Simulations without critical slowing down: Ising and three-state Potts models

Daniel Kandel and Eytan Domany
Department of Electronics, Weizmann Institute of Science, Rehovot 76100, Israel

Achi Brandt
Department of Applied Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel
(Received 25 January 1989)

We have developed a novel simulation method that combines a multigrid technique with a stochastic blocking procedure. Our algorithm eliminates critical slowing down completely, as we demonstrated previously by simulating the two-dimensional Ising model at criticality. Here we describe results of much more extensive simulations of the two-dimensional Ising and three-state Potts models and provide a heuristic argument to explain why our method works.

I. INTRODUCTION

Simulation techniques have considerable importance in the study of many-body systems, since rigorous analytic results can be obtained only for the simplest models. In particular, Monte Carlo simulations are widely used to calculate properties of systems in statistical physics[1] and in lattice gauge theories.[2] One of the major problems that simulations must overcome is that of system size; in a variety of cases the main theoretical (and experimental) interest in a model lies in the thermodynamic limit, i.e., when the number of degrees of freedom is macroscopic. To reliably estimate the behavior in the thermodynamic limit one has to perform simulations of large systems. Such large-scale simulations become prohibitively long (in terms of computer time) and expensive, since the basic time unit is proportional to the system size. Beyond this trivial scaling with size, simulations are even more problematic near critical points, because of *critical slowing down* (CSD). This phenomenon causes the divergence of a relaxation time τ as the critical point is approached.[3] Thus the times needed to equilibrate the system and to generate statistically independent configurations at equilibrium become exceedingly large. At critically, τ grows as

$$\tau \sim L^z \tag{1.1}$$

with the linear size L of the system. Here τ is measured in units that scale with the total number of sites (i.e., L^d), and z is the dynamic critical exponent.

In the last few years there have been several attempts to invent simulation techniques that eliminate CSD.[4-7] The physical origin of CSD is the manner in which the system moves in phase space, e.g., by changing one degree of freedom (spin) at a time. At a critical point fluctuations that involve creation of arbitrarily large correlated domains have significant statistical weight; reaching such configurations by "incoherent" single-spin moves takes a long time. Hence, in order to eliminate CSD one should clearly seek an algorithm that performs collective, nonlocal moves. Multigrid techniques[8-10] appear to be a sensible choice for such a task. They were developed for accelerating iterative solvers of partial differential equations, where slowness arises from similar reasons to the one discussed above. One of the first successful applications of multigrid methods in statistical mechanics, aimed at accelerating convergence to equilibrium of a model, was carried out by Goodman and Sokal.[4] They applied such a method to the ϕ^4 model on the square lattice and reduced the relaxation time by an order of magnitude; they did not, however, reduce the value of z. On the other hand Swendsen and Wang[6] (SW) did succeed in reducing z. They generalized a method due to Fortuin and Kasteleyn, based on a mapping between percolation and Potts models,[11] to accelerate simulations of the latter. For the Ising model on the square lattice, for example, their method yields $z \approx 0.35$ rather than $z \approx 2.1$, the value obtained in standard simulations.[12,13]

We have recently demonstrated that CSD can be eliminated completely for a nontrivial model.[14] Our algorithm combines multigrid ideas with a stochastic blocking (coarsening) technique similar to the one used by SW. In fact, the SW procedure can be viewed as a special case of ours. In this sense the new method is a generalization of the SW method, but it is essentially different from other known generalizations.[15] Ideas underlying our approach were first presented in Appendix B of Ref. 9. General aspects of the method are outlined there in a manner applicable to discrete, as well as continuous-state Hamiltonians, with a wider class of coarse-to-fine interpolations (not just simple blockings).

The basic idea, or strategy, of our method is as follows. We perform a few Metropolis sweeps[16] of the original lattice, using its associated Hamiltonian to generate (single-spin-flip) transitions in phase space. Next, a stochastic coarsening procedure is used to generate a new problem, with fewer degrees of freedom, coupled by a different Hamiltonian. This process is referred to as going from fine to coarser level, or *coarsening*. The reverse process, of restoring finer scale degrees of freedom and Hamiltonian, is called *uncoarsening*. In the course of our simulation we perform an organized self-similar sequence of coarsening and uncoarsening steps, with a few standard Metropolis sweeps at each level. Sweeps performed at a

level n use the Hamiltonian $\mathcal{H}_n$ that corresponds to that level to generate transitions in the phase space of the level being simulated.

This paper contains a detailed description of the technique and its performance on the two-dimensional Ising and three-state Potts models. We describe the algorithm in Sec. II and show in Sec. III that it satisfies detailed balance. Section IV contains results of extended simulations, performed on the (square lattice) Ising and three-state Potts models. All simulations were done at the critical temperature of the infinite lattice. We describe the manner in which we analyzed the data, and determined relaxation times for systems of varying sizes $4<L<128$. We measured equilibration times of the energy and the susceptibility, and demonstrated that they did not diverge as L was increased. We also studied the relaxation of time-delayed energy-energy correlation functions *in equilibrium*. Section V presents an argument that explains why our algorithm does eliminate CSD completely, whereas the SW procedure does not. The argument is based on a scaling ansatz, whose validity is checked in Sec. VI, which also contains new results on the SW method.

II. THE ALGORITHM

As outlined in the Introduction, our algorithm combines a multigrid approach with a stochastic method that creates blocks of spins. In what follows, we first give a general description of the stochastic blocking techniques. Next we explain how it was implemented in our previously reported study of the square-lattice Ising model, and show how stochastic blocking was incorporated in the multigrid cycle.

A. The stochastic coarsening procedure

Our coarsening procedure has two effects. First, from a problem with L^d degrees of freedom ("fine spins") we go to a new problem, with a smaller number, $(L/b)^d$ of "coarse spins." Each new degree of freedom is, in effect, a block of the original fine spins, all aligned. This "freezing" of the fine spins is equivalent to restricting the phase space of the original problem. The second effect is the replacement of the full original Hamiltonian with a new one. Thus our process contains Monte Carlo simulations of a stochastically generated, simplified Hamiltonian, over a restricted phase space with fewer degrees of freedom.

To explain the manner in which coarsening is done, denote the fine-level Hamiltonian as

$$\mathcal{H}=\mathcal{H}_c+V\ , \tag{2.1}$$

where factors of $1/k_BT$ have been absorbed into $\mathcal{H}$ and where $\mathcal{H}_c$, the coarse-level Hamiltonian, is in some sense easier to simulate than the original $\mathcal{H}$. Assume that the fine system is in a state Q, obtained for example in a standard simulation. We now may eliminate, or "kill" the interaction V by either "deleting" it with probability

$$p_d(Q)=C_V e^{V(Q)} \tag{2.2}$$

or "freezing" it with probability $p_f(Q)=1-p_d(Q)$. If the interaction is frozen, only states Q' with $V(Q')=V(Q)$ are considered in the ensuing simulation. If the interaction is deleted, no such restriction is placed on the states. In either case, the thermodynamics are subsequently governed only by the simplified Hamiltonian $\mathcal{H}_c$ for a fixed number of steps, after which the original Hamiltonian $\mathcal{H}$ and the original space of states are restored. One must choose C_V so that $p_d,p_f\in[0,1]$ for any Q. It is shown in Sec. III that the rule (2.2), with any such C_V, guarantees the entire process to be statistically legitimate.

