Journal of Statistical Physics, Vol. 66, Nos. 5/6, 1992

Dynamical Analysis of Low-Temperature Monte Carlo Cluster Algorithms

Fabio Martinelli[1]

Received April 5, 1991

We present results on the Swendsen–Wang dynamics for the Ising ferromagnet in the low-temperature case without external field in the thermodynamic limit. We discuss in particular the rate of convergence to the equilibrium Gibbs state in finite and infinite volume, the absence of ergodicity in the infinite volume, and the long-time behavior of the probability distribution of the dynamics for various starting configurations. Our results are purely dynamical in nature in the sense that we never use the reversibility of the process with respect to the Gibbs state, and they apply to a stochastic particle system with *non-Gibbsian* invariant measure.

KEY WORDS: Monte Carlo; stochastic Ising model; dynamical phase transitions.

INTRODUCTION

In this paper we continue our analysis of the Swendsen–Wang dynamics for the ferromagnetic Ising model (see, e.g., refs. 1–3) in the low-temperature regime, which was begun in refs 4 and 5 in collaboration with E. Olivieri and E. Scoppola. The SW algorithm, a particular random clulster dynamics reversible with respect to the Gibbs state of the Ising model, is based on the Fortuin–Kasteleyn[6,7] representation of the Ising model, and it has the advantage, with respect to the usual single-spin-flip Glauber dynamics, of updating in a very efficient way the configurations on large scales. The algorithm works as follows: starting from a configuration σ, we construct a new configuration σ' in two steps:

(i) First we construct the "bond configuration" $\{\gamma(b)\}$, $b=(x, x')$, $|x-x'|=1$, as follows: a bond (x, x') is defined to be "vacant,"

[1] Dipartimento di Matematica, Universita' "La Sapienza," Rome, Italy.

0022-4715/92/0300-1245$06.50/0

$\gamma(b)=0$, if $\sigma(x)\neq\sigma(x')$; if $\sigma(x)=\sigma(x')$, then the bond $\{x, x'\}$ is defined to be "occupied," $\gamma(b)=1$, with probability $1-\exp(-\beta)$ and "vacant" with probability $\exp(-\beta)$, β being the inverse temperature.

(ii) Then, given $\{\gamma(b)\}$, we consider the connected sets of sites C, called "clusters," in the graph whose edges are the occupied bonds b. The second step consists in updating simultaneously all the spins in every cluster C. The updating is such that all the spins in C become either $+1$ or -1 with equal probability, independently for each cluster. Homogeneous boundary conditions (b.c.) may be taken into account by imposing that the clusters which are connected to the boundary cannot flip and must preserve the same value of the spin as the boundary (e.g., $+1$).

A more detailed construction of the SW algorithm is given in Section 1.

The above algorithm was introduced Swendsen and Wang[1] in order to reduce or even completely eliminate the critical slowing down that greatly hampered Monte Carlo simulations of critical phenomena in ferromagnetic systems of statistical mechanics such as Potts models. Their initial ideas were further developed and improved by a number of people (see, e.g., ref. 8 and references therein) and made available for models different from the original ones, such as plane rotators[9] or completely frustrated systems.[10]

This type of stochastic algorithm (known as stochastic cluster algorithms) proved to be very efficient from the numerical point of view (see, e.g., ref. 11), and, because of the greatly reduced computer time, allowed very detailed studies of the statistical properties of the "physical" clusters of the Ising model.[12]

From a theoretical point of view and in connection with numerical simulations, the central point of this subject is to study the critical behavior of the dynamics. Unfortunately, very little is known rigorously on this difficult problem, with the exception of a rigorous lower bound on the dynamical critical exponent z obtained by Li and Sokal.[13] However, if one is interested in a rigorous analysis of purely nonequilibrium phenomena, such as the way equilibrium is approchad, metastability, or large deviations from equilibrium, then the SW turns out to be a very interesting model of random dynamics for which it is possible to develop new ideas and techniques that can be applied also to different contexts. For example, ref. 4 successfully discussed the approach to equilibrium in a finite but arbitrarily large volume at low temperature and in the presence of a small positive external field, by means of a novel multiscale analysis in space-time borrowed from statistical mechanics of disordered systems.[14] When boundary conditions were opposite to the external field, the dynamics was shown to exihibit metastable behavior, i.e., starting from the "wrong" phase, equilibrium was reached through homogeneous nucleation of droplets larger than a critical one of the "right" phase. This transition was analyzed in detail in ref. 5, where the existence of the critical droplet was shown together with sharp bounds on the tunneling time. A similar study for the more conventional Metropolis algorithm was carried out at the same time by Jordao-Neves and Schonmann.[15]

We stress here that the presence of an external field in the results of ref. 4 was absolutely crucial, since it allowed the use of expansions similar to low-temperature expansions of equilibrium statistical mechanics in order to have a rough control of the size of the clusters evolving in time, and therefore of the speed with which information propagates through the system. A very interesting question is therefore what happens at zero external field and low temperature. A zero-temperature analysis shows immediately that in this case the SW dynamics is very different from a single spin dynamics. In fact, for a traditional single-spin algorithm like Metropolis with plus boundary conditions in a box of side L in 2 dimensions, a spin configuration starting from all minuses will become identically equal to plus only in a time of order L^2, by a kind of erosion mechanism starting from the boundary of the chosen box. On the contrary, in the SW dynamics, the same configuration will flip to all pluses in a time of order one. Actually, one easily proves (see the discussion after Theorem 2.1) that any initial configuration will reach equilibrium, i.e., all pluses, in a time of order $\log(L)$. This fact suggests that also the low-temperature behavior, e.g., equal site time correlations at equilibrium, should be different between the two dynamics.

To this purpose, we recall that there is a very convincing argument by Huse and Fisher[16] (see also Sokal and Thomas[17]) predicting a stretched exponential $[\exp(-\sqrt{t})]$ convergence to equilibrium in two dimensions for the Metropolis algorithm, essentially based on the observation that large clusters of the wrong phase survive for a very long time (proportional to their area) under the dynamics. For the SW dynamics, however, big clusters of the wrong phase, which are therefore not attached to the boundary, can be flipped in a single move even without external field. Thus we conjecture that the SW dynamics should approach equilibrium exponentially fast in time. Although we are not able to prove this here, we show that the convergence is faster than $\exp(-t^{\alpha})$ with $\alpha = \ln(2)/\ln(3)$.

A second very interesting question concerns the behavior of the dynamics in the infinite volume $\mathbf{Z}^d$. In this case the associated Ising model exihibits a phase transition, and a nontrivial problem is to study the limit (if it exists) as t tends to infinity of the probability distribution of the dynamics at time t. For the usual Glauber dynamics, such as Metropolis or heath bath algorithms, absence of ergodicity is proved using the attractivity of the dynamics and the reversibility with respect to the Gibbs measure. Attractivity is equivalent to saying that for any time t, if $f(\sigma)$ is an increasing function of the spin configuration σ, then the expected value of f over the configuration σ_t is an increasing function of the starting configuration. This fact, toghether with reversibility, is sufficient to prove, for example, that, starting from all pluses, the expected value of the spin at the origin will always be greater than or equal to the magnetization in the plus phase (μ_+), while starting from all minuses, the same average will always be smaller than or equal to the magnetization in the minus phase μ_-.

If the temperature is below the critical point, there is spontaneous magnetization and therefore the system is no longer ergodic. It is, however, well known that it is very difficult to prove this result by purely dynamical methods, i.e., without using the atractivity and reversibility of the dynamics, and, to our knowledge, no rigorous results are available in this direction. A notable exception is represented by the beautiful work of Toom on stochastic cellular automata.[18]

For the SW dynamics the situation is in some respect more complicate than for a Glauber dynamics, since attractivity no longer holds and a detailed analysis of the dynamics in unavoidable.

As it will become clearer in the sequel, in order to be able to give even a partial answer to the above questions, one is forced to have a good probabilistic control on the occurrence during the time evolution of long paths of vacant bonds. In ref. 4 this control was achieved through the external field h, since each vacant bond at integer time corresponds to a spin opposite to the field. In the absence of the magnetic field the situation is much more complicated, since now a given path of vacant bonds may resist for a long time t with probability $(1/2)^t$. Therefore one cannot hope to get, uniformly in the starting configuration, estimates on the probability to observe at time t a path of vacant bonds of length L starting from a given site x, which are exponentially small in L, as is the case for the Gibbs state at low temperature, unless t is much larger than L. On the other hand, it is a central point of our strategy to think of the dynamics on a given length scale L as being built up by many local dynamics on a smaller length scale $L' \ll L$ evolving in time more or less independently one from the other. This picture is of course valid only if "information" in the system has not been able to travel in the given time scale t a distance L'. This fact impliers that, even with a bounded velocity of "propagation of information," length scales should be larger than time scales.

In this paper we provide a first solution to the above problems, certainly not the optimal one, for the SW dynamics by means of a multi-scale analysis which avoids completely any kind of Peierls argument and in general any *a priori* knowledge about the equilibrium Gibbs measure. This last feature of our approach is in our opinion the most important one, since it allows us to treat other models of interacting particle systems which do not have a Gibbsian invariant measure. This is the case of the model introduced in ref. 19 in dimension $d \geqslant 2$ which will be discussed in Section 5. As pointed out by Lebowitz and Schonmann,[20] invariant measure of nonequilibrium statistical mechanics should generically be expected to be *non-Gibbsian.*

The paper is organized as follows: In Section 1 we define precisely the dynamics. In Section 2 we prove the basic estimates on the probability of having long paths of vacant bonds and we show the existence in the infinite volume of an infinite cluster starting from a homogeneous configuration (all spins $= +1$ or -1). In Section 3 we study the rate of convergence to equilibrium in a finite volume. In Section 4 we give a dynamical proof of the existence of aphase transition. In Section 5 we briefly discuss the application of the techniques to another model of an interacting stochastic particle system with a *non-Gibbsian* invariant measure.

1. CONSTRUCTION OF THE DYNAMICS AND NOTATION

We start by constructing the dynamics with $+$ boundary conditions in a finite subset of the d-dimensional cubic lattice $\mathbf{Z}^d$. We first introduce the notation.

(i) Λ will denote a generic finite subset of $\mathbf{Z}^d$. Given a pair of sites x and y in $\mathbf{Z}^d$, we set

$$\delta(x, y) = \sum_{i=1..d} |x_i - y_i|$$

$$d(x, y) = \max_{i=1..d} |x_i - y_i|$$

$$\operatorname{diam}(\Lambda) = \sup_{x, y \in \Lambda} \delta(x, y)$$

The distance between two sets A, B, denoted by $d(A, B)$, is given by $\min_{x \in A, y \in B} d(x, y)$.

(ii) The unordered pair b in $\mathbf{Z}^d$, $b = (x, y)$, with $\delta(x, y) = 1$ is called a bond. Λ^* is the set of all bonds (x, y) such that either x or y or both belong to Λ. The set of all bonds in $\mathbf{Z}^d$ will be denoted by $\mathbf{Z}^{d*}$.

