases where the V_l's can have more than two possible energies. This type of operation has been used, for example, in [14].

A particular case of this operation is obtained if $\mu = 1$, and we set

$$C_1^l = \frac{1}{\beta} \ln(p_1) - E_1. \qquad (15)$$

Substitution into (13) yields

$$P_1^l(u) = \begin{cases} p_1 & \text{if } V_l(u) = E_1 \\ 0 & \text{otherwise}, \end{cases} \qquad (16)$$

and from (14), the modified interaction takes the form

$$\tilde{V}_1^l(\tilde{u}) = \begin{cases} 0 & \text{if } V_l(\tilde{u}) = E_1 \\ \infty & \text{otherwise}. \end{cases} \qquad (17)$$

This operation assigns infinite energy to any configuration $\tilde{u}$ for which $V_l(\tilde{u}) \neq E_1$. That is, in the ensuing simulation the interaction V_l is frozen at energy E_1. Freezing of V_l is assigned with probability p_1, and only when $V_l(u) = E_1$. It is easy to see that the SW freezing is precisely of this form.

C. Cluster method for the Ashkin-Teller model

We describe now the cluster algorithm we devised for the GAT model and the considerations that led us to it. It consists of a freeze-delete scheme which generates noninteracting clusters of aligned σ spins and of aligned τ spins. We will phrase it in terms of the general scheme described in Sec. III B.

The first decision one needs to make when coming to design a freeze-delete scheme is the choice of the basic interaction term V_l. Our choice is to associate all interactions that reside on an edge $\langle jk \rangle$ of the lattice to one V_l:

$$V_l = -[K_\sigma \sigma_k \sigma_j + K_\tau \tau_k \tau_j + L \sigma_k \tau_k \sigma_j \tau_j]. \qquad (18)$$

Since the model is invariant under any permutation of K_σ, K_τ, and L (to make this symmetry explicit, define a new Ising spin $s_j = \sigma_j \tau_j$ along with the constraint $s_j \sigma_j \tau_j = 1$), it is possible to choose $L \leq K_\sigma, K_\tau$ (the reason for this choice will become clear later). The interaction V_l depends on four independent Ising spins that

can have 16 states. Every V_l can take one of four possible energy values:

$$
\begin{aligned}
E_1 &= -K_\sigma - K_\tau - L \ , \\
E_2 &= K_\sigma - K_\tau + L \ , \\
E_3 &= -K_\sigma + K_\tau + L \ , \\
E_4 &= K_\sigma + K_\tau - L \ .
\end{aligned}
\tag{19}
$$

Every energy is fourfold degenerate; we denote by u_i all (four) states for which $V_l = E_i$. Let u_i represent the state of the spins σ_j, σ_k, τ_j, τ_k of a particular pair of nearest-neighbor sites $\langle j, k \rangle$. Four representative possible states are depicted in Fig. 2. u_1 is a ground state in which both the σ and the τ bonds are satisfied. u_2 and u_3 are excited states in which one Ising bond is broken and the other is satisfied. Our choice $L \leq K_\sigma, K_\tau$ makes u_4 the highest energy state [15], in which both Ising bonds are broken.

The general philosophy of our freeze-delete scheme is as follows. We wish to build clusters of σ spins and of τ spins since we know that these clusters are the basic excitations of the model. Clusters of $s = \sigma\tau$ spins are not important because L, the bond between them, is the weakest. In order to build clusters, one needs to freeze parallel σ spins to each other and parallel τ spins to each other, and delete the bonds between antiparallel spins. For example, in the state u_2 we wish to freeze the bond between the τ spins and delete the bond between the σ spins. This consideration leads us to include in our scheme two operations. The operation which we will identify as $i = 2$ freezes the bond between τ_j and τ_k and deletes the bond between σ_j and σ_k (see Fig. 3). According to the discussion above, we want to perform this operation with some probability $p_2 \neq 0$ for the state u_2, and with probability 0 for the states u_3 and u_4. This operation allows one to move from u_2 to u_1, so in order to maintain detailed balance, we must perform it with some nonvanishing probability $q_2 \neq 0$ on u_1 too. All this is achieved by assigning, in the modified interaction term $\tilde{V}_2^l$, infinite energy to u_3, u_4 and zero energy to u_1, u_2. This modification is precisely of the form of Eq. (14) so it is of the generalized freezing type described at the end of Sec. III B and hence the probabilities must be [see Eq. (13)] [16]

$$
P_2^l(u) = \begin{cases} e^{V_l(u)+C_2} & \text{if } V_l(u) = E_2 \text{ or } V_l(u) = E_1 \\ 0 & \text{if } V_l(u) = E_3 \text{ or } V_l(u) = E_4 \ . \end{cases}
\tag{20}
$$

The choice of C_2 will set the values of p_2 and q_2:

$$
p_2 = e^{V_l(u_2)+C_2}, \quad q_2 = e^{V_l(u_1)+C_2} \ .
\tag{21}
$$

	u_1	u_2	u_3	u_4
$\sigma_j \ \sigma_k$	+ +	+ −	+ +	+ −
$\tau_j \ \tau_k$	− −	+ +	+ −	− +

FIG. 2. States of spins at a pair of nearest-neighbor sites j, k. Each state u_i represents one out of four states with the same energy E_i.

		$i=1$	$i=2$	$i=3$	$i=d$
u_1	+ + / + +	p_1	q_2	q_3	$P_d(u_1)$
u_2	+ − / + +	0	p_2	0	$P_d(u_2)$
u_3	+ + / + −	0	0	p_3	$P_d(u_3)$
u_4	+ − / − +	0	0	0	$P_d(u_4)=1$

FIG. 3. Freeze-delete scheme for the AT model. The four states of two neighboring sites are denoted in the left column; four freeze-delete operations are denoted in the top row, and the respective probabilities are denoted in the table. A frozen bond is denoted by a double thick line, while a deleted bond is denoted by a blank.