In practice, killing a single interaction V does not leave an $\mathcal{H}_c$ that can be simulated trivially (i.e., without CSD). To achieve that, additional terms of the Hamiltonian must be killed. A sensible course of action to generate an $\mathcal{H}_c$ that is obviously trivial to simulate was taken by Swendsen and Wang. They chose to kill *all* the interactions in $\mathcal{H}$; this gives rise to a system which consists of frozen blocks of degrees of freedom that are completely decoupled from each other. Such a system is, of course, trivial to simulate—but is subject to a large set of restrictions on its states. To obtain good statistics these restrictions must be averaged over, by repeating the process. It is clear that the largest choice of C_V (under the constraint $p_d,p_f\in[0,1]$) produces the best statistics, since with this choice the probability for deletion p_d is maximal. Then coarsening leads to a minimal set of restrictions on the states, and the number of coarsening-uncoarsening cycles needed is the smallest.

For the two-dimensional ferromagnetic Ising model

$$\mathcal{H}=-\sum_{\langle ij\rangle}K_{ij}s_is_j\ , \tag{2.3}$$

where $K_{ij}\geq 0$, the procedure described above can be carried out by killing the interactions $K_{ij}s_is_j$ one at a time. The optimal probability for deletion of bond (ij), then, is

$$p_d^{(ij)}=e^{-K_{ij}(1+s_is_j)}\ . \tag{2.4}$$

With this choice, interactions between antiparallel spins will always be deleted; only parallel spins have positive probability to be frozen together, thereby forming blocks of aligned spins. The relative orientation of spins within a block may not be changed before the original Hamiltonian is restored. Hence, each such block can be viewed as a single coarse spin. Once all the interactions have been killed, we are left with a system of completely decoupled coarse spins. Any random assignment of coarse-spin values is a legitimate Monte Carlo move of this new system. The new coarse-spin configuration corresponds to a new configuration of the fine spins. Using this new fine-spin configuration with the original Hamiltonian on the fine lattice, we can repeat the coarsening process. This is precisely the procedure followed by SW. As mentioned above, it still exhibits CSD, although with a considerably reduced dynamical exponent. We discuss in Sec. V the reason this procedure does not eliminate CSD completely.

B. Stochastic multigrid method

To eliminate CSD completely, we combine this stochastic blocking technique with multigrid ideas. We use

the coarsening procedure to transform a lattice with L^2 spins (the fine lattice) to a lattice with $(L/b)^2$ spins (the coarse lattice). Each coarse spin is associated with a (stochastically determined) block of fine spins, and a new Hamiltonian is defined on the coarse lattice. There are many ways to do this; here we describe the algorithm we introduced recently, leaving description of computationally more efficient ones for future publication.[17] Our method is demonstrated in Fig. 1. We start from a particular fine-spin configuration [Fig. 1(a)], and choose the boxed sites as our coarse lattice; here the length rescaling factor is $b=2$. This is strongly reminiscent of a decimation transformation used frequently in position-space renormalization-group calculations. We, however, do not trace over the unboxed spins; rather, we coarsen the system by visiting sequentially[18] all pairs of coupled sites i,j. At any such bond we encounter one of three possible situations. First, we may find that *both* i and j have already been frozen to different coarse spins. By this we mean that there exists a path of frozen bonds connecting site i to a boxed spin (and the same holds for j). In this case we move on to the next pair, leaving the bond i,j *alive*. The second possibility is that at least one of the two sites has *not* been frozen to a coarse spin; in this case we kill (i.e., freeze or delete; see above) the interaction $K_{ij}s_is_j$. The third possibility is that both i and j have been frozen to the *same* coarse spin. In this case the bond between them is frozen. Since the decision to freeze or delete a bond is stochastic, a coarsening step, performed on a given fine-spin configuration, may have different outcomes. One possible outcome of coarsening the fine-spin configuration of Fig. 1(a) is shown in Fig. 1(b). Single (double) lines between sites indicate living (frozen) bonds; deleted bonds are not marked. Each cluster of fine spins, connected by frozen bonds, constitutes an irregular, stochastically generated block [Fig. 1(c)]. A block which contains a boxed spin is viewed as a single coarse spin. It is easy to see that a block that is not frozen to any boxed spin has to be completely disconnected from the rest of the lattice. These disconnected blocks are not part of the coarse lattice.[19] They will be considered again only when we uncoarsen, i.e., return to the fine lattice. Then, each of them is set to an arbitrary value of spin.

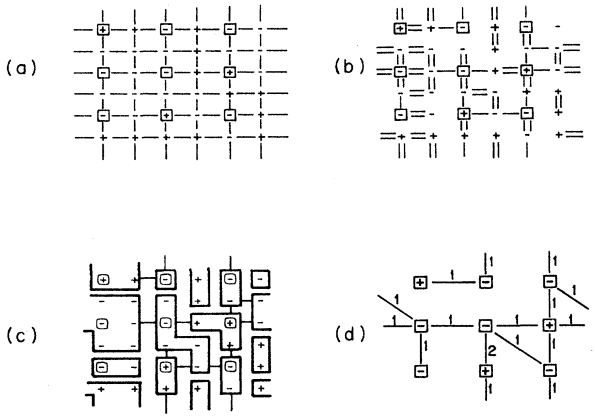

FIG. 1. A demonstration of the coarsening procedure. (a) A configuration of the fine lattice. Boxed spins constitute the coarse lattice and lines between sites denote bonds. (b) A possible coarsening. Single (double) lines denote living (frozen) bonds. Deleted bonds are not marked. (c) The corresponding block configuration, and the (original) bonds between the blocks. (d) The coarse lattice. Numerals near the lines denote coarse-lattice bond strengths.

Next, the coarse Hamiltonian is constructed: the coupling between coarse spins s_i and s_j is the sum of the couplings that are *alive* and connect fine spins associated with coarse spins s_i and s_j. This is illustrated in Fig. 1(d): numerals on the lines between sites indicate the number of fine couplings that contribute to the coarse bond. The new Hamiltonian is inhomogeneous, but contains pairwise interactions only, which may be of long range. However, finite couplings between distant spins are unlikely, and their probability drops exponentially with the distance. The same procedure can be reapplied to the coarse lattice to obtain multiply coarsened systems with their associated Hamiltonians. In all of them the average number of interactions per spin turns out to be small (less than 3).