(iii) $\sigma \in \{-1, 1\}^{\Lambda}$ denotes a generic configuration of plus or minus spins in Λ.

(iv) C_Λ is the family of all "geometric clusters" C in $\bar{\Lambda}=\{x; \exists b\in\Lambda^*; x\in b\}$. A geometric cluster C is a subset of $\mathbf{Z}^d$ which is connected in the following sense: $\forall x, y\in C$ there exists a chain of nearest neighbor sites in C connecting x to y:

$$x^1 \cdots x^n\colon \quad x_1=x, \quad x^n=y, \quad \delta(x^{i+1}, x^i)=1 \quad i=1,..., n-1$$

(v) Given a geometric cluster C, we define the "outermost boundary" of C as the set of sites x not in C such that there exists a nearest neighbor of x in C and there exists a chain of nearest neighbors sites $x_1, x_2,..., x_N$ in $\bar{\Lambda}\backslash C$, with $x_1=x$ and $x_N\in\partial\Lambda$, where $\partial\Lambda=\{x\in\bar{\Lambda}\backslash\Lambda\}$.

(vi) A collection $\gamma\equiv\{b_1, b_2,..., b_n\}$ of bonds in Λ^* is called "a path of bonds containing x" iff x is the endvertex of one of the bonds b_i and the distance between the endvertices of two consecutive bonds (as a set of two sites in $\mathbf{Z}^d$) is not greater than one. The length of the path γ, $|\gamma|$, is set equal to $\operatorname{diam}(V(\gamma))$, where $V(\gamma)$ is the set of endvertices of the bonds $\{b_1, b_2,..., b_n\}$.

Now, given Λ, let ν_b and $\xi(C)$ be numbers in $\{0, 1\}$ associated to each bond and to each geometric cluster $C\in C_\Lambda$, respectively. Given the numbers ν_b and $\xi(C)$, we construct out of a configuration σ a new configuration σ' as follows. From σ we first generate a configuration γ of occupied $[\gamma(b)=1]$ and vacant $[\gamma(b)=0]$ bonds, by setting

$$\gamma(b)=\left(\frac{1+\sigma_b}{2}\right)\nu_b$$

where $\sigma_b=\sigma(x)\,\sigma(y)$ if $b=(x, y)$ and $\sigma(x)= +1$ if $x\in\partial\Lambda$. The configuration γ can be identified as the subset of the occupied bonds in Λ^*. Sometimes in order to denote the configuration (and the corresponding subset of Λ^*) γ obtained starting from σ, we use the symbol γ_σ (γ_σ depends of course on the numbers ν_b). We will say that two n.n. sites (x, x') are connected in the bond configuration γ if $\gamma(x, x')=1$, i.e., the bond (x, x') is occupied in γ. The maximal connected components C (with respect to the configuration γ) are called "γ-clusters" or more simply clusters. They are of course in particular geometric clusters and may reduce to a single site.

For a geometric cluster C which is also a γ-cluster we often write $C\subset\gamma_\sigma$. Now for any $C\subset\rho_\sigma$ we set

$$\begin{aligned} &\sigma'(x)=1 \quad \forall x\in C \qquad &&\text{if either } \xi(C)=0 \text{ or } C\cap\partial\Lambda\neq\varnothing \\ &\sigma'(x)=-1 \quad \forall x\in C \qquad &&\text{if } \xi(C)=1 \text{ and } C\cap\partial\Lambda=\varnothing \end{aligned} \tag{1.1}$$

Let us now consider two sequences of numbers

$$\omega \equiv (\{v_b(t)\}_{t \in N, b \in \Lambda^*}; \{\xi(t, C)_{t \in N, C \in C_\Lambda})$$

that we think of as the realization of two mutually independent processes with values in $\{0, 1\}$ each of which is a collection of independent identically distributed random variables (i.i.d. rv) with distribution

$$v_b = 0 \qquad \text{with probability } \exp(-\beta)$$

$$v_b = 1 \qquad \text{with probability } 1 - \exp(-\beta)$$

and Bernoulli distribution with parameter 1/2 for the $\xi(s, C)$.

Given ω, we finally construct a random flow on $\{-1, 1\}^\Lambda$, $\{\phi_t^{\Lambda,\omega}(\cdot)\}_{t \in N}$ by applying at each time step t the rule (1.1) with numbers $v_b(t)$, $\xi(t, C)$. Sometimes, for notational convenience, we will write

$$\sigma_t^\omega(x) = \phi_t^{\Lambda,\omega}(\sigma)(x) \tag{1.2}$$

Sometimes for notational convenience we will say that some event E occurs at time $t + 1/2$ if E depends only on the random variables $\gamma(b)$ constructed from the configuration $\phi_t^{\Lambda,\omega}(\sigma)$ using the random variables $v_b(t)$.

Remark 1. (i) The boundary condition $+1$ at the boundary of Λ is taken into account by the condition that any cluster C touching $\partial\Lambda$ is set equal to $+1$. Other boundary conditions may be considered, e.g., periodic or open.

(ii) Notice that if $\Lambda' \subset \Lambda$, then one can compare the random flows $\phi_t^{\Lambda,\omega}$, $\phi_t^{\Lambda',\omega}$ as follows: given σ in Λ, one constructs $\hat{\sigma}$ in Λ' by the rule

$$\hat{\sigma}(x) = \sigma(x) \qquad \text{if} \quad x \in \Lambda'$$

$$\hat{\sigma}(x) = +1 \qquad \text{if} \quad x \in \partial\Lambda'$$

The evolutions $\phi_t^{\Lambda,\omega}(\sigma)$ and $\phi_t^{\Lambda',\omega}(\hat{\sigma})$ are constructed by means of the same random numbers $(v_b(t), \xi(t, C))$ if b and C are in Λ'. However a cluster C intersecting $\partial\Lambda'$ is set equal to $+1$ for the dynamics $\phi_t^{\Lambda',\omega}$ but may be -1 for the dynamics $\phi_t^{\Lambda,\omega}$. This observation will be exploited in a crucial way in Section 3.

It is easy to see that the above-defined dynamics satisfies the detailed balance condition for the Gibbs state of the Ising model on Λ, with $+$ boundary conditions on $\partial\Lambda$, at inverse temperature β. The proof of this statement can be found, for example, in refs. 1 and 4.

2. THE BASIC ESTIMATE

In this section we will establish a basic estimate on the probability of having a path of vacant bonds [i.e., with the associated random variables $\gamma(b)$ equal to zero] of length L containing a fixed point x at a given time t, with t greater than some time scale $t(L)$ related to L. Such an estimate will play a crucial role in establishing the results of the subsequent sections. For simplicity we will discuss only the two-dimensional case; the result, however, holds also in higher dimension. Let us first fix some notations. For any integer k we define:

(i) $L_k = 4^{k^2}$.

(ii) $t_k = 3^k$.

(iii) $\Lambda_L(x) = \{y \in \mathbf{Z}^2;\ d(x, y) \leqslant L\}$, $\Lambda_L = \Lambda_L(0)$, $\Lambda_k = \Lambda_{L_k}$.

(iv) Given Λ_L with $L > L_k$, we denote with the name $(k, +)$-dynamics in Λ_L the algorithm described in the previous section with the following extra condition:

$$\xi(s, C) = 0 \qquad \text{if} \quad \operatorname{diam}(C) > L_k$$

(v) $\Omega_{L,x,t,k,\sigma}$ will denote the event that there exists at time $t + 1/2$ a path of vacant bonds in $\mathbf{Z}^{2*}$ of length $n \geqslant L_k$ containing x for the dynamics in Λ_L with + b.c. starting from σ. $\Omega^k_{L,x,t,k,\sigma}$ will denote the same event, but computed for the $(k, +)$-dynamics in Λ_L starting from σ.

(vi) $P_k = \sup_{L \geqslant L_k; x \in \Lambda_L; t \geqslant t_k; \sigma \in \{-1,1\}^{\Lambda_L}} \max(P(\Omega_{L,x,t,\sigma}),\ P(\Omega^k_{L,x,t,k,\sigma}))$.

For convenience and whenever this will not lead to confusion, we will denote with $P(L, x, t, k, \sigma)$ either $P(\Omega^k_{L,x,t,k,\sigma}))$ or $P(\Omega_{L,x,tg,k,\sigma}))$ without specifying the dynamics for which it is evaluated. Then we will prove the following result:

Theorem 2.1. There exists $\beta_0 > 0$, $c > 0$, $k_0 > 0$, and $a > 0$ such that for any $\beta \geqslant \beta_0$ there exists a positive constant $m(\beta)$ with $m(\beta) \geqslant c$ such that

$$P_k \leqslant \frac{1}{L_k^{2a}} \exp[-m(\beta) 2^k] \qquad \forall k > k_0$$

Before proving the theorem, it is important to understand the case of zero temperature, $\beta = \infty$. In this case no bond is made vacant during the dynamics and the only possibility to observe a path of vacant bonds at time $t + 1/2$ is that the same path was already present at any previous time including $t = 0$. The probability of this last event is bounded from above by $(1/2)^t$; however, if the path under consideration separates exactly two different clusters at time $t=0$, then the above bound becomes exact. This discussion suggests that any bound on P_k will be at most exponential in the time scale t_k with rate constant $m(\beta)$ at most equal to $\ln(2)$ and in particular that to obtain a rigorous bound on P_k by means of some kind of Peierls argument should be a very difficult task, since the number of paths grows exponentially fast on the length scale $L_k \gg t_k$.

One should at this point be puzzled by our choice of the length and time scales ($L_k \gg t_k$), since the above arguments seem to indicate that time scales of the same order as the length scales should be more appropriate. As will appear clear in the course of the proof, it is a central point of our strategy that if the dynamics starts from a configuration with paths no longer than L_k and at a later time t_k a path longer than L_{k+1} is present, then there are at least two pieces of it, each of length greater than L_k, that have been created independently one from the other. A proof of this fact requires, however, that time scales are much smaller than length scales (more precisely: $t_k L_k \ll L_{k+1}$).

The actual result, although it is sufficient for our purposes, is unfortunately much weaker than the naive guess made on the basis of the above discussion, since it is only an exponential of $t_k^{(\ln 2/\ln 3)}$. A substantial improvement seems to require new ideas.