The operation which we identify as $i = 3$ freezes the bond between σ_j and σ_k and deletes the bond between τ_j and τ_k. It follows the same logic as the operation $i = 2$ and its probability is

$$
P_3^l(u) = \begin{cases} e^{V_l(u)+C_3} & \text{if } V_l(u) = E_3 \text{ or } V_l(u) = E_1 \\ 0 & \text{if } V_l(u) = E_2 \text{ or } V_l(u) = E_4 \ . \end{cases}
\tag{22}
$$

Again we define the probabilities:

$$
p_3 = e^{V_l(u_3)+C_3}, \quad q_3 = e^{V_l(u_1)+C_3} \ .
\tag{23}
$$

The third operation we wish to perform is a freezing operation of the form defined in Sec. III B, Eqs. (16) and (17). In our case it freezes the bond between σ_j and σ_k and the bond between τ_j and τ_k, so it also builds the τ and σ clusters. We will denote it by $i = 1$ and its probability is

$$
P_1^l(u) = \begin{cases} p_1 & \text{if } V_l(u) = E_1 \\ 0 & \text{otherwise} \ . \end{cases}
\tag{24}
$$

The fourth and last operation is the deletion operation defined in Sec. III B, Eqs. (11) and (12), with the probability

$$
P_d^l(u) = e^{V_l(u)+C_d}.
\tag{25}
$$

These four operations define completely our freeze-delete scheme, summarized by Fig. 3. In the first column, in each row, one or two representative configurations of u_i appear, depicting the state of $\sigma_j, \tau_j, \sigma_k$, and τ_k. In the first row the four modified interactions $\tilde{V}_i^l$ are depicted. An upper double line, for example, denotes a frozen bond between σ_j and σ_k, a blank denotes a deleted bond. We have yet to determine the constants C_2, C_3, p_1, C_d. In the case of u_4, we want to delete the bond between the τ spins and the bond between the σ spins with probability 1, as we never want an unsatisfied bond to be frozen, so we choose $C_d = -E_4$, to get $P_d(u_4) = 1$. Having chosen C_d, all deletion probabilities are set. The rest of the

constants are determined by the normalization condition (7), i.e.,

$$p_2 = 1 - P_d(u_2) \ ,$$
$$p_3 = 1 - P_d(u_3) \ . \qquad (26)$$

Consequently, Eqs. (21) and (23) determine the constants C_2, C_3 and the probabilities q_2, q_3. Finally, p_1 is determined again by the normalization condition

$$p_1 = 1 - P_d(u_1) - q_2 - q_3. \qquad (27)$$

For completeness we will list all the probabilities of our scheme which follow from Eqs. (20)–(27):

$$
\begin{aligned}
P_d(u_i) &= e^{E_i - E_4} & \forall i \ , \\
p_i &= 1 - e^{E_i - E_4} \ , & i = 2, 3 \ , \\
q_i &= p_i e^{E_1 - E_i} \ , & i = 2, 3 \ , \\
p_1 &= 1 - P_d(u_1) - q_2 - q_3 \ .
\end{aligned}
\qquad (28)
$$

Checking that all the probabilities of the scheme fulfill the condition $0 < P_i(u) < 1$ for all i and u is trivial for all probabilities except for p_1 which is a bit more laborious, but still straightforward.

We have described how we generate a new Hamiltonian of noninteracting clusters of σ spins and of τ spins. Now we have to choose some legitimate MC procedure to simulate this Hamiltonian. We chose to do it in a similar fashion to the single-cluster algorithm of Wolff [9], but the SW version is equally applicable. To be more explicit, we choose a site j of the lattice at random and a random spin, either σ_j or τ_j (the choice of σ or τ was done with probability $\frac{1}{2}$ which is sensible in the case $K_\sigma = K_\tau$, but in general any probability is acceptable, and an optimal choice can be made to minimize the autocorrelation time τ). This spin will belong to the cluster we will flip, and the cluster will contain only τ (σ) spins if we initially chose τ_j (σ_j). We perform the freeze-delete operations only on bonds belonging to the surface of the cluster. For example, suppose at some stage τ_j was joined to the cluster, and suppose that the term $V_{j,k}$ was modified according to $i = 3$, then τ_k will not be joined to the cluster and we need not perform freeze-delete operations on the other three bonds connecting the site k. If it is modified according to $i = 1$ or $i = 2$ then τ_k is joined to the cluster. Since our cluster algorithm fits into the general scheme, we do not need to prove detailed balance, while ergodicity is ensured by the deletion operation.

We have explained the reasoning behind our algorithm. We believe that the clusters we build of parallel σ spins and of parallel τ spins are the basic excitations of the model. We therefore believe that it will be efficient. It is encouraging to notice that in the decoupled Ising subspace and the four-state Potts subspace our freeze-delete scheme is identical with SW's freeze-delete operations for these models. In the Appendix we show that our algorithm is equivalent to the idea of embedding into the AT model an Ising model and simulating it by Wolff's single-cluster procedure. In Sec. IV we present numerical evidence for our algorithm's efficiency.

D. "Naive" SW option

We find it illuminating to compare our algorithm to a cluster algorithm which one could regard as the naive generalization of the SW method to the AT model. Such an algorithm would define V_l in the same manner as our scheme does [see Eq. (18)]. For each u_i, the bonds between the two neighboring sites get either deleted [with our $P_d(u_i)$] or frozen with $P_f(u_i) = 1 - P_d(u_i)$. Figure 4 can clarify how this scheme fits into the general one of Sec. III B and how it compares with our scheme. Since it fits into the general scheme we do not need to prove detailed balance, while the deletion operation ensures its ergodicity except for $T = 0$.