Our algorithm moves back and forth between coarser and finer levels. In order to do this, we need, in addition to coarsening, also an "uncoarsening" prescription. Assume we arrived at a particular coarse-spin configuration. To "uncoarsen" the system, every decoupled block is set to some arbitrary value of spin, and all finer-lattice spins take the value of the block spin to which they were frozen. Finally, the couplings that were present at this finer level are restored.

The final ingredient of our procedure is the manner in which we move in phase space at any level of coarsening. This is done by standard Monte Carlo updates, performed at any length scale, using a Metropolis algorithm[16] on the coarse spins at that scale. The corresponding Hamiltonian is used to determine transition probabilities.

Our dynamic procedure "cycles" through all length scales, starting from the finest. At each intermediate length scale, one has to decide on the next step; whether one should continue by going to a coarser level or by uncoarsening. This decision is made according to the following rule: Uncoarsening will take place only if the last γ "visits" of the present level were continued by coarsening. Otherwise, one proceeds by coarsening. In other words, each level is coarsened γ times before it is uncoarsened. Figure 2 is an example of a cycle with $\gamma=2$ and $b=2$, where coarsening (uncoarsening) is denoted by an arrow pointing downwards (upwards). The finest lattice consists of 64 spins ($L=8$), and since $b=2$, there are three coarser levels ($L=4$, 2, and 1). Consider, for example the third level ($L=2$); it is coarsened twice before each time it is uncoarsened. The rationale for choosing such a self-similar cycle is explained in Sec. V. When the coarsest level is reached, the fine spins of the original problem are grouped into decoupled blocks, one of which belongs to the coarsest lattice. Each time this level is reached a new coarse-spin configuration is (trivially) generated, and it is immediately uncoarsened. The cycle

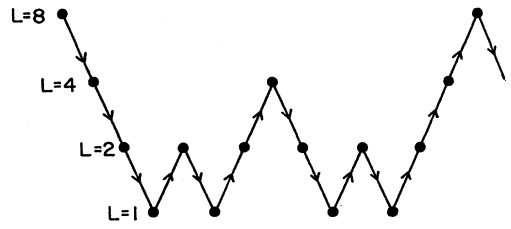

FIG. 2. An example of a cycle on an 8×8 lattice, with length rescaling factor $b=2$ and $\gamma=2$. Coarsening (uncoarsening) is denoted by an arrow pointing downwards (upwards).

ends when the finest level is reached. Whenever we arrive at a level, we perform a few Metropolis sweeps before going to the next level. The number of sweeps performed has some effect on the relaxation time, as described in detail in Sec. IV. If no Metropolis sweeps are used at all, transitions are generated only by coarsening and uncoarsening (note that during uncoarsening disconnected blocks may get flipped). Even with no sweeps, we get $z\approx 0.0$ (when using a $\gamma\geq 2$, $b=2$ cycle), but the approach to this result is slower (i.e., happens at larger sizes L), as discussed below.

Lastly, note that the SW procedure[6] corresponds to $\gamma=1$ with $b=L$, going directly from the finest to the coarsest level and back, without Metropolis sweeps at the finest level. In fact, we argue that any $\gamma=1$ cycle is equivalent to the SW procedure, independent of the rescaling factor used, as long as no Metropolis sweeps are performed. To see this, note that differences between the SW scheme and the $\gamma=1$ scheme with $b<L$ may arise only from the fact that whenever a coarse level is created, all fine bonds that connect the same two coarse spins are combined and treated as a single bond. On the other hand, in the SW procedure bonds are never combined, are treated separately. Consider two coarse spins that were created at a certain level of the $b<L$ cycle. In the SW scheme, at any intermediate stage of coarsening, the probabilities of freezing or disconnecting the two blocks, represented by the coarse spins, depend through Eq. (2.4) on the strengths of all the living bonds that connect the two blocks. In the $b<L$ scheme, however, the probabilities of occurrence of the same events depend (through the same equation) on the strength of a single bond whose value is the sum of living bonds connecting the two blocks. One might have thought that the resulting distribution of blocks at the coarsest level differs in the two schemes. Remembering, however, that p_d depends exponentially on the bond strength, the probability of deletion of a sequence of bonds one after the other is the same as that of deletion of a single bond, whose value is the sum of the bonds in the sequence. Therefore the probability of disconnecting the two coarse spins from one another is the same in the two procedures, and so is the probability of freezing, since the sum of the two probabilities is 1. Thus we have excluded the only possible difference between the schemes, and deduced that the SW method is equivalent to $\gamma=1$ cycles with no Metropolis sweeps.

III. PROOF OF DETAILED BALANCE

If every point in phase space can be reached in the course of a simulation, and transition probabilities satisfy detailed balance, the simulation procedure is legitimate; i.e., it takes the system to equilibrium and allows evaluation of thermal averages. Clearly, our procedure is strongly ergodic (even without Metropolis steps) since there is always a nonzero probability that no restriction will be placed on the simulation, allowing nonzero transition probabilities between all states. Transitions between coarse-spin states are generated either by Metropolis sweeps or by flipping disconnected blocks during uncoarsenings. Both processes satisfy detailed balance with respect to the coarse Hamiltonian. Consider such a transition, at coarse level n. It can be viewed as taking place between the two corresponding spin states of level $n-1$. We want to show that if the probability of generating $\mathcal{H}_n$ is also taken into account, detailed balance with respect to the Hamiltonian $\mathcal{H}_{n-1}$ is satisfied. Once proven, this implies (by induction) that the transitions satisfy detailed balance with respect to the *finest* Hamiltonian.

Start from some level in state Q with a Hamiltonian $\mathcal{H}$, and "kill" an interaction V to get the coarse Hamiltonian $\mathcal{H}_c=\mathcal{H}-V$. Then, a transition from Q to another state Q', with $V(Q)\neq V(Q')$ can take place only if V has been deleted, i.e., by (2.2):

$$W(Q\to Q')=C_V e^{V(Q)}W_c(Q\to Q')\ , \tag{3.1}$$

where W and W_c are transition probabilities that satisfy detailed balance with respect to $\mathcal{H}$ and $\mathcal{H}_c$, respectively. Hence,

$$\frac{W(Q\to Q')}{W(Q'\to Q)}=\frac{e^{V(Q)}e^{-\mathcal{H}_c(Q')}}{e^{V(Q')}e^{-\mathcal{H}_c(Q)}}=\frac{e^{-\mathcal{H}(Q')}}{e^{-\mathcal{H}(Q)}}\ , \tag{3.2}$$

where the fact that the rates W_c satisfy detailed balance with respect to $\mathcal{H}_c$ has been used. Alternatively, if Q and Q' are such that $V(Q)=V(Q')$, the interaction V may have either been deleted or frozen:

$$\begin{aligned}W(Q\to Q')&=C_V e^{V(Q)}W_c(Q\to Q')\\&\quad+(1-C_V e^{V(Q)})W_c^f(Q\to Q')\\&=[C_V e^{V(Q)}W_c(Q'\to Q)\\&\quad+(1-C_V e^{V(Q)})W_c^f(Q'\to Q)]\frac{e^{-\mathcal{H}_c(Q')}}{e^{-\mathcal{H}_c(Q)}}\ ,\end{aligned} \tag{3.3}$$

where W_c^f are transition probabilities with the reduced Hamiltonian over the restricted space. We have used here the detailed balance condition:

$$\frac{W_c(Q\to Q')}{W_c(Q'\to Q)}=\frac{W_c^f(Q\to Q')}{W_c^f(Q'\to Q)}=\frac{e^{-\mathcal{H}_c(Q')}}{e^{-\mathcal{H}_c(Q)}}\ . \tag{3.4}$$

Remembering that $V(Q')=V(Q)$ yields

$$W(Q\to Q')=W(Q'\to Q)\frac{e^{-\mathcal{H}(Q')}}{e^{-\mathcal{H}(Q)}}\ , \tag{3.5}$$

completing the proof of detailed balance.