Proof. For β large enough we will show that the quantity P_{k+1} "on scale $k+1$" can be estimated in terms of the same quantity "on scale k" P_k by

$$P_{k+1} \leqslant L_k^a P_k^2 \tag{2.1}$$

for a suitable positive constant a independent of k and β. If (2.1) holds and if $f_k = L_k^{2a} P_k$, then, by explicit computation,

$$f_{k+1} \leqslant f_k^2 \tag{2.2}$$

provided $k \geqslant 5$. Thus

$$f_k \leqslant (f_{k_0})^{2^{k-k_0}} \tag{2.3}$$

for any $k_0 \geqslant 5$, i.e., using the definition of f_k,

$$P_k \leqslant \frac{1}{L_k^{2a}} (L_{k_0}^{2a} P_{k_0})^{2^{k-k_0}} \tag{2.4}$$

Therefore the theorem follows with $m(\beta) = -2^{-k_0} \log(L_{k_0}^{2a} P_{k_0})$ provided that

$$\lim_{\beta \to \infty} (L_{k_0}^{2a} P_{k_0}) < 1 \tag{2.5}$$

for some $k_0 \geqslant 5$.

To prove (2.5), we first observe that using the Markov property of the dynamics, we have

$$P_k = \sup_{L \geqslant L_k; x \in \Lambda_L; \sigma \in \{-1,1\}^{\Lambda_j}} P(L, x, t_k, \sigma) \tag{2.6}$$

Moreover, for fixed L and fixed $\sigma \in \{-1, 1\}^{\Lambda_L}$ and x in Λ_L we have that

$$P(L, x, t_k, \sigma) \leqslant L_{k_0}^2 t_{k_0} e^{(-\beta)} + L_{K_0}^2 (\tfrac{1}{2})^{t_{k_0}} \tag{2.7}$$

The rhs of (2.7) is in fact a rough upper bound of the probability that at least one of the $L_{k_0}^2$ bonds b at distance from x smaller then or equal to L_{k_0} either was made vacant nat some time $s + 1/2$ smaller than t_{k_0} [i.e., $v(b, s+1/2) \leqslant e^{-\beta}$] or that $\sigma_s(b) = -1$, $\forall s \leqslant t_{k_0}$. It is easy to check that indeed the rhs of (2.7) is smaller than 1 if k_0 is taken large enough and β is large enough depending on k_0. We are therefore left with the proof of the basic recursion inequality (2.1). Using (2.6), it is sufficient to estimate $P(\Omega_{L,x,t_{k+1},k+1,\sigma})$, computed either for the dynamics with + b.c., of for the $(k+1, +)$-dynamics in Λ_L, with the rhs of (2.1) uniformly in $L > L_{k+1}$, $x \in \Lambda_L$, and in the initial configuration σ. Thus, let us fix $L \geqslant L_{k+1}$, $x \in \Lambda_L$, $\sigma \in \{-1, 1\}^{\Lambda_L}$, and one of the two dynamics in Λ_L and let us introduce the auxiliary event Ω_1:

$$\begin{aligned}\Omega_1 = \{&\text{there exists } y \text{ in } \Lambda_L \text{ with } d(y, x) \leqslant L_{k+1} \text{ and } s \text{ in } [t_k, 2t_k]\\ &\text{such that at time } s + \tfrac{1}{2} \text{ there exists}\\ &\text{a path of vacant bonds in } \mathbf{Z}^{2*} \text{ of length } n \geqslant L_k \text{ containing } y\}\end{aligned}$$

Remark 1. If Ω_1^c holds, then it holds also at integer times $s \in [t_k, 2t_k]$. This follows by the simple observation that the number of vacant bonds at time s is always smaller than or equal to the same number at time $s + 1/2$.

Then we write

$$P(\Omega_{L,x,t_{k+1},k+1,\sigma}) = P(\Omega_{L,x,t_{k+1},k+1,\sigma} \cap \Omega_1) + P(\Omega_{L,x,t_{k+1},k+1,\sigma} \cap \Omega_1^c) \tag{2.8}$$

where Ω_1^c denotes the complement of Ω_1. Using again the Markov property and the definition of P_k, we estimate the first term in the rhs of (2.8) uniformly in x, L, σ by

$$t_k L_{k+1}^2 P_k^2 \tag{2.9}$$

In fact, if in the event $\Omega_{L,x,t_{k+1},k+1,\sigma} \cap \Omega_1$ we also fix the site y and the time $s \in [t_k, 2t_k]$ entering in the definition of Ω_1, then we are examining a subset of the event that there are two paths of vacant dual bonds each one of length greater than L_k, one at time $s+1/2$ containing y and the other at time $3t_k+1/2$ containing x. Since both s and $t_{k+1}-s$ are greater than t_k and $L \geqslant L_{k+1} > L_k$, we can use P_k and the Markov property to obtain the estimate P_k^2 for the probability of this last event. The factor $t_k L_{k+1}^2$ takes into account the number of choices of y and s.

We now turn to the estimate of the second term in the rhs of (2.8). For simplicity we first describe the estimate for the case $x=0$. After that we will explain the (trivial) modifications which are necessary in order to consider the general case.

Definition 1. Given integers L and $l \leqslant L$, we set

$$\Lambda_{L-l} = \{x \in \Lambda_L;\, \mathrm{dist}(x, \partial\Lambda_L) \geqslant l\}$$

Definition 2. (a) Let

$$\begin{aligned}\tau_1 = \min\{s > t_k;\, \exists y \in \Lambda_{L_{k+1}-(s-t_k)l_k}\ &\text{such that at time } s+\tfrac{1}{2} \text{ there exists a path of vacant}\\ &\text{bonds in } \mathbf{Z}^{2^*} \text{ of length } n \geqslant L_k \text{ containing } y \text{ where } l_k = 4t_kL_k\}\end{aligned}$$

(b) Let y_1 be the leftmost and uppermost of the site y's appearing in the definition of the random time τ_1. Then we set

$$\begin{aligned}\tau_2 = \min\{s \geqslant \tau_1;\, \exists y \in \Lambda_{L_{k+1}-(s-t_k)l_k} &\text{ with } d(y, y_1) \geqslant (s+2-\tau_1)l_k\\ &\text{such that at time } s+\tfrac{1}{2}\\ &\text{there exists a path of vacant bonds in } \mathbf{Z}^{2^*} \text{ of length } n \geqslant L_k \text{ containing } y\}\end{aligned}$$

For convenience the leftmost and uppermost of the sites y's appearing in the definition of τ_2 will be denoted by y_2.

Remark 2. (i) By direct computation we have that $t_k l_k = L_{k+1}(9/16)^k$ and therefore $t_k l_k \ll L_{k+1}$ for k large.

(ii) If the event $\Omega_{L,x=0,t_{k+1},k+1,\sigma} \cap \Omega_1^c$ holds, then $\tau_1 > 2t_k$ and, more important, $\tau_2 \leqslant 3t_k$. In fact, necessarily $\tau_1 \leqslant 3t_k$, since otherwise for any time s between $2t_k$ and $3t_k+1/2$ the path of vacant bonds containing $x=0$ would have length smaller than L_k, which is impossible because of $\Omega_{L,x=0,t_{k+1},k+1,\sigma}$. In fact, if the site y_1 defined above is such that $d(y_1, 0) \geqslant l_k t_k$, then again $\tau_2 > 3t_k$ would imply that the path of vacant bonds starting from $x=0$ at time $3t_k+1/2$ has length smaller than L_k. The same occurs if $d(y_1, 0) \leqslant t_k l_k$ and $\tau_2 > 3t_k$, since $L_{k+1} \gg t_k l_k$.

We will therefore estimate $P(\Omega_{L,x=0,t_{k+1},k+1,\sigma} \cap \Omega_1^c)$ by

$$P(\{2t_k \leqslant \tau_1 \leqslant \tau_2 \leqslant 3t_k\} \cap \Omega_1^c)$$

$$\leqslant \sum_{s_1=2t_k}^{3t_k} \sum_{s_2=s_1}^{3t_k} \sum_{x_1 \in A_{k+1}} \sum_{x_2 \in A_{k+1}} P(\{\tau_1 = s_1;\, y_1 = x_1;\, \tau_2 = s_2;\, y_2 = x_2\} \cap \Omega_1^c) \tag{2.10}$$

From here the discussions are restricted to fixed τ_1, τ_2, y_1, y_2 without notice. The main idea to estimate the rhs of (2.10) is to show that, due to the condition expressed by Ω_1^c and by the definition of τ_1 and of τ_2, the paths γ_1 and γ_2 of vacant dual bonds with length greater than L_k starting at times s_1 and s_2 from x_1 and x_2, respectively, have been created by the random dynamics starting from the configuration σ_{t_k} independently one from the other within the time intervals $[t_k, s_1]$ and $[t_k, s_2]$. If this is the case, then each term in the sum (2.10) can be estimated by P_k^2 and the theorem follows. In order to carry out this program, we first prove two simple geometric results on the structure of the configuration σ_t for $t \leqslant \tau_2$.

Let Λ_{k+1}^s denote $\Lambda L_{k+1} - (s - t_k) l_k$, and let Λ_i be the square $\Lambda_{l_k}(y_i)$, $i = 1, 2$. We also denote by $C_s(x)$ the cluster containing the site x at time $s + 1/2$.

Lemma 2.1. For any s such that $t_k \leqslant s < \tau_1$ there exists a cluster denoted $C_s^\infty(1)$ with the property that its outermost boundary does not intersect Λ_{k+1}^s and such that for any $x \in \Lambda_{k+1}^s$ only one of the following two possibilities holds:

(1) $\operatorname{diam}(C_s(x)) < L_k$.

(2) $C_s(x)$ coincides with $C_s^\infty(1)$.

Lemma 2.2. For any $t_k \leqslant s < \tau_2$ there exists a cluster denoted $C_s^\infty(2)$ with the property that its outermost boundary does not intersect Λ_2 and such that for any $x \in \Lambda_2$ only one of the following two possibilities holds:

(1) $\operatorname{diam}(C_s(x)) < L_k$.

(2) $C_s(x)$ coincides with $C_s^\infty(2)$.

Proof of Lemma 2.1. Let us first show that for any $s < \tau_1$ there must exist in Λ_{k+1}^s a site x such that $\operatorname{diam}(C_s(x)) > L_k$. In fact, if for some $s \leqslant \tau_1 - 1$ and any $x \in \Lambda_{k+1}^s$, $\operatorname{diam}(C_s(x)) < L_k$, then necessarily there exists a site x_0 in Λ_{k+1}^s such that at time s there exists a path of vacant bonds in $\mathbf{Z}^{2*}$ of length $n \geqslant L_k$ containing x_0. This fact is of course in contradiction with the definition of τ_1.

Let us now fix $x \in \Lambda^s_{k+1}$ and let us assume that $\mathrm{diam}(C_s(x)) > L_k$. Since $s < \tau_1$, the outermost boundary of $C_s(x)$ cannot intersect the box Λ^s_{k+1}. We then take as $C^\infty_s(1)$ the cluster $C_s(x)$. It remains to show that if $y \in \Lambda^s_{k+1}$ is any other site such that $\mathrm{diam}(C_s(y)) > L_k$, then $C_s(y) = C^\infty_s(1)$. This is quite clear, since the outermost boundary of $C_s(y)$ cannot intersect Λ^s_{k+1} and therefore it must coincide with the outermost boundary of $C_s(x) = C^\infty_s(1)$.