This scheme also generates clusters of σ spins and clusters of τ spins, but with the "naive" scheme, antiparallel Ising spins can be found in the same cluster, so this scheme does not identify the elementary excitations of the model. Moreover, the σ clusters' structure is forced to match the τ clusters (and vice versa). This is totally unphysical since in practice it is energetically favorable for the two cluster structures not to match (for $L < K_\sigma, K_\tau$). At $T = 0$ this scheme could freeze the whole lattice into a single cluster even when it is not in the ground state. This is in contrast to the ability of our scheme to relax excitations on any scale even at $T = 0$ (in other words, at $T = 0$ our algorithm will bring any initial configuration to the ground state in a short time). At a finite temperature the "naive" SW scheme will produce clusters that are too large. In Sec. IV A we list results of simulations using the single-cluster (1C) version of this naive SW scheme, which show that it is indeed much less efficient than our algorithm.

IV. SIMULATIONS OF THE AT CRITICAL LINE

In addition to checking the efficiency of our method at the AT critical line, we also wanted to check a prediction made by Li and Sokal [6], who have proved a rigorous lower bound

σ q τ τ		i=1	i=2	i=3	i=d
u_1	+ + / + +	$P_f(u_1)$	0	0	$P_d(u_1)$
u_2	+ − / + +	0	$P_f(u_2)$	0	$P_d(u_2)$
u_3	+ + / + −	0	0	$P_f(u_3)$	$P_d(u_3)$
u_4	+ − / − + / + − / + −	0	0	0	$P_d(u_4)=1$

FIG. 4. "Naive" SW freeze-delete scheme for the AT model. The four states of two neighboring sites are denoted in the left column; four freeze-delete operations are denoted in the top row, and the respective probabilities are denoted in the table. A frozen bond is denoted by two thick lines, while a deleted bond is denoted by a blank.

$$z_{\text{SW}} \geq \alpha/\nu \qquad (29)$$

for the dynamical critical exponent of the SW algorithm for the ferromagnetic q-state Potts model. More precisely, the Li-Sokal bound is $\tau_{\text{exp,int}} \geq \text{const} \times C_h$, from which Eq. (29) follows (for definitions of $\tau_{\text{exp/int}}$ and $z_{\text{exp/int}}$ see [17]). This bound relates dynamics to the static properties of a model and is thus of great importance. Their proof is for both z_{int} and z_{exp}. We wanted to check whether this bound holds all along the AT critical line, which connects the decoupled Ising critical point with $\frac{\alpha}{\nu} = 0$ at one end to the four-state Potts critical point with $\frac{\alpha}{\nu} = 1$ at the other end. At both of these points our algorithm is identical to Wolff's 1C version of the SW method for Potts models. The Li-Sokal bound was proved for the SW dynamics, but perhaps it is valid for Wolff's 1C dynamics as well (at least for $d = 2$). We are unaware of any rigorous proof for that, but our results for z for $q = 2$ and 4 seem to indicate $z_{1C} = z_{\text{SW}}$ and are in accordance with previous results for the Ising critical point (see [8] and [18]). Besides estimating the dynamical exponents, we have estimated, using finite-size scaling, the critical exponents $\frac{\alpha}{\nu}$ and $\frac{\gamma}{\nu}$ along the AT critical line. As stated in Sec. II the AT critical line is given by

$$Z = 1 - 2X. \qquad (30)$$

Our measurements were done at the decoupled Ising critical point $\mathbf{X}_0$ and at the four-state Potts critical point $\mathbf{X}_4$ (see Fig. 1 and exact definition in Sec. II). Three additional measurements were carried out at three intermediate equidistant points in the $X - Z$ plane, on the AT critical line. So a total of five measurements were done at the points

$$\mathbf{X}_i = \mathbf{X}_0 + \tfrac{i}{4}(\mathbf{X}_4 - \mathbf{X}_0), \quad i = 0, \dots, 4. \qquad (31)$$

All the points $\mathbf{X}_i$ are marked in Fig. 1. We simulated lattices with periodic boundary conditions of up to size 128×128, and up to 5×10^5 clusters were flipped for each lattice size.

We calculated the energy [19] per site:

$$E = \langle E \rangle = -\frac{1}{L^2} \left\langle \sum_{\langle k,j \rangle} [K\sigma_k\sigma_j + K\tau_k\tau_j + L\sigma_k\tau_k\sigma_j\tau_j] \right\rangle, \qquad (32)$$

where L is the linear lattice size, and the angular brackets denote the usual thermal MC average. The specific heat

per site follows from the energy fluctuations

$$C = \frac{C_h}{k_B} = L^2(\langle E^2 \rangle - \langle E \rangle^2). \qquad (33)$$

We calculated the magnetic susceptibility of the σ spins defined as [19]

$$\chi = L^2 \langle M^2 \rangle, \quad M = \frac{1}{L^2} \sum_k \sigma_k \qquad (34)$$

and also measured c, the size of the cluster flipped at each step, and calculated $\langle c \rangle$. There is a connection between the size of the clusters and the susceptibility, [20] which is common for algorithms which build noninteracting clusters of spins (if all spins in a cluster have the same value). For 1C algorithms it has the simple form [9]

$$\chi = \langle c \rangle. \qquad (35)$$

To get the dynamic properties we calculated the time-dependent autocorrelation functions $\phi_E(t)$ and $\phi_\chi(t)$ defined as

$$\phi_A(t) = \frac{\langle A(0)A(t) \rangle - \langle A \rangle^2}{\langle A^2 \rangle - \langle A \rangle^2}, \qquad (36)$$

where A stands for E or χ. A typical plot of $\phi_\chi(t)$, measured from the X_2 model at $L = 128$, is presented in Fig. 6.