One final remark is, though, in order. Going back to our actual procedure, as described in Sec. II, one may notice that when we visit a bond, we not only decide whether to delete or freeze it. Prior to that, we check whether it connects two (effective) coarse spins; if it does, we leave the bond *alive*. One may wonder whether this decision on the fate of a bond, which depends on previous decisions, made about other bonds, does or does not destroy detailed balance. Obviously, detailed balance is "threatened" only by decisions that depend on the spin configuration Q. The decision to leave a bond alive depends on Q *indirectly*, through the fact that whether the other relevant bonds were or were not frozen, did depend on Q. However, this dependence arises only through previous freezing and/or deleting decisions, for which our proof of detailed balance definitely holds.

IV. DETAILS OF SIMULATIONS

In this section details and results of simulations of the $d=2$ Ising and three-state Potts models on the square lattice are given. First, we describe tests of the computer program. We compare equilibrium values of energy, susceptibility, and specific heat of the Ising model (at various temperatures) obtained using the new algorithm, with analytic results,[20] values from simulations done using the Metropolis algorithm, and scaling laws at the critical point.[21]

In the second part we show results of relaxation of energy and susceptibility of the Ising model from a fully magnetized state to equilibrium, and compare them to standard (Metropolis algorithm) and SW results. We analyze in detail the influence of two parameters (the multigrid cycle parameter γ and the number of Metropolis sweeps at intermediate levels, μ) on the results. In the third part relaxation of time-delayed energy-energy correlations is examined. Lastly, we report results of simulations performed on the three-state Potts model. In particular, we show the improvements of our algorithm on the SW method by measuring relaxation times from energy decay.

A. Equilibrium properties

In order to be sure that our computer program works properly, we performed a number of simulations. We measured the equilibrium values of the energy, specific heat, and susceptibility, for lattices of linear sizes $L=4$, 8, 16, 32, and 64 with periodic boundary conditions, at various temperatures ($T/T_c=0.6, 0.8, 1.0, 1.3, 1.5$; T_c is the critical temperature of the infinite lattice). Multigrid cycles characterized by $\gamma=1$, 2, and 4, with rescaling factor $b=2$ were used to perform these measurements. Each run consisted of 30 equilibration cycles followed by a large number of measurement cycles (from 4000 cycles for $L=64$ to 10^6 cycles for $L=4$). Averages were calculated using a single data point from each cycle, measured at the finest level. Standard deviations were measured as well. The deviation of a single measurement of the energy E, for example, from the average $\langle E\rangle$, was estimated as $\sigma_E=(\langle E^2\rangle-\langle E\rangle^2)^{1/2}$. As shown below, for our multigrid algorithm a typical relaxation time of E is of the order of two cycles. Hence, when $\langle E\rangle$ is estimated as an average over N measurements, the standard deviation is well approximated by $\sigma_E/\sqrt{N/2}$. We measured energy with accuracy of 0.1%, while the accuracy of susceptibility measurements ranged from 0.1% for $L=4$ to 2% for $L=64$. Specific heat is much more difficult to measure with high precision; in our measurements the error ranged from 2% for a 4×4 lattice to 30% for a 64×64 lattice.

After taking averages over the measurement cycles we compared our results to known ones. First, we calculated the equilibrium values of the energy and specific heat from the exact expressions obtained by Ferdinand and Fisher.[20] The upper part of Fig. 3 shows the deviation of our results from the exact energy values for various lattice sizes, evaluated at the critical temperature of the infinite lattice. Such comparisons were done for all temperatures and for the specific heat as well, and the agreement with the analytical results is satisfactory. Secondly, we compared our results for the susceptibility with values obtained from simulations with the Metropolis algorithm. Lastly, we used the scaling of the specific heat C and the susceptibility χ with the linear size of the system, L, at the critical temperature:[21]

$$C\sim\log_{10}(L);\quad \chi\sim L^{1.75}\ .$$

In the lower part of Fig. 3 our results for the specific heat are shown to behave linearly with $\log_{10}(L)$, while the middle part of the Figure is a log-log plot of our values for $\chi(L)$, with the solid line marking the slope 1.75. The boxed data points denote results from simulations with

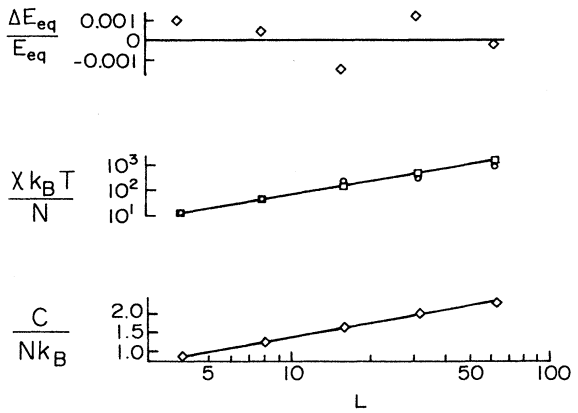

FIG. 3. Tests of the program: The upper graph shows ΔE—the deviation of equilibrium energy values, measured with the new algorithm, from exact values (Ref. 20)—plotted vs linear system size L (on logarithmic scale). The deviations are of the order of 0.1%. The middle graph is a log-log plot of the dimensionless susceptibility per spin, $\chi k_B T/N$, vs linear system size, with the solid line marking a slope of 1.75. Squares denote results from simulations using the multigrid algorithm, while results from Metropolis-type simulations are denoted by circles. In the bottom graph the specific heat per site, $C/(Nk_B)$, is plotted vs L (on logarithmic scale).

the multigrid algorithm, while results from Metropolis simulations are marked by circles. These comparisons clearly show that our algorithm yields the correct equilibrium values for the two-dimensional Ising model.