Proof of Lemma 2.2. Let us fix $s < \tau_2$. Then any site x inside the box Λ_2 is at distance from y_1 greater than or equal to $(\tau_2 - \tau_1 + 2) l_k - l_k \geqslant (s - \tau_1 + 2) l_k$. Therefore, by the definition of τ_2, the longest path of vacant bonds at time $s + 1/2$ containing x and intersecting the box Λ_2 has length smaller than L_k. Thus, the same proof of Lemma 2.1 applies with the box Λ^s_{k+1} replaced by the box Λ_2.

Remark 3. (i) It is clear that the cluster $C^\infty_s(i)$, $i = 1, 2$, may coincide with the cluster of the boundary of the chosen box Λ_L.

(ii) If the dynamic s under consideration is the $(k+1, +)$-dynamics in Λ_L and if the diameter of $C^\infty_s \geqslant L_{k+1}$, then necessarily the sign of C^∞_s will be plus.

(iii) By construction, the outermost boundary of $C^\infty_s(1)$ cannot intersect the boxes Λ_i, $i = 1, 2$, for $s < \tau_1$.

The next step consists in establishing a coupling between the dynamics σ_s inside the boxes Λ_i and the $(k, +)$-dynamics in Λ_i, for $s \in [t_k, \tau_i]$, $i = 1, 2$. For this purpose, given a site x inside $\Lambda_{L_{k+1} - l_k}$, given $s_0 \in [t_k, 3t_k]$, and given a realization ω of the basic random variables $\{\nu(s, b)\}_{s \in [t_k, 3t_k]}$, $\{\xi(s, C)\}_{s \in [t_k, 3t_k]}$, we define in general $\eta^{(x, s_0)}_s$ as the evolution of the configuration σ_{t_k} inside the square $\Lambda_{l_k}(x)$ with the following new set of basic random variables:

(1) $\xi(s, C)' = \{\xi(s, C) + \xi(s, C^\infty_s)\} \bmod 2$ if $s < s_0$, if $\mathrm{diam}(C) \leqslant L_k$ and if C is strictly contained inside $\Lambda_{l_k}(x)$. Here C^∞_s, if it exists, is the unique cluster C for the true dynamics σ_s evolving with the given realization ω starting from σ_{t_k}, with the property that is outermost boundary does not intersect $\Lambda_{l_k}(x)$ and such that for any $y \in \Lambda_{l_k}(x)$ only one of the following two possibilities holds:

(i) $\mathrm{diam}(C_s(y)) < L_k$.

(ii) $C_s(y)$ coincides with C.

If C^∞_s does not exist, then $\xi(s, C^\infty_s)$ is set equal to one.

(2) $\xi(s, C)' = \xi(s, C)$ if $s \geqslant s_0$, if $\mathrm{diam}(C) \leqslant L_k$ and C is strictly contained inside $\Lambda_{l_k}(x)$.

(3) $\xi(s, C)' = 0$ otherwise.

(4) $v(s, b)' = v(s, b)$.

Remark 4. It is very important to realize that, given x and s_0 as above, the distribution of the new random variables $\{\xi(s, C)'\}$ is again Bernoulli with parameter 1/2 for those clusters C such that the diameter of the cluster C is smaller than L_k and $C \subset \Lambda_{l_k}(x)$, so that the probability distribution of the new evolution $\eta_s^{(x, s_0)}$ coincides with the distribution of the $(k, +)$-dynamics inside $\Lambda_{l_k}(x)$ with starting point σ_{t_k}. This fact depend in a crucial way on the fact that the probability distribution of the random variables $\xi(s, C)$ is symmetric $[P(\xi(s, C) = +1) = P(\xi(s, C) = 0) = 1/2]$.

Let now assume that the realization ω was such that $\tau_1 = s_1$, $\tau_2 = s_2$, $y_1 = x_1$, and $y_2 = x_2$, and let us compare the true evolution σ_s with the evolution $\eta_s^{(x_i, s_i)}$ as defined by the above rules inside the squares Λ_i. We have the following important result:

Lemma 2.3. Let $s \leqslant s_i$ and let $b = (x, y)$ be a bond in Λ_i such that $d(b, \partial\Lambda_i) \geqslant 2sL_k$. Then

$$\eta_s^{(x_i, s_i)}(b) = \sigma_s(b)$$

Proof. The proof is by induction and it is the same for $i = 1$ or $i = 2$. Let us assume that the result of the lemma is true up to time s with $s + 1 \leqslant s_i$, and let us show that it holds also at time $s + 1$. From the inductive hypothesis, at time $s + 1/2$ the bond variables inside the box $\Lambda_{l_k - 2sL_k}(x_i)$ for η and σ are equal. Therefore, if $b = (x, y)$ is as in the hypothesis and the clusters $C_s(x)$ and $C_s(y)$ computed for σ_s have diameter less than L_k, then necessarily they must coincide with the clusters of x and y computed for $\eta_s^{(x_i, s_i)}$: in this case $\eta_{s+1}^{(x_i, s_i)}(b) = \sigma_{s+1}(b)$ by construction.

If $C_s(x)$ and $C_s(y)$ computed for σ_s both have diameter greater than L_k, then, by Lemmas 1.2 and 2.2, they must coincide with C_s^∞; on the other hand, again by the inductive hypothesis, the clusters of x and y computed for η_s also must have diameter greater than L_k. Therefore, in this case, by construction, $\eta_{s+1}^{(x_i, s_i)}(b) = \sigma_{s+1}(b) = 1$.

The last case, when only one between $C_s(x)$ and $C_s(y)$ has diameter less than L_k, follows by a similar reasoning.

Remark 5. It follows immediately from the above lemma and from the fact that $2s_i L_k \ll l_k/2$ that at time $s_i + 1/2$ for the new dynamics $\eta_s^{(x_i, s_i)}$ there exists a path of vacant bonds of length at least L_k containing x_i. For notational convenience we will denote this last event by $\Omega^n_{l_k, x_i, s_i, \sigma_{t_k}}$.

Using the above remark, we can finally estimate the generic term in the sum in the rhs of (2.10) by

$$P(\{\tau_1=s_i;\ y_1=x_1;\ \tau_2=s_2;\ y_2=x_2\}\cap\Omega_1^c)$$
$$\leqslant P(\Omega^{\eta}_{l_k,x_1,s_1,\sigma_{t_k}}\cap\Omega^{\eta}_{l_k,x_2,s_2,\sigma_{t_k}}) \tag{2.11}$$

Since $d(x_1, x_2)>2l_k$, the two dynamics $\eta_s^{(x_i,s_i)}$, $i=1,2$, are clearly independent and therefore, using Rermark 4, the r.h.s. of (2.11) can be estimated by P_k^2, which gives the estimate

$$t_k^2L_{k+1}^4P_k^2 \tag{2.12}$$

for the rhs of (2.10).

If we combine now (2.9) with (2.12), we get that

$$P(\Omega_{L,x=0,t_{k+1},k+1,\sigma})\leqslant(t_kL_{k+1}^2+t_k^2L_{k+1}^4)P_k^2 \tag{2.13}$$

As we have already anticipated, the same estimate can be obtained also for $x\neq 0$. In this last case all the steps that led to the estimate (2.12) are unchanged except that now the square Λ_{k+1} has to be replaced by $\Lambda_{k+1}(x)\cap\Lambda_L$ and the same for Λ_{k+1}^s and Λ_i.

Thus, the basic recursion inequality (2.1) is proved with, e.g., $a=18$ and the theorem follows.

Before closing this section, we would like to comment about our particular choice of the boundary conditions $(+)$ for the two dynamics involved in Theorem 2.1. Our choice was not at all essential for the result to hold and other b.c., such as open or periodic ones, can be accommodated as well. In these cases the proof is unchanged provided that one defines the quantity $P(L, x, k, \sigma)$ appearing in the definition of P_k as the largest of (1) the probability that there exists at time $t+1/2$ a path of vacant bonds in $\mathbf{Z}^{2*}$ of length $n\geqslant L_k$ containing x for the dynamics in Λ_L with the chosen b.c. starting from σ and (2) the same quantity computed for the $(k, +)$-dynamics in Λ_L starting from σ.

Rather interesting for later applications are the $(p, 1-p)$-b.c. defined as follows: all the clusters which touch the boundary of Λ_L are part of a unique cluster, called the boundary cluster, which is set equal to $+1$ with probability p and to -1 with probability $1-p$. In this case the random variable $\xi(s, C_s^\infty)$ may have Bernoulli distribution with parameter p, as well as Bernoulli distribution with parameter 1/2, depending on whether the cluster C_s^∞ touches the boundary of Λ_L or not. In both cases the random variables $\xi\omega'(s, C)$ will be distributed according to the Bernoulli distribution of parameter 1/2 and the proof will remain unchanged.

There are, however, limitations that come from those b.c. that naturally produce already in the Gibbs state long paths of vacant bonds, such as the $(+, -)$ b.c. in two dimensions (i.e., + b.c. in the upper half-plane and − b.c. in the lower half-plane of $\mathbf{Z}^2$). The proof in this case breaks down and the reason is that the analog of estimate (2.7) no longer holds. In fact, a given bond b may remain vacant for a long time with large probability if its endpoints belong to two different boundary clusters pinned to opposite sign by the b.c. It seems, however, that if one modifies the definition of $P(L, x, t, k, \sigma)$ in such a way that one considers only those paths that do not intersect the long path which joins the opposite side of Λ_L separating the plus phase from the minus phase, then the technique illustrated above can be applied again.

Also, the case of an external magnetic field parallel to the b.c. could be accommodated as well, but it was not included in our discussion, since this case was already successfully discussed in ref. 4.

From the above proof and particularly from the discussion made right after the theorem it appears that the main reason for connecting length scales with time scales comes from taking in the definition of the quantity P_k the supremum over all possible initial configurations, since for starting configurations with many paths of vacant dual bonds (e.g., a chessboard configuration) it takes a long time to become more regular and to look like a typical configuration of the equilibrium Gibbs state. This suggests that if we start already with a "regular configuration," then the probability of having a long path of vacant dual bonds at time t should decay fast enough in the length of the path already for short times. There is, however, a problem to solve since in the course of the proof of the theorem and particularly in the estimate of the first term in the rhs of (2.8), we made use of the Markov property due to the fact that in the definition of P_k we took the supremum over σ. Thus the following result becomes rather interesting. Let P_k^+ be defined as

$$P_k^+ = \sup_t \sup_{L > L_k} P(L, x, t, k, +) \tag{2.14}$$

where + denotes the configuration identically equal to plus one. Then we have the following result.