A. MC results and discussion

1. Susceptibility and specific heat

According to finite-size scaling theory [21] one expects

$$\chi \sim L^{\frac{\gamma}{\nu}}, \quad C \sim L^{\frac{\alpha}{\nu}} \qquad (37)$$

for large enough L. Fitting our measurements of χ to Eq. (37) for lattice sizes $L \geq 16$ fits the exact universal value $\frac{\gamma}{\nu} = \frac{7}{4}$ within errors (see Table I) as expected. Our measurements confirm the equality (35), where $\langle \chi \rangle$ and $\langle c \rangle$ have an error of the same magnitude. Plots of $\log C$ vs $\log L$ for the five models can be seen in Fig. 5. From linear fits to the log-log plots, estimates for $\frac{\alpha}{\nu}$ are obtained which do not agree with the exact known values. For comparison with the exact values see Table I and Fig. 13. The differences may be due to finite-size effects and corrections to scaling. An exception to this

TABLE I. Results from the AT critical line. The errors in parentheses are only the statistical errors of the fit in our fitting interval. They do not include systematic errors stemming from finite-size effects and corrections to scaling.

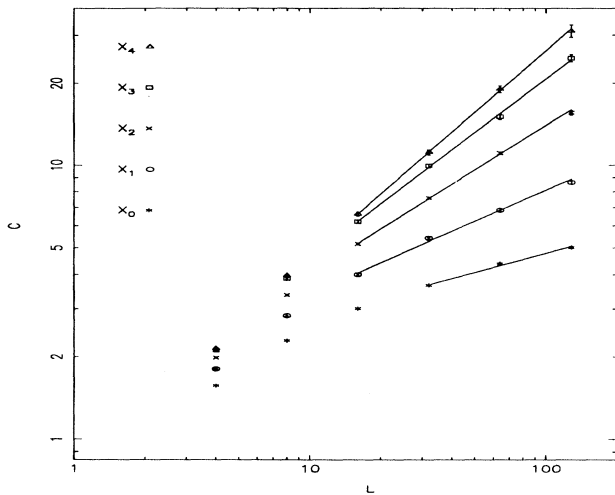|	$\frac{\gamma}{\nu}$	$\frac{\alpha}{\nu}$	$z_{\text{int,E}}$	$z_{\text{exp,E}}$	$z_{\text{int,}\chi}$	$z_{\text{exp,}\chi}$	$\frac{\alpha}{\nu}$ exact
X_0	1.751(4)	0.23(1)	0.26(3)	0.23(2)	0.13(1)	0.267(4)	0
X_1	1.751(1)	0.38(2)	0.396(5)	0.40(3)	0.273(3)	0.47(2)	0.2096
X_2	1.752(3)	0.542(8)	0.61(1)	0.68(3)	0.532(5)	0.62(1)	0.4182
X_3	1.763(6)	0.655(13)	0.719(2)	0.74(2)	0.717(8)	0.72(3)	0.6383
X_4	1.752(4)	0.747(3)	0.92(1)	0.99(5)	0.99(3)	0.94(3)	1

FIG. 5. A log-log plot of specific heat vs linear lattice size at the five critical models $X_0, \ldots, X_4$. The solid lines are the linear fits. The critical exponents $\frac{\alpha}{\nu}$ are given in Table I.
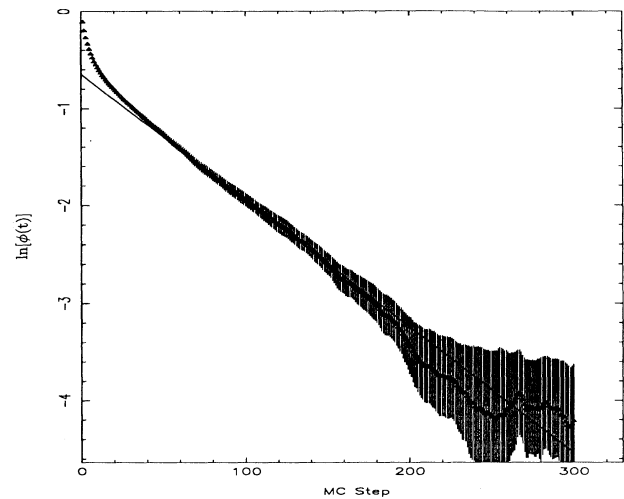

FIG. 6. The ln of the time autocorrelation function of the susceptibility $\phi_\chi(t)$, measured for the X_2 model at $L = 128$. The unit of time is a single MC step or a single-cluster flip. The vertical lines are the error bars. The exponential fit is also shown.

mismatch is the decoupled Ising point for which the value of $\frac{\alpha}{\nu} = 0$ fits nicely according to the semilog plot in Fig. 8 (for completeness we also quote in Table I an estimate for $\frac{\alpha}{\nu}$ at X_0 from the log-log fit of Fig. 5). The slope of the C curves at X_1 and X_2 does tend to a lower value with increasing lattice size. The four-state Potts model is known to have a logarithmic correction to scaling: $C \sim L/\ln^{3/2} L$ [6]; in any case, our result $\frac{\alpha}{\nu} = 0.747(3)$ agrees with previous MC results obtained from lattices of sizes up to $L = 256$ [6].

2. τ_{exp} and τ_{int}

In order to check the efficiency of our algorithm and to see whether the Li-Sokal bound holds, we measured τ_{exp} and τ_{int} for both the energy E and the susceptibility χ. This was done by the following procedure. $\phi_{E/\chi}(t)$ was plotted on a semilog plot, the unit of time being a single cluster flip. τ_{exp} was then consistently estimated by the slope of $\ln \phi(t)$ in a window from τ_{exp} to $3\tau_{exp}$. A typical example for $\phi(t)$ and the linear fit to extract τ_{exp} can be seen in Fig. 6. To calculate τ_{int}, we integrated numerically $\phi(t)$ in the interval from $t = 0$ up to $t \approx 1.5\tau_{exp}$ and estimated the tail $(t > 1.5\tau_{exp})$ of $\phi(t)$ by the exponential fit. The error of both τ's was estimated from repeated experiments when the statistics were large enough and otherwise from subjective estimates from the fluctuations in τ_{exp} in the window $\tau_{exp} - 3\tau_{exp}$. Each MC step, or cluster flip, involves on the average the flipping of $\langle c \rangle$ spins. Since the natural unit of time is a sweep over all spins of the lattice, we multiplied τ by $\langle c \rangle L^{-2}$.