B. Relaxation to equilibrium — Ising dynamic exponent

We used the multigrid Monte Carlo (MGMC) algorithm to simulate lattices of linear sizes $L=4$, 8, 16, 32, 64, and 128 with periodic boundary conditions. Cycles of $\gamma=1$, 2, and 4 with rescaling factor $b=2$ were used. To check the influence of Monte Carlo sweeps at different scales we performed runs with $\mu=1$, 2, or 4 Metropolis sweeps at each level. Starting from a fully magnetized state we measured the energy and the square of the magnetization as a function of time. After repeating this measurement many times (up to 3×10^5 systems for $L=4$ and 300 for $L=128$), we averaged over the ensemble to get time-dependent averages, which in the long-time limit decay exponentially to an equilibrium value. The logarithm of such an average (after subtracting the appropriate equilibrium value) should decay linearly with time, with slope $1/\tau$, where τ is the relaxation time. Figs. 4(a), 5(a), and 6(a) contain log plots of the deviation of the energy from its equilibrium value versus time for $L=8$, 32, and 128, respectively. Error bars denote one standard deviation with respect to the ensemble at any given time. The multigrid cycle (MGC) served as our basic time unit. Here a cycle of $\gamma=2$ with a single Metropolis sweep at each level was used. A measurement was taken once at each level, arrived at after coarsening or uncoarsening (measurement algorithms for the energy and susceptibility at coarse levels are described in the Appendix). Note that our data points fall into clusters; measurements within the same cycle seem to be highly correlated. Thus one cannot gain efficiency by performing many measurements in a cycle. This fact tells us that the natural time unit in the simulations is the multigrid cycle.

In order to eliminate the influence of fine details of the multigrid cycle, we also present values of the energy averaged over each cycle. Thus each cycle is represented by a single point in the plot. These cycle averages do not reduce fluctuations in a significant manner, since measurements in the same cycle are strongly correlated. The cycle averages that correspond to Figs. 4(a), 5(a), and 6(a) are shown in Figs. 4(b), 5(b), and 6(b), respectively. From the slopes of such plots we extracted the equilibration times of the energy, $\tau_E(L)$, and of the square of the magnetization, $\tau_\chi(L)$. Similar behavior was found for the square of the magnetization, although with different relaxation times. The exact (finite size) equilibrium values of the energy were calculated from the analytic formulas of Ref. 20, while equilibrium values of the susceptibility were taken from our simulations, described in Sec. IV A.

One should note that here, in contrast with standard simulations, the transients do not get longer as the size of the lattice is increased. This indicates that the nonlinear dynamic exponent ($z_{\rm nl}$) vanishes; the Racz scaling law,[22] derived for local dynamics, $z_{\rm nl}=z_l-\beta$ (β is the order parameter exponent), does not apply for MGMC. In standard simulations the time beyond which the system is linear (i.e., relaxes with a simple exponential time dependence) also diverges at criticality. Our method, as we demonstrate below, has no divergent time scale at all; hence the uncommon fast termination of the nonlinear regime.

In all our measurements the time unit was taken to be one multigrid cycle, parametrized by γ and b. In order to calculate the amount of work (or computer time) per cycle, $W(\gamma,b)$, we assume that the amount of work performed at each level is proportional to the number of spins at that level. Calculation of W involves summing a geometric series, where each term is the number of times a level is coarsened in a cycle weighted by the number of spins at that level:

$$W(\gamma,b)\sim L^d\sum_{i=0}^{m-2}\left[\frac{\gamma}{b^d}\right]^i,$$

where $m=\log_b L+1$ is the number of levels. For large systems the asymptotic behavior of this quantity is given by

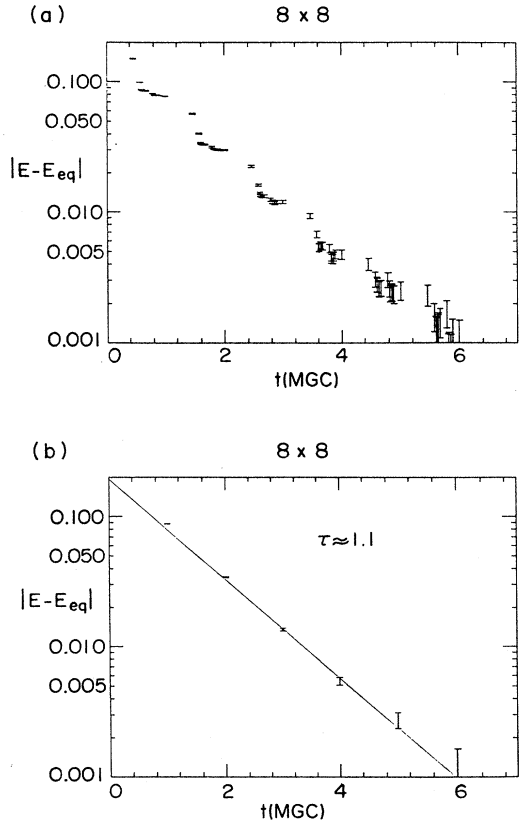

FIG. 4. (a) Log plot of $|E-E_{\rm eq}|$—the deviation of the energy from its equilibrium value—vs time t [in units of multigrid cycles (MGC)], for a lattice of linear size $L=8$. Measurements are taken at each level, arrived by coarsening or uncoarsening, and error bars denote one standard deviation with respect to the ensemble. (b) Cycle averages of (a). The decay is exponential with a slope consistent with $\tau\approx1.1$ MGC.

$$\frac{W(\gamma,b)}{L^d} \sim \begin{cases} 1 & \text{for } \gamma < b^d \\ \log_b L & \text{for } \gamma = b^d \\ L^{\log_b \gamma - d} & \text{for } \gamma > b^d \,. \end{cases} \tag{4.1}$$

In the simulations presented here we used cycles with $b=2$. Hence, for large L the work per cycle is proportional to the number of spins provided $\gamma < b^d = 4$, and diverges with L for $\gamma \geq 4$. Whenever we make detailed comparisons between different cycles, and between our algorithm and other methods, the computer time per cycle should be taken into account.

In Figs. 7, 8, and 9 the equilibration times are plotted versus L for the three types of multigrid cycles we used. As expected the $\gamma=1$ results (Fig. 7) agree with those of SW ($z \approx 0.35$). In fact, when no Metropolis sweeps are performed, not only the scaling, but also the actual values of the relaxation times agree. For $\gamma > 1, z$ is much smaller; the manner in which $\tau(L)$ flattens as L increases clearly indicates a tendency for asymptotic saturation, i.e., $z=0$. Figures 8 and 9 from our finite size simulations yield upper bounds: $z_E(\gamma=2) < 0.02$ $z_\chi(\gamma=2) < 0.04$, $z_E(\gamma=4) < 0.09$, and $z_\chi(\gamma=4) < 0.2$. Better bounds can be obtained by going to larger lattices and by performing longer simulations.