Theorem 2.2. The exists $\beta_0 > 0$, $c > 0$, and $a > 0$ such that for any $\beta \geqslant \beta_0$ there exists a positive constant $m_0(\beta)$ with $m_0(\beta) \geqslant c$ such that

$$P_k^+ \leqslant \frac{1}{L_k^{2a}} \exp[-m_0(\beta) 2^k], \qquad \forall k$$

Proof. Let k_0 and c be given by Theorem 2.1 and let $m_0(\beta)$ be given by $m_0(\beta) = \frac{1}{2} m(\beta) \wedge 2c$, where the constant $m(\beta)$ is as in Theorem 2.1. The above estimate on P_k^+ is clearly true as $\beta \to \infty$ for $k < k_0$. For $k \geqslant k_0$, in analogy with (2.1), we will prove that

$$P_{k+1}^+ \leqslant L_k^{2a}(P_k^+)^2 + \exp[-m(\beta)2^{k+1}] \tag{2.15}$$

where the constants m and a are as in Theorem 2.1. As in the proof of Theorem 2.1, it is easy to see that the result follows from (2.15) with, e.g., if for some fixed k_1, e.g., $k_1 = k_0$, we have

$$P_{k_1}^+ \leqslant \frac{1}{L_{k_1}^{2a}} \exp[-m_0(\beta)2^{k_1}]$$

This inequality is certainly true as $\beta \to \infty$, since we start with a configuration with no vacant bonds. Thus we will concentrate on the proof of the above modified recursion inequality.

Let us consider $P(L, x, t, k+1, +)$ with $L > L_{k+1}$ fixed and $x = 0$ for simplicity. If $t > t_{k+1}$, then by Theorem 2.1,

$$P(L, x, t, k+1, +) \leqslant \exp[-m(\beta)2^{(k+1)}] \tag{2.16}$$

If instead $t \leqslant t_{k+1}$, then, following the proof of Theorem 2.1, we define the random rimes τ_1 and τ_2 as follows.

Definition 3. (a) $\tau_1 = \min\{s \geqslant 0;\ \exists y \in \Lambda_{L_{k+1} - sl_k}$ such that at time $s + 1/2$ there exists a path of vacant bonds in $\mathbf{Z}^{2*}$ of length $n > L_k$ containing $y\}$, where $l_k = 4t_k L_k$.

(b) Let y_1 be the leftmost and uppermost of the sites y's appearing in the definition of the random time τ_1. Then we set

$$\begin{aligned}\tau_2 = \min\{s \geqslant \tau_1;\ &\exists y \in \Lambda L_{k+1} - sl_k \text{ with } \delta(y, y_1) \geqslant (s + 2 - \tau_1)l_k\\ &\text{such that at time } s + \tfrac{1}{2} \text{ there exists a path of}\\ &\text{vacant bonds in } \mathbf{Z}^{2*} \text{ of length } n \geqslant L_k \text{ containing } y\}\end{aligned}$$

For convenience the leftmost and uppermost of the sites y's appearing in the definition of τ_2 will be denoted by y_2.

It is easy to see, following the same arguments explained in Remark 2, that necessarily

$$\begin{aligned}&P(L, x=0, t, k, +)\\ &\quad \leqslant P(\tau_1 \leqslant \tau_2 \leqslant t)\\ &\quad \leqslant \sum_{s_1 \leqslant t} \sum_{s_2 = s_1}^{t} \sum_{x_1 \in \Lambda_{k+1}} \sum_{x_2 \in \Lambda_{k+1}} P(\{\tau_1 = s_1;\ y_1 = x_1;\ \tau_2 = s_2;\ y_2 = x_2\})\end{aligned} \tag{2.17}$$

In order to estimate a generic term in the sums in (2.17), we proceed exactly as we did for the estimate of the rhs of (2.10) and obtain the same result as in (2.11), (2.12):

$$P(\{\tau_1=s_1;\, y_1=x_1;\, \tau_2=s_2;\, y_2=x_2\})\leqslant (P_k^+)^2 \tag{2.18}$$

that is, the bound

$$t_{k+1}^2 L_{k+1}^4 P_k^2 \tag{2.19}$$

for the rhs of (2.17). If we now combine (2.19) together with (2.16), we get (2.15) and the theorem.

Remark 6. Exactly as for Theorem 2.1, the result of Theorem 2.2 applies also for the SW dynamics with different boundary conditions, e.g., open periodic or $(p, 1-p)$ b.c., which do not induce by themselves a long path of vacant bonds in the corresponding Gibbs state.

We conclude with a rather standard application of the above result to thne problem of the existence of an infinite cluster containing the origin for the SW dynamics on the full lattice $\mathbf{Z}^d$. Of course, on the whole lattice $\mathbf{Z}^d$ we need to give a prescription for the updating of cluster of infinite size. We decide to choose the $(p, 1-p)$ rule, which corresponds to setting an infinite cluster equal to $+1$ with probability p and to -1 with probability $1-p$. Thus, the $(1, 0)$ rule should correspond to the $+$ b.c. in the finite volume. It is easy to check that the result of Theorem 2.2, being uniform in the volume, applies also to the infinite-volume case with the $(p, 1-p)$ rule. Then we have the following result.

Corollary 2.1. There exists $\beta_0<\infty$ such that for any $\beta>\beta_0$ there exists $c(\beta)\in(0, 1)$ with $\lim_{\beta\to\infty} c(\beta)=1$ such that for any value of $p\in[0, 1]$, if $\eta(x)=1$, $\forall x\in\mathbf{Z}^d$, then: (a) We have

$$P_\eta(\operatorname{diam}(C_t^\infty(0))=+\infty)\geqslant c(\beta) \qquad \forall t$$

where $C_t^\infty(0)$ denotes the cluster containing the origin at time $t+1/2$ for the dynamics starting from η in $\mathbf{Z}^d$ with the $(p, 1-p)$ rule.

(b) For any time t there exists a unique infinite cluster with $P=1$.

Proof. (a) For any $\bar{k}$ we estimate $P_\eta(\operatorname{diam}(C_t^\infty(0))<+\infty)$ by

$$\begin{aligned}&P_\eta(\exists b\in\Lambda_{\bar{k}};\, b \text{ is vacant at time } t+\tfrac{1}{2})\\ &\quad+\sum_{k>\bar{k}} P_\eta\,(\text{there exists a path of vacant bonds of length } L>L_{k-1}\\ &\quad\text{containing } x, \text{ for some } x \text{ in } \Lambda_k)\end{aligned} \tag{2.20}$$

The first term in (2.20) is bounded by

$$(2L_{\bar{k}})^d[(\tfrac{1}{2})^\beta + (e^{-\beta})\beta]$$

If $t > \beta$ and by $(2L_{\bar{k}})^d \beta e^{-\beta}$ if $t \leqslant \beta$, while the second term is small for large $\bar{k}$, by Theorem 2.2. Thus (a) follows.

(b) If at time t there exist more than one infinite cluster, then there exists a path of vacant bonds of infinite length. That is excluded by Theorem 2.2. The existence with $P = 1$ of one infinite cluster is assured by (a) and by the ergodicity of the probability distribution at time t.

3. RATE OF CONVERGENCE TO EQUILIBRIUM IN A FINITE VOLUME

In this section we discuss the implication of our basic result, Theorem 2.1, for one of the most important problems of Monte Carlo algorithms, namely the rate of convergence of the probability distribution of the random dynamics at time t as $t \to \infty$ to the equilibrium measure given by the Gibbs state in a finite but large volume Λ. Thus, let us fix a box of side $2L$ centered at the origin Λ_L; let $f: \{-1, 1\}^{\Lambda_L} \to \mathbf{R}$ be an arbitrary observable; and let $\mu^+_{\Lambda_L}(f)$ and $E_\sigma f(\sigma_t)$ denote the expected value of f with respect to the Gibbs state in Λ_L with + b.c. and to the SW dynamics with + b.c. at inverse temperature β, respectively. Then we will prove the following result:

Theorem 3.1. There exists $\beta_0 < \infty$ and $c > 0$ such that for any $\beta > \beta_0$ there exists a positive constant $m(\beta)$ with $m(\beta) > c$ such that for any $t > t(L) = \exp[\ln(3)(\ln(L)/\ln(4))^{1/2}]$

$$\sup_\sigma |\mu^+_{\Lambda_L}(f) - E_\sigma f(\sigma_t)| \leqslant 2\, |f|_\infty \exp[-m(\beta)\, t^\alpha]$$

where $\alpha = \ln(2)/\ln(3)$ and $|f|_\infty = \sup_\sigma |f(\sigma)|$.

Before proving the theorem, it is important to understand the reason for the restriction on the time: $t > t(L)$. In a finite volume at zero temperature the equilibrium measure is totally concentrated on the plus configuration; moreover, any initial configuration after a time of the order $\ln(L)$ also becomes identically equal to plus one. More precisely,

$$\sup_\sigma P^{(\beta=\infty)}(\sigma_t(x) \neq 1 \text{ for some } x) \leqslant (2L)^d(\tfrac{1}{2})^t$$

It is also easy to check that for certain anomalous initial configurations (e.g., a chessboard configuration) the above upper bound becomes almost exact. Thus, equilibrium is only reached after a time of the order of the logarithm of the volume. At very low but positive temperature the result of Theorem 2.1 proves that any configuration after a suitable time becomes almost one in the sense that with large probability a given site x connected to the boundary of the chosen box; however, due to our choice of the time and length scales, the upper bound on the time necessary for this to happen is not logarithmic in the volume, but only of the form expressed in the theorem.

Proof of Theorem 3.1. The proof follows very closely the pattern of the proof of the same result given in ref. 4 for the SW dynamics with an external positive field. Let $p(L, t)$ be defined by

$$p(L, t) \equiv \sup_{\sigma, \eta} P(\eta_t \neq \sigma_t) \tag{3.1}$$

where η_t and σ_t are coupled together in the sense that $\eta_t = \phi_t^{\Lambda_L, \omega}(\eta)$ and the same for σ, where ω is the same realization of the random variables $\{v(s, b)\}$ and $\{\xi(s, C)\}$ for both. The quantity $p(L, t)$ is in some sense a way to measure the memory of the dynamics of the initial condition. It is in fact easy to see that[4]

$$\sup_{\sigma} |\mu_{\Lambda_L}^{+}(f) - E_\sigma f(\sigma_t)| \leqslant 2\, |f|_\infty\, p(L, t) \tag{3.2}$$

Thus, it will be sufficient to prove

$$p(L, t) \leqslant \exp[-m(\beta)\, t^\alpha] \qquad \forall t > t(L) \tag{3.3}$$

Following ref. 4, the above inequality will be proved by means of a multiscale analysis similar to that involved in the proof of Theorem 2.1. Let the length and time scales L_j, t_j be as in the previous section: $L_j = 4^{j^2}$, $t_j = 3^j$. Then we will set

$$p_j \equiv \sup_{(L_j \leqslant L < L_{j+1})} \max\{p(L, t_j), p(L, t_j, (+, j-1))\} \tag{3.4}$$

where $p(L, t(+, j-1))$ is the same quantity as $p(L, t)$ but computed for the $(+, j-1)$-dynamics in Λ_L. It is quite simple to relate $p(L, t)$ to p_j for suitable j. To this purpose, let, for a given integer L, $k \equiv k(L)$ be such that $L_k \leqslant L < L_{k+1}$ and let for any integer t, $n(t) = [t/t_k]$; then, using the Markov property, we obtain

$$p(L, t) \leqslant (p_k)^{n(t)} \tag{3.5}$$

Therefore the theorem follows if we can show that

$$p_j \leqslant \exp[-m(\beta)2^j] \tag{3.6}$$

for any j large enough, where $m(\beta)$ is a suitable constant uniformly bounded away from zero for all large enough β.