In Figs. 7–12 we plot measured values of τ_{exp} and τ_{int} for E and χ as a function of the linear lattice size L. In Figs. 7–10 we display the measurements of τ_χ at different models X_i separately, while in Figs. 11 and 12 we display $\tau_{exp,E}$ and $\tau_{int,E}$ for the five models jointly. Fitting our results to the forms

$$\tau_{exp} \sim L^{z_{exp}} \quad ; \quad \tau_{int} \sim L^{z_{int}} , \qquad (38)$$

for large enough L, yielded the dynamical critical exponents listed in Table I. Figures 7 and 8 are both from the decoupled Ising model point X_0. In Table I we list values of z obtained from the power-law fits of Fig. 7. None the less, comparing these fits to the $\log L$ fits of Fig. 8, we feel that the latter ones are better and that the true value is $z = 0$ for all four autocorrelation times
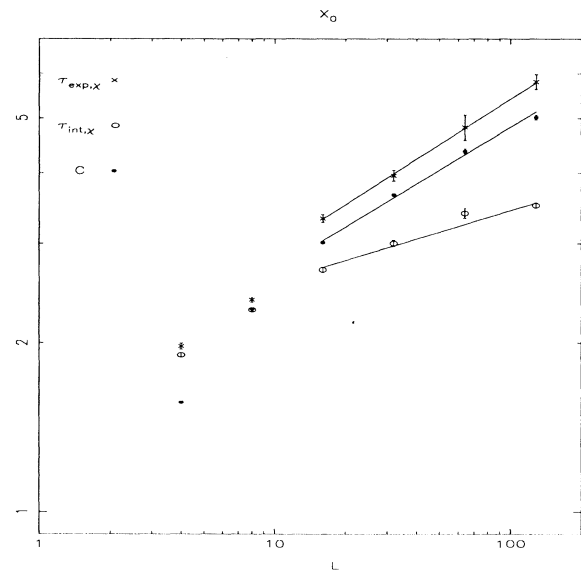

FIG. 7. Log-log plots of autocorrelation times τ_χ and the specific heat C as a function of linear lattice size L, at the decoupled Ising point X_0. The values of z listed in Table I are the slopes of the fits.
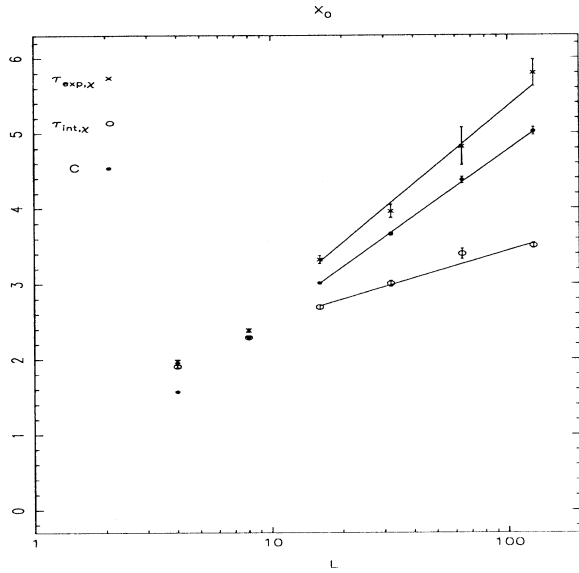
FIG. 8. Semilog plots of autocorrelation times τ_χ and the specific heat C as a function of linear lattice size L, at the decoupled Ising point X_0. This fit seems to be better than the log-log fit, yielding values of $z = \frac{\alpha}{\nu} = 0$.

FIG. 10. Log-log plots of autocorrelation times τ_χ and the specific heat C as a function of linear lattice size L, at the point X_3. The values of z listed in Table I are the slopes of the fits.

τ. At X_0 (and all along the $Z = X^2$ or $L = 0$ line) our algorithm is identical with the Wolff 1C algorithm for the Ising model. Our results for $z_{\rm int}$ from the fits in Fig. 7 are consistent with those of Wolff [18]. A $\log L$ behavior has been measured for $\tau_{\rm exp,M}$ and $\tau_{\rm int,M}$ for SW dynamics in [22].