We did not detect any influence of the number of Metropolis sweeps at intermediate levels on the dynamic exponent. However, such sweeps reduce the relaxation times by a constant factor, with a relatively low cost in computer time. For example, the energy relaxation time of the 128×128 lattice, in the cycle characterized by $\gamma=2$ and $b=2$, is reduced from 2.8 cycles when no Metropolis sweeps are performed to 1.4 cycles, as obtained when four sweeps are performed at each level. As can be seen from Figs. 8 and 9, the saturation of the relaxation time occurs at smaller systems when more intermediate Metropolis sweeps are performed.

Figure 10 is a comparison of the dependence of relaxation times on system size, between the three simulation techniques: Metropolis,[6,13] the SW procedure,[6] and the $\gamma=2$, $\mu=1$ cycle of our algorithm (MGMC). Taking into account computer time overheads, the crossing point between MGMC and the SW procedures is at $L \approx 100$. In more complicated models we expect the dynamic exponent of the SW procedure to be higher and the crossing to occur at smaller lattice sizes.

Needless to say, our algorithm uses more computer storage than that needed for the SW method. We cannot give a definite quantitative statement about this storage overhead, since it depends crucially on the specific com-

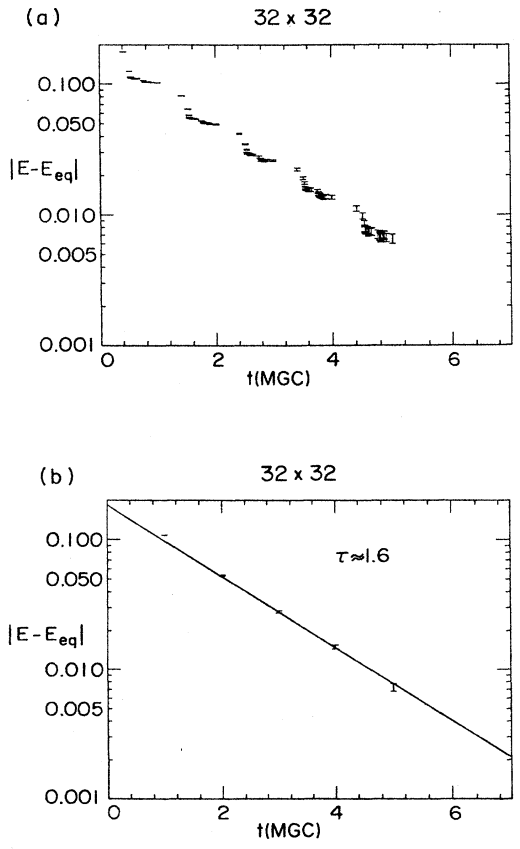

FIG. 5. (a) The same as Fig. 4(a), for a lattice of linear size $L=32$. (b) Cycle averages of (a). $|E-E_{eq}|$ decays exponentially with time, with a relaxation time $\tau \approx 1.6$ MGC.

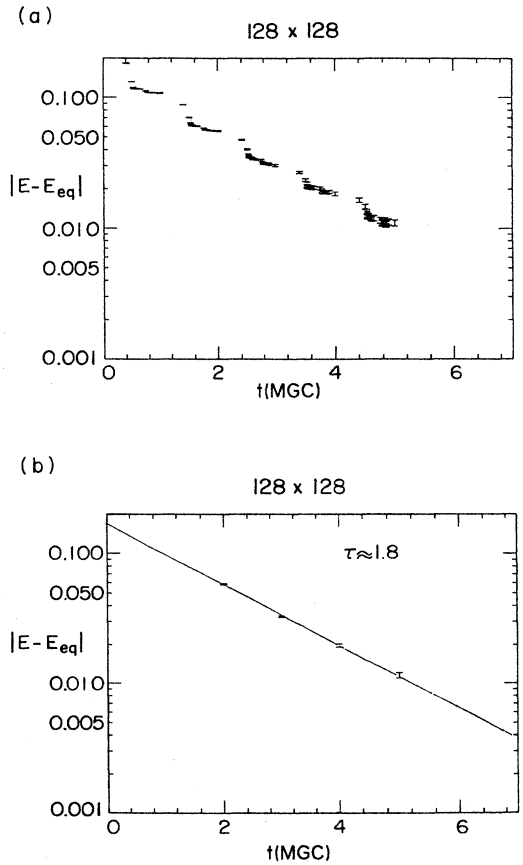

FIG. 6. (a) The same as Fig. 4(a), for a lattice of linear size $L=128$. (b) Cycle averages of (a). $|E-E_{eq}|$ decays exponentially with time, with a relaxation time $\tau \approx 1.8$ MGC.

puter program one is using. We can, however, state that even when no attempt is made to reduce the amount of storage used, the overhead factor is less than 10.

C. Time-delayed energy-energy correlations in equilibrium

In Sec. II B we estimated relaxation times from the decay with time of various quantities starting from a fully magnetized state. There are, however, other ways that are commonly used to extract relaxation times. For example, one can measure the time dependence of time-delayed energy-energy correlations *at equilibrium* (see Ref. 6). The method we have used in Sec. II B is easier to implement since energy-energy correlations fluctuate more strongly than the energy. One should, however, be careful since in such a nonequilibrium measurement the system may not encounter typical equilibrium states. To ensure that the complete elimination of CSD, implied by our results, is not an artifact of the particular method of measurement we have used, we performed equilibrium measurements of time-delayed energy-energy correlations.

The simulations were done on systems of linear sizes $4 \le L \le 128$, using three different multigrid cycles characterized by the parameters $\gamma=1$, $\mu=0$; $\gamma=2$, $\mu=1$; and $\gamma=2$, $\mu=4$. The first cycle was expected to be consistent with the SW results, while the remaining two were used to detect the elimination of CSD. Each run consisted of 80 000 cycles; the first 200 were discarded, and the remaining 79 800 were used to calculate averages. Once

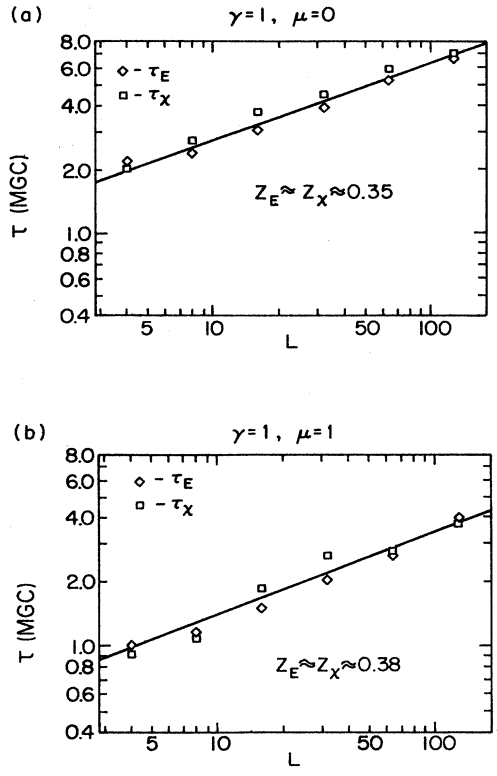

FIG. 7. Log-log plot of the energy and susceptibility relaxation times, τ_E and τ_χ, vs linear system size for the Ising model. Cycles with $\gamma=1$ were used. (a) $\gamma=1$ cycle without intermediate Metropolis sweeps ($\mu=0$); this is equivalent to SW (see text). Both relaxation times grow exponentially with exponents $z_E \approx z_\chi \approx 0.35$. (b) $\gamma=1$ cycle with a single intermediate sweep at each level ($\mu=1$). Both relaxation times grow exponentially with exponents $z_E \approx z_\chi \approx 0.38$.