The basic inequality (3.6) will in turn follow from the usual recursive inequality

$$p_{j+1} \leqslant (2L_{j+2})^d p_j^2 + \exp[-m'(\beta)2^j] \tag{3.7}$$

provided that for some finite j_0, p_{j_0} is small enough. That is certainly true, since for finite j_0 and sufficiently large β the dynamics [or the $(+, j_0 - 1)$-dynamics] is undistinguishable from the dynamics at zero temperature.

Thus let us prove (3.7). To this end, let us fix one of the two dynamics [i.e., the usual SW dynamics with + b.c. or the $(+, j)$-dynamics] in the box Λ_L, $L \in [L_{j+1}, L_{j+2}]$; let us fix two arbitrary initial conditions η and σ; and let us introduce the event Ω_0 defined by

$$\begin{aligned}\Omega_0 = \{&\text{there exists } x \in \Lambda_L \text{ such that either}\\ &\Phi_{3t_j, t_j}^{\Lambda_j^x, +}(\eta_{t_j})(x) \neq \eta_{3t_j}(x) \text{ or } \Phi_{3t_j, t_j}^{\Lambda_j^x, +}(\sigma_{t_j})(x) \neq \sigma_{3t_j}(x) \text{ or both}\}\end{aligned} \tag{3.8}$$

where Λ_j^x is the unique box of side $2L_j$ inside Λ_L containing x and maximizing the distance of x from the part of its boundary not contained in the boundary of Λ_L, and $\Phi_{s,t_j}^{\Lambda_j^x +}(\eta_{t_j})(x)$ denotes the $(+, j-1)$-evolution in the time interval $[t_j, 3t_j]$ in the box Λ_j^x of the restriction to the box Λ_j^x of the configuration η_{t_j}. It is most important to outline that the random flow $\Phi_{3t_j, t_j}^{\Lambda_j^x +}(\cdot)$ is coupled to the random flow $\Phi_{3t_j, t_j}^{\Lambda_L}(\cdot)$ (see Remark 1, Section 1) because they use the same sequence of random variables $\{\nu(s, b)\}$, $\{\xi(s, C)\}$.

With the even Ω_0 so defined, we write

$$P(\eta_{t_{j+1}} \neq \sigma_{t_{j+1}}) = P(\Omega_0^c \cap \{\eta_{t_{j+1}} \neq \sigma_{t_{j+1}}\}) + P(\Omega_0) \tag{3.9}$$

The event appearing in the first term in the rhs of (3.9) is contained in the event that for some x in Λ_L,

$$\Phi_{3t_j, t_j}^{\Lambda_j^x, +}(\eta_{t_j})(x) \neq \Phi_{3t_j}^{\Lambda_j^x, +}(\sigma_{t_j})(x)$$

Thus, using the Markov property and the definition of p_j, we get

$$P(\Omega_0^c \cap \{\eta_{t_{j+1}} \neq \sigma_{t_{j+1}}\}) \leqslant (2L)^d p_j^2 \tag{3.10}$$

It remains to estimate the probability of Ω_0. What we will prove below is that the event Ω_0 implies that for some $s \in [t_j, 3t_j]$, η_s has a cluster not connected to the boundary of Λ_L of diameter greater than L_{j-1}. The probability of this last event, which we call Ω_1, is estimated, using Theorem 2.1, by

$$P(\Omega_1) \leqslant 2(2t_j)(2L)^d \exp[-m(\beta)2^{j-1}] \tag{3.11}$$

for β large enough.

It is clear that (3.6) follows from (3.9), (3.10) if we take $m'(\beta) = m(\beta)/2$ and $j > j_0$ with j_0 large enough independently of β.

Thus we are left with the proof of the inclusion

$$\Omega_0 \subset \Omega_1 \tag{3.12}$$

Lemma 3.1. If for any x in Λ_L and any s in $[t_j, 3t_j]$, $\mathrm{diam}(C_s(x)) \leqslant L_{j-1}$, then

$$\Phi_{3t_j, t_j}^{\Lambda_j^x, +}(\eta_{t_j})(x) = \eta_{3t_j}(x) \tag{3.13}$$

Proof. The proof is similar to Lemma 2.3 and it is by induction. For simplicity, we only discuss the case where x coincides with the center of the box Λ_j^x; the other case require only minor geometric modifications to the argument.

Let us assume as induction hypotheses that

$$\Phi_{s, t_j}^{\Lambda_j^x, +}(\eta_{t_j})(y) \tag{3.14}$$

for any y inside $\Lambda_j^x - 2(s - t_j)L_{j-1}$ and let us show that it propagates also to $s+1$. Let y be given inside $\Lambda_j^x - 2(s+1-t_j)L_{j-1}$; then from the inductive hypothesis, at time $s+1/2$ the bond variables inside the box $\Lambda_j^x - 2(s-t_j)L_{j-1}$ for η_s and $\Phi_{s, t_j}^{\Lambda_j^x, +}(\eta_{t_j})$ are equal. Therefore, if the cluster $C_s(y)$ computed for η_s has diameter less than L_{j-1}, then necessarily it must coincide with the cluster of y computed for $\Phi_{s, t_j}^{\Lambda_j^x, +}(\eta_{t_j})$ and in this case $\eta_{s+1}(y) = \Phi_{s, t_j}^{\Lambda_j^x, +}(\eta_{t_j})(y)$ by construction. If diam $(C_s(y)) > L_{j-1}$, then it must touch the boundary of Λ_L; on the other hand, again by the inductive hypothesis, also the cluster of y computed for $\Phi_{s, t_j}^{\Lambda_j^x, +}(\eta_{t_j})$ must have diameter greater than L_{j-1}. Therefore, by construction, $\eta_{s+1}(y) = \Phi_{s, t_j}^{\Lambda_j^x, +}(\eta_{t_j})(y)$, since the random flow $\Phi_{s, t_j}^{\Lambda_j^x, +}$ sets all clusters with diameter greater than L_{j-1} equal to plus one. The lemma is proved.

4. ABSENCE OF ERGODICITY IN INFINITE VOLUME AT LOW TEMPERATURE

In this section we apply the result of Theorem 2.3 to prove that the SW dynamics on the whole lattice $\mathbf{Z}^d$ is not ergodic at low temperature. Contrary to the case of short-range attractive Glauber dynamics as in the Metropolis or the heath bath algorithms, where the nonergodicity is easily proved using the attractivity of the dynamics, the problem of the ergodicity of the SW dynamics in the infinite volume at zero external magnetic field is not trivial and, moreover, in order to be well defined, it needs a prescription for the updating of clusters of infinite size. In fact, when a constant external field is applied, then any infinite cluster will flip in the direction of the field with probability one; obviously this is no longer true in the absence of the field. The prescription for the updating of an infinite cluster depends on course on which of the various Gibbs states of the Ising model at low temperature we choose as a candidate for the invariant measure for our dynamics. For example, if we decide that any infinite cluster is set equal to plus one with probability $p \in [0, 1]$, then it is natural to ask whether the invariant measure μ_β of the dynamics in $\mathbb{Z}^2$ (if it exists) is given for β large by

$$\mu_\beta = p\mu_+ + (1-p)\mu_- \tag{4.1}$$

where μ_+ and μ_- are the extremal states of the two-dimensional Ising model.

We will prove below that this is not the case for $p \neq 1/2$; more precisely, we will show that for $p \neq 1/2$ it is possible to find two different initial configurations σ and η such that $E(\sigma_t(0)) = 0$ and $|(E\eta_t(0))| > c > 0$ for any integer time t and a suitable constant c depending on β. It is clear that such a result cannot depend only on the local properties of the dynamics, but that it involves a control of the dynamics arbitrarily far away from the origin uniformly in the time t.

Let η and σ be given by

$$\eta(x) = 1 \qquad \forall x \in \mathbf{Z}^d$$

$$\sigma(x) = (-1)^k \qquad \text{with } d(0, x) \in ((k-1)^2, k^2], \quad k = 2,...$$

and let $P(\xi(s, C) = 1) = 1 - p$ whenever $\mathrm{diam}(C) = \infty$. Then we have the following result.

Theorem 4.1. There exists $\beta_0 < \infty$ such that for any $\beta > \beta_0$ and any $p \in [0, 1]$ there exists $c(\beta, p) \in (0, 1)$ with $\lim_{\beta \to \infty} c(\beta, p) = |2p - 1|$ such that

$$\text{(a)} \quad |E(\eta_t(0))| \geqslant c(\beta, p) \qquad \forall t$$

$$\text{(b)} \quad |E(\sigma_t(0()))| = 0 \qquad \forall t$$

Proof. (a) We write

$$|(E\eta_t(0))| = |(p-(1-p))|\, P_\eta(\text{diam}(C(t-1/2, 0)) = \infty) \tag{4.2}$$

and the result then follows from Corollary 2.1.

(b) For simplicity we only describe the two-dimensional case. Using (4.2), it is sufficient to prove that $P_\sigma(\text{diam}(C(t-1/2, 0)) = \infty) = 0$ for any t. Let us assume that for any $s \leqslant t-1/2$ and any $x \in \mathbf{Z}^d$, $P_\sigma(\text{diam}(C(s, x)) = \infty) = 0$; let us fix $k \geqslant 1$; and let Λ_k be equal to the anulus $\{x;\ \|x\| \in [(k-1)^2 + k/4,\ k^2 - k/4]\}$. Then it is not difficult to see, by the same methods used in the proof of Lemma 2.3 and 3.1, that we can couple the dynamics starting from σ in the whole lattice $\mathbf{Z}^d$ with the dynamics in Λ_k with (1/2, 1/2) b.c. in such a way that if the latter did not have a path of vacant bonds of length greater than $k/2t$ before time t, then up to time $t-1/2$ the two dynamics are identical for all x in Λ_k with $\text{dist}(x, \partial\Lambda_k) > k/4$. Since the starting configuration σ is homogeneous inside Λ_k, we can use Theorem 2.2 to get that the probability for the above event to happen tends to 1 as $k \to \infty$ for t fixed.