At the point X_4, the four-state Potts critical point (and all along the $X = Z$ line), our algorithm is again identi-
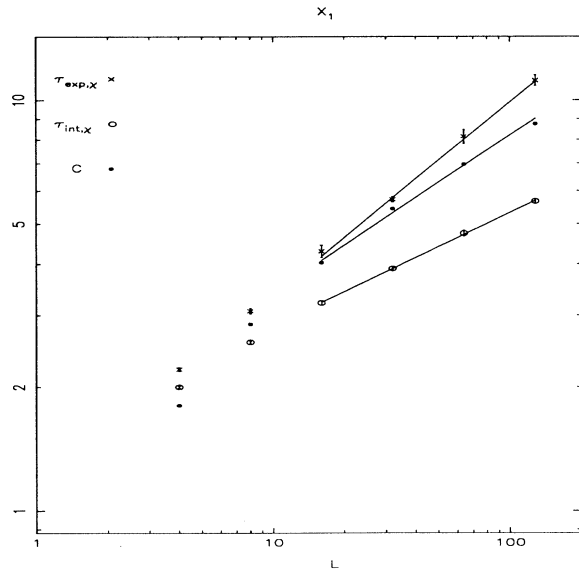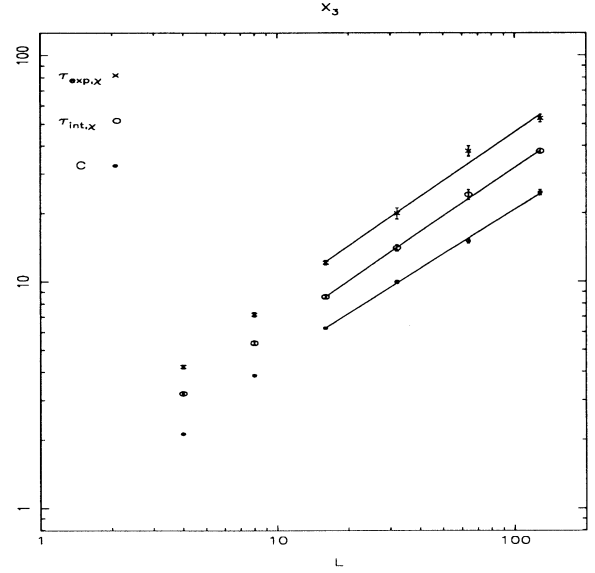
cal with the Wolff 1C algorithm for the four-state Potts model. Our result for $z_{\rm int,E}$, obtained from the fit in Fig. 12, is consistent with the result of Ref. [6] for SW dynamics. This result can be added to the accumulating evidence, indicating $z_{\rm SW} = z_{\rm Wolff}$ for $d = 2$.

The continuous variation of z along the AT critical line can be explicitly seen in Figs. 11 and 12 where we plot $\tau_{\rm exp,E}$ and $\tau_{\rm int,E}$ for the five points X_i. In Fig. 12 we also plot results from simulations at X_2 using the Metropolis method and the "naive" SW method described in Sec. III D. For the Metropolis method, we measured a value of $z_{\rm int,E} = 1.64(8)$, which should be compared

FIG. 9. Log-log plots of autocorrelation times τ_χ and the specific heat C as a function of linear lattice size L, at the point X_1. The values of z listed in Table I are the slopes of the fits.

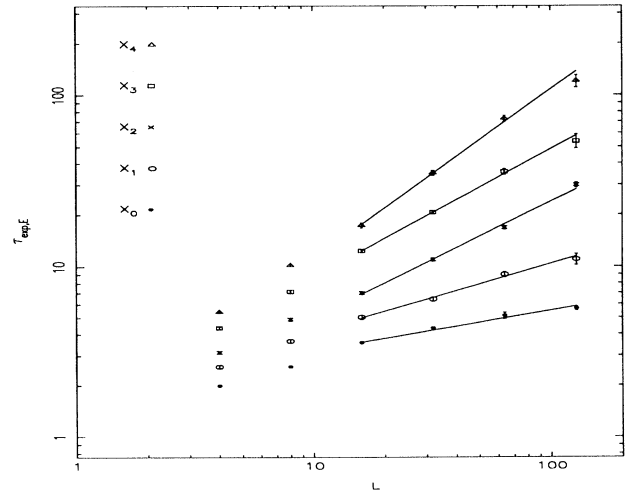

FIG. 11. A log-log plot of $\tau_{\rm exp,E}$ vs linear lattice size at the five critical models $X_0, \ldots, X_4$. The solid lines are the linear fits. The critical exponents $z_{\rm exp,E}$ are given in Table I.
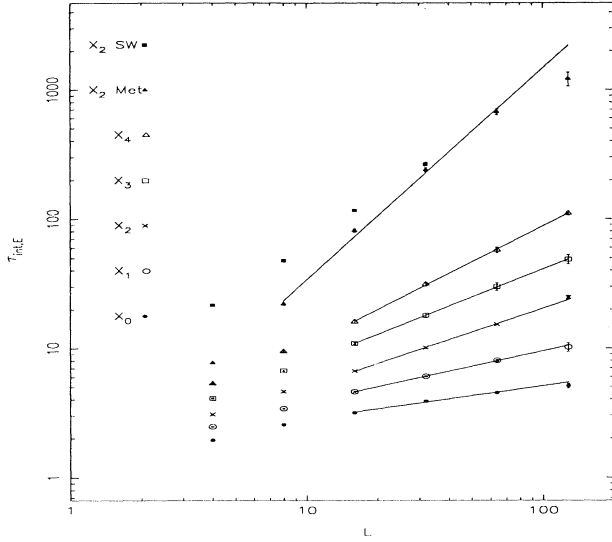
FIG. 12. Log-log plots of $\tau_{\text{int,E}}$ as a function of linear lattice size L, at all five points X_i, $i = 0, \ldots, 4$. The continuous variation of the slope z along the AT critical line is easily seen. Metropolis results and "naive" SW results for X_2 are also plotted.
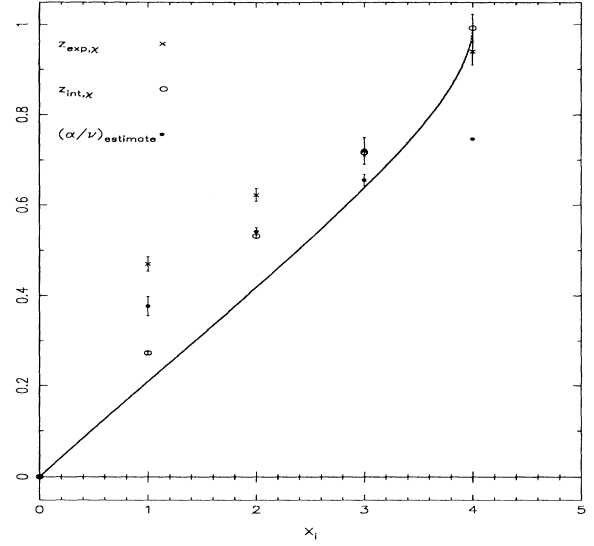