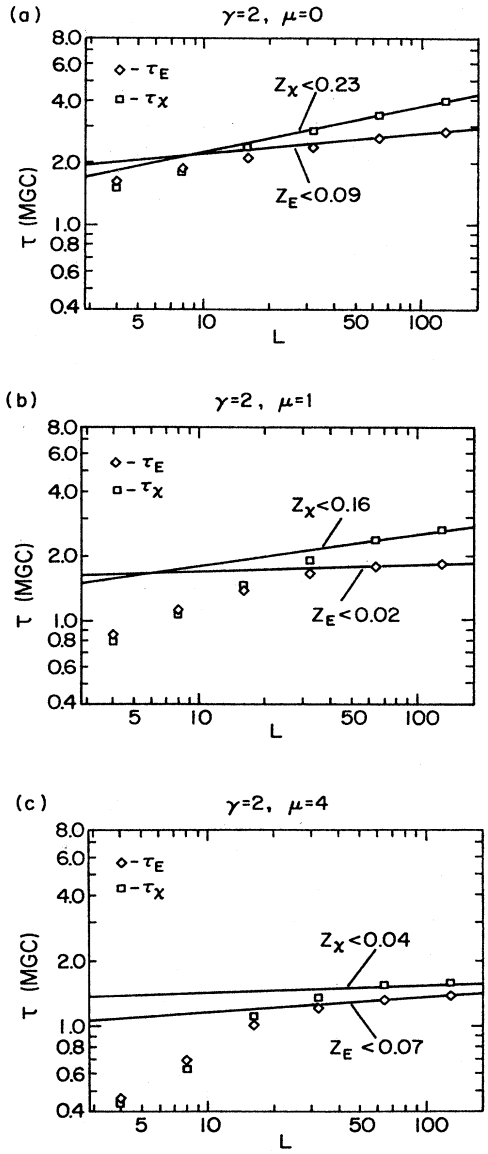

FIG. 8. Log-log plot of the energy and susceptibility relaxation times, τ_E and τ_χ, vs linear system size for the Ising model. Three types of cycles, all with $\gamma=2$, were simulated. (a) $\gamma=2$ cycle without intermediate Metropolis sweeps ($\mu=0$). Upper bounds for the dynamic exponents are found to be $z_E<0.09$ and $z_\chi<0.23$. (b) $\gamma=2$ cycle with a single intermediate sweep at each level ($\mu=1$). These results are consistent with the following upper bounds for the dynamic exponents: $z_E<0.02$ and $z_\chi<0.16$. (c) $\gamma=2$ cycle with four intermediate sweeps at each level ($\mu=4$). Upper bounds for the dynamic exponents are found to be $z_E<0.07$ and $z_\chi<0.04$.

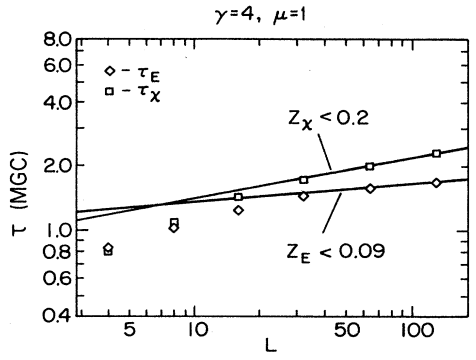

FIG. 9. Log-log plot of the energy and susceptibility relaxation times, τ_E and τ_χ, vs linear system size for the Ising model. A cycle with $\gamma=4$ and $\mu=1$ was used. The results yield upper bounds for the dynamic exponents: $z_E<0.09$ and $z_\chi<0.2$.

in each cycle we measured the average energy, and from the sequence of energy values we calculated the time-delayed energy-energy correlations $\langle E(0)E(t)\rangle$; here $\langle\ \rangle$ denotes time averaging:

$$\langle E(0)E(t)\rangle=\frac{1}{T-t}\sum_{t'=1}^{T-t}E(t')E(t'+t)\ , \tag{4.2}$$

where T is the length of the sequence of energy values obtained from the simulation. We extracted the relaxation times from the time dependence of the normalized correlations:

$$\Delta(t)=\frac{\langle E(0)E(t)\rangle-\langle E\rangle^2}{\langle E^2\rangle-\langle E\rangle^2}\ , \tag{4.3}$$

where exact values[20] of the average energy $\langle E\rangle$ were used, and values of $\langle E^2\rangle$ were taken from our simulation.

Examples of $\Delta(t)$ for various lattice sizes are shown in Fig. 11. The results are taken from a simulation of a cycle with $\gamma=2$ and $\mu=4$. $\Delta(t)$ decays exponentially with time, and relaxation times can be estimated from this decay as was demonstrated in Sec. II B. The relaxation times obtained from the $\gamma=1$, $\mu=0$ cycle are in perfect agreement with Swendsen and Wang's results. Results from cycles with $\gamma=2$ are shown in Fig. 12; relaxation times are plotted versus linear system size. We see that the behavior of $\tau(L)$ as a function of L is similar to what

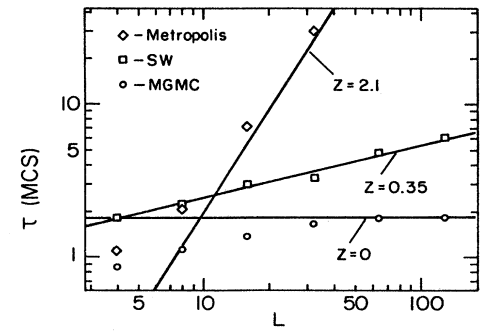

FIG. 10. Relaxation time τ vs linear system size L. Three methods are compared: Metropolis algorithm (Ref. 13), Swendsen and Wang's method (Ref. 6) (SW), and the $\gamma=2$, $\mu=1$ cycle of the multigrid Monte Carlo technique (MGMC). Time is measured in Monte Carlo steps (MCS). A MCS is equivalent to a single sweep over the lattice in the Metropolis algorithm, and to a single cycle in the SW and MGMC methods.