In conclusion, for k large enough depending to t, for each time s less than or equal to $t-1/2$, it is possible to find with large probability a path $\gamma_k(s)$ inside Λ_k at a distance from its boundary greater than $k/4$ of nearest neighbor sites $x_1, x_2,...$ encircling the origin with $\sigma_s(x_i) = \sigma_s(x_j)$. The same is of course true for $k+1$. Since, however, the value of σ at time $t=0$ inside Λ_k is equal to minus the value of σ inside Λ_{k+1}, then with probability greater than $(1/2)^t$, we can impose that for each s less than t, the value of σ attained in the sites belonging to the path $\gamma_k(s)$ is opposite to its value in the sites belonging to the path $\gamma_{k+1}(s)$. Thus, with probability one, for any given k_0, it is possible to find an integer $k > k_0$ and two paths of nearest neighbor sites γ_k and γ_{k+1} inside Λ_k and Λ_{k+1}, respectively, encircling the origin, on which the configuration σ_t assumes opposite values. Thus, for any $x \in \mathbf{Z}^d$, $P(\text{diam}(C(t, x)) = \infty) = 0$ and (b) follows.

Remark 1. It would be interesting to generalize the above results to different initial conditions, e.g., spin configurations with a density of spin flips either very low or very high. This does not seem to be a trivial question, since already for a chessboard spin configuration in two dimensions it is not clear whether the dynamics is able to create in a finite time an infinite cluster already at zero temperature.

We conclude this section with two positive results concerning the long-time behavior of the probability distribution of the algorithm at time t in two dimensions. The first says that, starting from a configuration identically equal to plus or minus one, the time distribution converges weakly to the right mixture of the extremal states, while the second states that the convergence holds no matter which starting configuration we choose, provided that we consider the dynamics given by the (1/2, 1/2) rule.

Theorem 4.2. Let η_t denote the evolution at time t according to the $(p, 1-p)$ rule in two dimensions, where η is identically equal to either plus or minus one. Let A be a cylindrical event. Then for β large anough we have:

(a) $\lim_{t\to\infty} P(\eta_t \in A) = p\mu_+(A) + (1-p)\,\mu_-(A)$.

(b) There exists $t_0(A)$ such that for any $t > t_0(A)$ and some positive m

$$|P(\eta_t \in A) - p\mu_+(A) - (1-p)\,\mu_-(A)| \leqslant \exp(-mt^\alpha)$$

where α is as in Theorem 3.1.

Theorem 4.3. With the same notations as in previous theorem, let p be equal to 1/2. Then we have:

(a) $\lim_{t\to\infty} P(\sigma_t \in A) = \frac{1}{2}\mu_+(A) + \frac{1}{2}\mu_-(A)$, $\forall \sigma \in \{-1, +\}^{z^2}$.

(b) There exists $t_0(A)$ such that for any $t > t_0(A)$ and some positive m

$$|P(\sigma_t \in A) - \tfrac{1}{2}\mu_+(A) - \tfrac{1}{2}\mu_-(A)| \leqslant \exp(-mt^\alpha)$$

where α is as in Theorem 3.1.

Proof of Theorem 4.2. Let L_k and t_k be the usual length and time scales, let k_0 be so large that the event A depends only on the spins inside Λ_{k_0-1}, and let t be such that $t_k < t \leqslant t_{k+1}$ for some $k > k_0 + 1$. By our standard coupling technique (see Lemmas 2.3 and 3.1), we can couple the $(p, 1-p)$ rule in the full lattice and the dynamics in Λ_{k-1} with $(p, 1-p)$-b.c. starting from η in such a way that if for both dynamics any site x of the box Λ_{k-1} and any $s \leqslant t$ either $\operatorname{diam}(C_s(x)) < L_{k-2}$ or x belongs to the infinite cluster (boundary cluster), then a time t they must coincide inside the box Λ_{k_0}. Using Theorem 2.2, we find that the probability for this not to happen is bounded from above by

$$t_{k+1}(2L_{k-1})^2 \frac{1}{L_{k-2}^{2a}} \exp[-m_0(\beta)\,2^{k-2}] \tag{4.3}$$

Therefore we immediately get that

$$|P(\eta_t \in A) - P(\eta_t^{\Lambda_{k-1}} \in A)| < t_{k+1}(2L_{k-1})^2 \frac{1}{L_{k-2}^{2a}} \exp[-m_0(\beta)\,2^{k-2}] \tag{4.4}$$

Moreover, using Theorem 3.1, we have

$$|P(\eta_t^{\Lambda_{k-1}} \in A) - p\mu_+^{\Lambda_{k-1}}(A) - (1-p)\,\mu_-^{\Lambda_{k-1}}(A)| \leqslant 2\exp[-m(\beta)t^{\alpha}] \tag{4.5}$$

with α as in Theorem 3.1.

Thus, part (a) follows from the triangle inequality and the standard fact that

$$|\mu_+^{\Lambda_{k-1}}(A) - \mu_+(A)| < \exp[-c(\beta)L_{k-1}]$$

as $k \to \infty$ and the same for $\mu_-^{\Lambda_{k-1}}$.

Part (b) follows from the above explicit bounds.

Proof of Theorem 4.3. The proof is almost identical to the proof of the previous theorem, although some care has to be paid to the coupling argument. Let k_0 and k be as in proof of Theorem 4.2 and let, for $s > t_{k-1}$,

$$\tilde{\sigma}_s \equiv \Phi_{s,t_{k-1}}^{\Lambda_{k-1},p}(\sigma_{t_{k-1}})$$

be the evolution at time s in the box Λ_{k-1} of the configuration $\sigma_{t_{k-1}}$ according to the dynamics which sets the sign of all clusters with diameter greater than L_{k-2} equal to the sign of the cluster of the boundary of the box Λ_{k-1}. For $p=1$ this dynamics coincides with the usual $(+, k-2)$-dynamics. In analogy with Lemma 3.1, let Ω_1 be the event that, for the dynamics in the infinite volume, for any x in Λ_{k-1} and any s in $[t_{k-1}, t]$ either $\operatorname{diam}(C_s(x)) \leqslant L_{k-2}$ oer the outermost boundary of $C_s(x)$ does not intersect the boundary of the box Λ_{k-1}; in this last case $C_s(x)$ does not depend on x in the sense that if $\operatorname{diam}(C_s(y)) > L_{k-2}$, then $C_s(x) = C_s(y)$ (see, e.g., Lemma 2.1) and, with an abuse of notation, it will be denoted C_s^{∞}. If the event Ω_1 occurred, then we couple $\tilde{\sigma}_s$ with the dynamics in the full lattice σ_s in the usual way, namely:

(i) Bonds are made vacant at the same time for both, and equal clusters which do not touch the boundary of Λ_{k-1} do the same thing.

(ii) The sign of the boundary cluster for $\tilde{\sigma}_s$ is the same as the sign of the cluster C_s^{∞} for σ_s.

Remark 2. It is at this stage that $p = 1/2$ is important. In fact, the cluster C_s^{∞} may or may not be an infinite cluster; however, if $p = 1/2$, this will not affect its probability distribution.

With this coupling it is easy to prove, following the proof of Lemma 3.1, that at time t, $\tilde{\sigma}_t(x) = \sigma_t(x)$ for any x in Λ_{k_0}. Thus we get

$$|P(\tilde{\sigma}_t \in A) - P(\sigma_t \in A)| < P(\Omega_1) < t_{k+1}(2L_{k-1})^2 \frac{1}{L_{k-2}^{2a}} \exp[-m_0(\beta)2^{k-2}] \tag{4.6}$$

where we used Theorem 2.1 in order to get the last inequality.

Next we compare $P(\tilde{\sigma}_t \in A)$ with $P(\sigma_t^{\Lambda_{k-1}} \in A)$, where $\sigma_t^{\Lambda_{k-1}}$ is the evolution at time t of σ in the box Λ_{k-1} with the usual dynamics with (1/2, 1/2)-b.c. Using (3.5), we get that

$$|P(\tilde{\sigma}_t \in A) - P(\sigma_t^{\Lambda_{k-1}} \in A)| < p_{k-2} < \exp(-m2^{k-2}) \tag{4.7}$$

Thus the triangle inequality together with (4.6) and (4.7) yields the analog of (4.5) and the rest of the proof is the same as the proof of Theorem 4.2.

5. EXTENSION OF THE RESULTS TO OTHER MODELS

In this final section we briefly discuss the extension of the ideas and results presented in the previous sections to a different model of random cluster dynamics introduced in ref. 19, sharing with the SW algorithm without external field the property that the updating of clusters of dynamical variables (particles in our case) occurs with a probability independent of the geometry of the cluster. The two dynamics are, however, very different one from the other; the invariant measure of the SW dynamics is the usual Gibbs measure of the Ising model, while the invariant measure of the second dynamics in dimension greater than two is not the Gibbs state for any absolutely summable interaction.[19]

We do this in order to illustrate with a concrete example the genuine nonequilibrium character of our techniques and to show that they work equally well for a dynamics which is not reversible with respect to an *a priori* probability measure.

The setting is as follows: at each point x in the box $\Lambda = [-L, L]^d \cap \mathbf{Z}^d$ we associate an occupation variable $\sigma(x)$ with values 0 or 1; given a configuration σ_t at time t, in order to define the new configuration σ_{t+1} at time $t+1$ we first consider all connected clusters of particles (site in which the configuration σ_t is equal to one) and we remove each cluster independently with probability 1/2; as a second step we create particles in each empty site independently with probability p.

The above dynamics is similar to a model considered by Swindle and Grannan,[21] although in their model clusters disappear with a rate proportional to their size. We were primarily interested in the long-time behavior of the above stochastic cluster dynamics and in particular in questions like ergodicity, approach to equilibrium, and mixing properties of the invariant measure. In turned out that in order to carry out this program it is crucial to have a good control of the range of the interaction, namely of the typical size of the clusters. In the one-dimensional case we could prove by means of a novel path expansion in space-time that the probability that the origin belongs to a big cluster consisting of N particles is bounded by a negative exponential in N for any $p \in (0, 1)$. As a consequence we can then prove for all p the exponential convergence as t tends to infinity of the distribution of the process at time t to the unique invariant measure together with the exponential decay of correlations of the latter.

In two or more dimensions the situation changes radically. For any $p \in (0, 1)$ we prove that the above probability cannot be bounded from above by a negative exponential in the number of sites of the cluster. More precisely, if $p_N(t)$ denotes the probability that the cube Q_N of side N centered at the origin is filled with particles at time t, then we prove that for suitable constants c_1, c_2, α, and α if $t \geqslant c_1 N$ we have

$$P_N(t) \geqslant \exp(-c_2 N) \qquad \forall p > 0 \tag{5.1a}$$

$$\frac{1}{N^\beta} \geqslant p_N(t) \geqslant \frac{1}{N^\alpha} \qquad \text{if} \quad 1-p \ll 1 \tag{5.1b}$$

However, the question of an upper bound for the above probability for small values of p remained open, together of course with the problem of existence and uniqueness of the invariant measure in the same range of values of p. It is our goal here to fill this gap by stating results very similar to those already explained for the SW dynamics; the proofs are, however, a duplicate of those given in the previous sections, and we will therefore only describe the results and the main ideas.