FIG. 13. Comparison of z and $\frac{\alpha}{\nu}$ at the five points X_i, $i = 0, \ldots, 4$. The line denotes the exact value of $\frac{\alpha}{\nu}$, while our estimated values for the five models are denoted by full circles. Values of $z_{\text{int},\chi}$ ($z_{\text{exp},\chi}$) are denoted by empty circles (crosses).

with $z_{\text{int,E}} = 0.542(8)$ using our method. The naive SW method almost froze the whole lattice, into a single cluster, the size of which was almost independent of the lattice size L. For example, the average cluster size for $L = 32$ was $\frac{\langle c \rangle}{L^2} = 0.868(2)$, as compared with $\frac{\langle c \rangle}{L^2} = 0.481(1)$ using our algorithm. We found determination of τ for the naive SW method of lattice sizes $L > 32$ so time consuming that it was impractical. The advantage of our method is clearly demonstrated. The three methods yielded the same results for the static observables (within errors).

The main question we wish to answer is whether the Li-Sokal bound is fulfilled and whether it is sharp. In Fig. 13 we compare, for the five points X_i, $i = 0, \ldots, 4$, the exact values of $\frac{\alpha}{\nu}$ and our estimated values of $\frac{\alpha}{\nu}$, $z_{\text{int},\chi}$, and $z_{\text{exp},\chi}$. We see that the rise of our estimated values of z follows that of $\frac{\alpha}{\nu}$, from the decoupled Ising point X_0 to the four-state Potts point X_4. Except for the point X_4 the Li-Sokal bound is fulfilled with respect to $\left(\frac{\alpha}{\nu}\right)_{\text{exact}}$. The anomalously low values for z at X_4 are probably caused by the multiplicative logarithmic correction for C described in the discussion of the specific-heat results. This explanation has been suggested by Li and Sokal [6] to explain the low value of z_{SW} for this model. From Table I we see that $z \geq \left(\frac{\alpha}{\nu}\right)_{\text{estimate}}$ with only one exception at the point X_1 (see Fig. 9). The fact that $z_{\text{int},\chi} < \left(\frac{\alpha}{\nu}\right)_{\text{estimate}}$ could be due to a difference in the finite-size corrections or in the corrections to scaling. $\tau_{\text{int},\chi}$ is probably less influenced by one or both of these two factors, in comparison with the large difference between $\left(\frac{\alpha}{\nu}\right)_{\text{estimate}}$ and $\left(\frac{\alpha}{\nu}\right)_{\text{exact}}$. At the point X_3, where the difference between $\left(\frac{\alpha}{\nu}\right)_{\text{estimate}}$ and $\left(\frac{\alpha}{\nu}\right)_{\text{exact}}$ is the smallest, no anomalies in the Li-Sokal bound occur (see Fig. 10).

We conclude that the Li-Sokal bound is fulfilled in a moderately sharp manner. The smallest z (which is $z_{\text{int},\chi}$) fulfills $\left(\frac{\alpha}{\nu}\right)_{\text{estimate}} \leq z_{\text{int},\chi} \leq \left(\frac{\alpha}{\nu}\right)_{\text{estimate}} + 0.1$ (not including the anomalies of the models X_4 and X_1 discussed above). Note that we are comparing estimated values of z with *estimated* values of $\frac{\alpha}{\nu}$. This is the correct comparison to make, assuming that τ and C have similar finite-size corrections and similar corrections to scaling.

V. SUMMARY AND DISCUSSION

The correct identification of the basic excitations of the model lead, along with the guidelines of a general scheme for cluster algorithms, to the construction of a cluster algorithm for the AT model. The algorithm was shown to be identical with the one obtained by embedding Ising spins into the AT model. Our algorithm is suitable for spatially varying coupling constants under the restriction $L^{k,j} \leq K_\sigma^{k,j}, K_\tau^{k,j}$ everywhere on the lattice. We are currently carrying out intensive simulations of a random-bond version of the AT model.

The dynamical behavior of the cluster algorithm was examined on the AT critical line. Critical slowing down of the algorithm ($0 \leq z < 1.0$) was found to be significantly smaller than that of the standard Metropolis method. A continuous variation of the dynamical exponent z along the AT critical line was seen, along with continuous variation of the static exponent $\frac{\alpha}{\nu}$. The Li-Sokal bound $z \geq \frac{\alpha}{\nu}$, that was proved only for q-state Potts models with SW dynamics, is satisfied for the AT model with single-cluster dynamics. The bound is moderately sharp. Another result is $z_{\text{Wolff}} = z_{\text{SW}}$ for the four-state Potts model, which

can be added to the accumulating results indicating that in two dimensions $z_{\text{Wolff}} = z_{\text{SW}}$.