Let C_Λ be the collection of all possible connected subsets I of Λ. Here I is connected iff for any two sites x and y there exists a path of nearest neighbor sites in I going from x to y.

Let also $\{v(x, s)\}_{x \in \Lambda, s \in N}$ and $\{\xi(I, s)\}_{I \in C_\Lambda, s \in N}$ be i.i.d. random variables with values in $\{0, 1\}$ with

$$P(v(x, s) = 1) = p, \qquad P(\xi(I, s) = 1) = \tfrac{1}{2}$$

For brevity, a realization of the $v(x, s)$ $[\xi(I, s)]$ variables will be denoted by v (ξ). On each site x of Λ we will associate an occupation variable $\sigma(x)$ taking values in $\{0, 1\}$; for brevity the collection of the variables $(\sigma(x))_{x \in \Lambda}$ will be denoted by σ. Thus, σ is an element of the configuration space $S = \{0, 1\}^\Lambda$. Using the random variables v, ξ, we now construct on S a random dynamics starting at the configuration σ at time $t = 0$ as follows:

(i) Given $\sigma_t^\Lambda \in S$, we set for any $x \in \Lambda$

$$\sigma_{t+1/2}^\Lambda(x) = 1 \qquad \text{if} \quad \sigma_t^\Lambda(x) = 1 \quad \text{and} \quad \xi(I_x, t) = 1$$

where I_x is the maximal element of C_Λ containing x such that $\sigma_t^\Lambda(y) = 1$, $\forall y \in I$.

(ii) For any $x \in \Lambda$

$$\sigma^{\Lambda}_{t+1}(x) = 0 \qquad \text{iff} \quad \sigma^{\Lambda}_{t+1/2}(x) = 0 \quad \text{and} \quad \nu(t+1, x) = 0$$

For brevity we will refer to part (i) of the updating as the killing of particles and to part (ii) as the creation of particles. Note that both processes occurs simultaneously (i.e., the updating is parallel) and that the nontrivial interaction fo the model is all contained in the killing process.

We will refer to the above rules as the "basic dynamics in Λ." The associated Markov process will always be denoted by σ_r, omitting the suffix Λ for brevity whenever it does not produce confusion. The time t will always take integer values; however, sometimes we will consider events involving the values of the process both at time t and at time $t + 1/2$.

It is very easy to check that in any finite volume Λ there exists a unique invariant measure, which will be denoted by μ_{Λ}.

Later, when discussing the approach to equilibrium for the process, we will need to compare the dynamics of a given site x produced by two different bowes Λ and Λ' with $\Lambda' \subset \Lambda$ both containing x. This will be done by establishing a coupling between the two dynamics according to the following rules:

(a) The variables $\nu(x, s)_{x \in \Lambda'}$ are exactly the same variables that one chooses for the dynamics in Λ, i.e., if a particle is created inside Λ' for the dynamics in Λ, then it is created also for the dynamics in Λ' and vice versa.

(b) The value of $\xi(I, s)$ is the same for both dynamics if $I \subset \Lambda'$.

In some sense the above coupling is the most natural way to restrict the dynamics in Λ to Λ'.

In one dimension, however, there is a more efficient way to realize this coupling in such a way that the value of the process at a given site x inside Λ' will always be equal for the two dynamics. We will now state a basic estimate on the probability that the cluster of a given fixed point x at some time t has a diameter greater than L, with t greater than some scale $t(L)$ related to L. Let us first fix some notations similar to those used in Section 2.

For any integer k we define

$$p_k = \sup_{l \geqslant L_k;\, x \in \Lambda_L;\, t \geqslant t_k;\, \sigma \in \{-1.1\}^{\Lambda_L}} P(L, x, t, k, \sigma)$$

where $P(L, x, t, k, \sigma)$ is the probability that the cluster of particles containing the site x has diameter greater than L_k, where L_k and t_k are as in Section 2. Then the following result holds:

Theorem 5.1. There exist $p_0 > 0$, $c > 0$, $k_0 > 0$, and $a > 0$ such that for any $p < p_0$ there exists a constant $m \equiv m(p)$ with $m > c$ such that

$$p_k < \frac{1}{L_k^{2a}} \exp(-m2^k) \qquad \forall k > k_0$$

Corollary 5.1. For p sufficiently small there exists a unique invariant measure μ for the basic dynamics in $\mathbb{Z}^d$ such that:

(a) Given a cylindrical event A, there exists $t(A)$ such that

$$|P(\sigma_t \in A) - \mu(A)| < \exp(-mt^\alpha)$$

for some positive m and any $t > t(A)$, where $\alpha = \ln(2)/\ln(3)$.

(b) Let f and g be local observables depending only on the value of the configuration $\sigma(x)$ for x in A and in B, respectively, with $d(A, B) = L$, and let $|f|_\infty$ denote the sup norm. If $\langle f; g\rangle$ denotes the expression

$$\int d\mu(\sigma)\, f(\sigma)\, g(\sigma) - \left[\int d\mu(\sigma)\, f(\sigma)\right]\left[\int d\mu(\sigma)\, g(\sigma)\right]$$

then we have

$$\langle f; g\rangle \leqslant \exp[-m(p) 2^{k(L)}]$$

for a suitable positive constant $m(p)$, where $k(L)$ is such that $L_k \leqslant L < L_{k+1}$.

Sketch of the Proof of Theorem 5.1. The proof follows step by step the proof of Theorem 2.2 with the simplification that now it is much more straightforward to prove that the pieces of the cluster inside the boxes Λ_1 and Λ_2 have been created by the dynamics independently one from the other. Lemmas 2.1 and 2.2 are in fact no longer necessary and, more important, there is no longer any need to define the analog of the random variables $\xi'(s, C)$.

Sketch of the Proof of Corollary 5.1. Part (a) is proved in a very similar way to the proof of Theorem 4.3:

(i) One first shows, following the proof of Theorem 3.1, that in a finite cube of side L two arbitrary initial configurations become identical after a time

$$t(L) = \exp\left[\ln(3)\left(\frac{\ln(L)}{\ln(4)}\right)^{1/2}\right]$$

with large probability (as $L \to \infty$) and thus one establishes a fast convergence to equilibrium.

(ii) Using Theorem 5.1, one then proves that locally the dynamics σ_t at time t in the infinite lattice is, with probability tending to one as $L \to \infty$ uniformly in the initial configuration σ, identical to the dynamics in the cube of side L starting at time $t - t(L)$ from $\sigma t(L)$, provided that $t > t(L)$ and

$$t(L) = \exp\left[\ln(3)\left(\frac{\ln(L)}{\ln(4)}\right)^{1/2}\right]$$

(iii) By combining (i) and (ii), one approximates the average at time

$$t > \exp\left[\ln(3)\left(\frac{\ln(L)}{\ln(4)}\right)^{1/2}\right]$$

of a local observable with its average with respect to the equilibrium measure associated to dynamics in the cube of side L, with an error that tends to zer as $L \to \infty$ uniformly in the initial configuration σ.

(iv) Step (iii) immediately implies the convergence as $t \to \infty$ of the average at time t of any local observable uniformly in the initial configuration and thus also the uniquiness of the invariant measure. It also follows that the finite-volume equilibrium measure converges weakly to the invariant measure of the process in the infinite volume.

Part (b) follows simply by the explicit bounds that one has in the previous step.

REFERENCES

1. R. H. Swendsen and J. S. Wang, Nonuniversal critical dynamics in Monte Carlo simulations, *Phys. Rev. Lett.* **58**:86 (1987).
2. R. G. Edwards and A. D. Sokal, Generalization of the Fortuin–Kasteleyn–Swendsen–Wang representation and Monte Carlo algorithm, *Phys. Rev. D* **38**:2009 (1988).
3. J. Godman and A. D. Sokal, Multigrid Monte Carlo method: Conceptual foundations, *Phys. Rev. D* **40**:2035 (1989).
4. F. Martinelli, E. Olivieri, and E. Scoppola, On the Swendsen and Wang dynamics I. Exponential converges to equilibrium, *J. Stat. Phys.* **62**:117 (1991).
5. F. Martinelli, E. Oliveri, and E. Scoppola, On the Swendsen and Wang dynamics II: Critical droplets and homogeneous nucleation at low temperature for the 2 dimensional Ising model, *J. Stat. Phys.* **62**:135 (1991).
6. P. W. Kasteleyn and C. M. Fortuin, On the random cluster model. I: Introduction and relation to other models, *Physica* **57**:536 (1972).
7. P. W. Kasteleyn and C. M. Fortuin, *J. Phys. Soc. Jpn. Suppl.* **26**:11 (1990).

8. R. H. Swendsen and J. S. Wang, Cluster Monte Carlo algorithms, *Physica A* **167**:565 (1990).
9. U. Wolff, *Phys. Lett. B* **62**:361 (1989).
10. D. Kandel, R. Ben-Av, and E. Domany, Cluster dynamics for fully frustrated systems, *Phys. Rev. Lett.* **65**:941 (1990).
11. U. Wolff, *Phys. Lett. B* **228**:379 (1989).
12. M. D'Onofrio De Meo, D. W. Heermann, and K. Binder, Monte Carlo study of the Ising model phase transition in terms of the percolation transition of "physical clusters," *J. Stat. Phys.* **60**:585 (1990).
13. Xiao-Jian Li and A. D. Sokal, Rigorous lower bound on the dynamical critical exponent of the Swendsen–Wang dynamics, *Phys. Rev. Lett.* **63**:827 (1989).
14. R. Von Dreyfus, On the effect of randomness in ferromagnetic models and Schröedinger operators, Ph.D. thesis, New York (1987).
15. E. Jordao-Neves and R. Schonmann, Critical droplets and metastability for a Glauber dynamics at very low temperature, *Commun. Math. Phys.*, to appear.
16. D. A. Huse and D. S. Fisher, *Phys. Rev. B* **35**:6841 (1987).
17. A. D. Sokal and L. E. Thomas, Absence of mass gap for a class of stochastic contour models, *J. Stat. Phys.* **51**:907 (1988).
18. A. L. Toom, Stable and attractive trajectories in multicomponents random fields, in *Multicomponent Random Fields*, R. L . Dobrushin and Ya. G. Sinai, eds. (Marcel Dekker, New York, 1980).
19. F. Martinelli and E. Scoppola, A simple stochastic cluster dynamics: rigorous results, CARR Reports in Mathematical Physics 24/90 Rome; *J. Phys. A: Math. Gen.*, to appear.
20. J. L. Lebowitz and R. H. Schonmann, Pseudo-free energies and large deviations for non-Gibbsian FKG measure, *Prob. Theory Rel. Fields* **77**:49 (1988).
21. G. Swindle and E. Grannan, A particle system with massive destruction, *J. Phys. A: Math. Gen.* **23**:L73 (1990).