APPENDIX: ISING EMBEDDING

Our algorithm can also be seen from a totally different point of view; as an example of an embedding algorithm [9,11]. The main idea is to embed into the AT model an Ising model of space-dependent couplings J_{jk} and simulate it using the SW or Wolff procedure for the Ising model. To be explicit, consider the Hamiltonian (1), and take the τ variables as fixed, so we can write

$$\mathcal{H} = \mathcal{H}_1 + \mathcal{H}_2 = - \sum_{\langle k,j \rangle} (K_\sigma + L\tau_k \tau_j)\sigma_k \sigma_j$$
$$- \sum_{\langle k,j \rangle} K_\tau \tau_k \tau_j. \qquad \text{(A1)}$$

$\mathcal{H}_2$ represented by the second sum is a constant, and remembering that we chose $L < K_\sigma, K_\tau$, $\mathcal{H}_1$ is a ferromagnetic Ising model in the σ variables with couplings $J_{jk} = K_\sigma + L\tau_k \tau_j$. Simulating this Ising model with any procedure that will maintain detailed balance with respect to $\mathcal{H}_1$ and will not change the value of $\mathcal{H}_2$ will also maintain detailed balance with respect to $\mathcal{H}$. So we can use, for example, the SW or Wolff procedure for the Ising model, explained in Sec. III B. This by itself will maintain detailed balance but will not be ergodic since the τ variables will not be updated. Obviously to update the τ variables the same process should be repeated, holding the σ variables fixed and simulating an Ising Hamiltonian with the τ variables. To summarize, a possible procedure would go as follows: Choose at random whether to em-

bed into the AT Hamiltonian an Ising Hamiltonian in the σ spins or in the τ spins. Pick a random site in the lattice, grow a cluster of σ or τ spins using the Wolff (1C) procedure using the Ising Hamiltonian and flip it. As we will now show, this process (we will call it IE, denoting Ising embedding) is exactly equivalent to the AT algorithm (ATA) we suggested in Sec. III C .

A first argument which is really sufficient goes as follows. Suppose we perform the IE freeze-delete scheme for the σ spins for the whole lattice, viewing the τ spins as constant. We could say we have generated a new Hamiltonian of noninteracting clusters of σ spins and which assigns infinite energy to any configuration which differs from the current configuration of τ spins. Now suppose we perform the ATA freeze-delete scheme for the whole lattice. Then we get a new Hamiltonian of noninteracting clusters of σ spins and of τ spins, but if at this stage we decide to flip only σ spins, then in practice we have assigned infinite energy to any configuration which differs from the current configuration of τ spins. In practice we flip only one cluster of the σ spins, so what we do is identical to the Wolff (1C) version of the IE. Since both processes maintain detailed balance using the same new Hamiltonian they must do so using the same probabilities (the general scheme in Sec. III B shows a one to one correspondence between probabilities and the new Hamiltonian), which completes our argument that the two methods are actually identical.

One can, of course, check that the probabilities of the two procedures are the same. For example, denote the probability to delete a bond between σ spins in ATA as $P_{d,\sigma}^{\text{ATA}}(u)$ for u_1, for example, from Fig. 3:

$$P_{d,\sigma}^{\text{ATA}}(u_1) = P_d(u_1) + q_2 = e^{E_1 - E_2} = e^{-2(K_\sigma + L)}. \qquad \text{(A2)}$$

Now with the Ising embedding algorithm, according to (5)

$$P_{d,\sigma}^{\text{IE}}(u) = e^{-(K_\sigma + L\tau_k \tau_j)(\sigma_k \sigma_j + 1)}, \qquad \text{(A3)}$$

so $P_{d,\sigma}^{\text{IE}}(u_1) = P_{d,\sigma}^{\text{ATA}}(u_1)$. This check can be carried out for all u_i.

* Electronic address: shai@elect1.weizmann.ac.il
† Electronic address: fedomany@WEIZMANN.weizmann.ac.il
[1] J. Ashkin and E. Teller, Phys. Rev. **64**, 178 (1943).
[2] F.Y. Wu and K.Y. Lin, J. Phys. C **7**, L181 (1974).
[3] E. Domany and E.K. Riedel, Phys. Rev. B **19**, 5817 (1979).
[4] R.V. Ditzian, J.R. Banavar, G.S. Grest, and L.P. Kadanoff, Phys. Rev. B **22**, 2542 (1980).
[5] B. Nienhuis, in *Phase Transitions and Critical Phenomena*, edited by C. Domb and J.L. Lebowitz (Academic, New York, 1987), Vol. 11.
[6] X.J. Li and A.D. Sokal, Phys. Rev. Lett. **63**, 827 (1989).
[7] A.D. Sokal, Nucl. Phys. B (Proc. Suppl.) **20**, 55 (1991).
[8] R.H. Swendsen and J.S. Wang, Phys. Rev. Lett. **58**, 86

(1987).
[9] U. Wolff, Phys. Rev. Lett. **62**, 361 (1989).
[10] D. Kandel and E. Domany, Phys. Rev. B **43**, 8539 (1991).
[11] S. Caracciolo, R.G. Edwards, A. Pelissetto, and A.D. Sokal, Nucl. Phys. B (Proc. Suppl.) **20**, 72 (1991).
[12] R.J. Baxter, Ann. Phys. (N.Y.) **70**, 193 (1972).
[13] Actually the algorithm is for the q-state Potts model.
[14] H.G. Evertz, G. Lana, and M. Marcu, Phys. Rev. Lett. **70**, 875 (1993).
[15] If $L = K_\sigma$ or $L = K_\tau$ then $E_4 = E_3$ or $E_4 = E_2$. We do not discuss here the behavior of our algorithm in such a case.
[16] From now on we will absorb β into the coupling constants of the Hamiltonian. We will also omit the l dependence of all constants of the algorithm since we are considering,

for the time being, constant K_σ, K_τ, and L. Generalization to spatially varying couplings is trivial as long as $K_\sigma, K_\tau \geq L$ everywhere.

[17] A.D. Sokal, Nucl. Phys. B (Proc. Suppl.) **20**, 55 (1991).

[18] U. Wolff, Phys. Lett. B **228**, 379 (1989).

[19] A factor of $k_B T$ has been omitted, but is irrelevant since all measurements are at a single temperature.

[20] U. Wolff, Nucl. Phys. B **300**, 501 (1988).

[21] M.N. Barber, in *Phase Transitions and Critical Phenomena*, edited by C. Domb and J.L. Lebowitz (Academic, New York, 1983), Vol. 8.

[22] D.W. Heermann and A.N. Burkitt, Physica A **162**, 210 (1990).