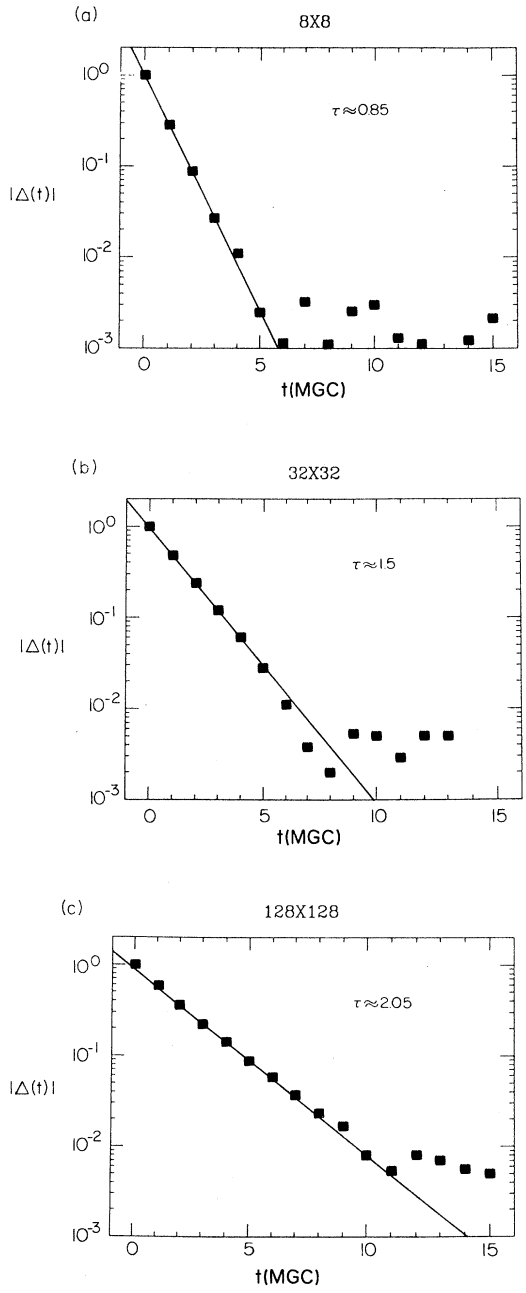

FIG. 11. Log plots of normalized time-delayed energy-energy correlations $\Delta(t)$ [see Eq. (4.3) for exact definition] vs time [in units of multigrid cycles (MGC)] for various system sizes: (a) $L=8$, (b) $L=32$, and (c) $L=128$. Relaxation times (estimated from the exponential decay) are given on each plot.

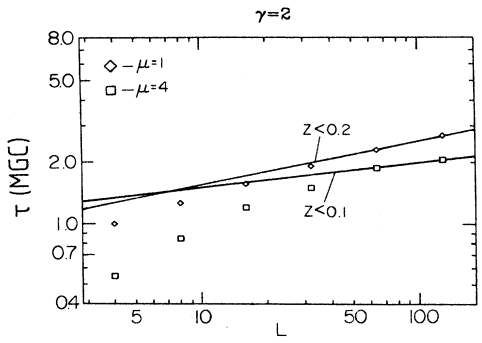

FIG. 12. Log-log plots of relaxation times (in units of MGC) obtained from $\Delta(t)$ (see Fig. 11) vs linear system size L, for two types of multigrid cycles: $\gamma=2$, $\mu=1$ (diamonds) and $\gamma=2$, $\mu=4$ (squares). The upper bounds on the dynamic exponent are $z<0.2$ from the first cycle, and $z<0.1$ from the second one.

was observed in nonequilibrium measurements (see Fig. 8); $\tau(L)$ tends to saturate with increasing L, indicating that $z=0$. Upper bounds on the dynamic exponent, determined from our data, are $z<0.2$ from the $\mu=1$ cycle, and $z<0.1$ from the $\mu=4$ cycle. Thus the time dependence of energy-energy correlations confirms our previous statement, that CSD is eliminated completely.

D. Three-state Potts model

To see that the MGMC method eliminates CSD for the three-state Potts model, we measured energy relaxation times using a cycle of $\gamma=2$ and $\mu=2$. Relaxation times are plotted versus lattice linear size in Fig. 13, on a log-log scale. From this plot we can deduce an upper bound for the dynamic critical exponent, $z<0.2$, while SW reported[6] $z_{SW}\approx 0.6$ for this model. Since finite size effects are more pronounced here than in the Ising model, we have to simulate much larger lattices in order to get a better upper bound for z. Because of these finite size effects we suspect that the SW result for the three-state Potts model is inaccurate; as is explained in Sec. VI our simulations are more consistent with $z_{SW}\approx 0.4$ in this case.

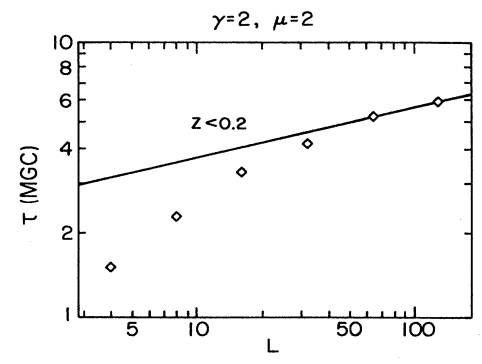

FIG. 13. Log-log plot of the energy relaxation time τ_E vs linear system size for the three-state Potts model. A cycle with $\gamma=2$ and $\mu=2$ was used. These results are consistent with $z<0.2$.

V. WHY DOES IT WORK?

In order to understand why CSD is eliminated by the MGMC procedure, we first review briefly the physical origin of slowing down, and explain why the SW algorithm does not eliminate it.

At criticality, configurations with correlated clusters of spins on all length scales are probable, and hence must be reached by the simulation in order to produce correct equilibrium measurements. The simulation must be able to induce transitions between configurations with very different clusters on arbitrarily large scales. To do this, large clusters have to be broken up, to be replaced by different ones. That is, the process has to destroy correlations on all length scales, present in any typical configuration. CSD is a reflection of the fact that elimination of long-range correlations takes a long time for a process that uses single-spin-flip dynamics. Both the MGMC and SW algorithms attempt to overcome this difficulty by allowing flips of large domains. This is done by the coarsening transformation, which identifies a possibly large block of fine spins with a single coarse spin. Long-range correlations on the fine lattice are transformed into short-range ones, and hence are easily eliminated by block-flip dynamics.

Had this been the entire picture, the SW algorithm should have eliminated CSD completely. However, coarsening introduces into the system correlations that do not allow the elimination of CSD. If two correlated fine spins i and j have been absorbed into the same block of a certain coarse level, their correlation cannot be eliminated. This happens when bonds along a path that connects these spins are frozen during the coarsening process. The correlation between such spins can decay only after a finer level, in which spins i and j are in separate blocks, is reached. This is a manifestation of the restrictions on the phase space caused by the coarsening transformation. Due to these restrictions it is impossible to get independent fine-lattice configurations by simulating a single realization of a coarse level. To be more specific, let us start from a given fine-lattice configuration, and use the coarsening procedure to obtain a new lattice with a new Hamiltonian. Even if we simulate the coarse lattice for a very long time, we cannot get a coarse-level configuration whose fine-lattice representation is independent of the initial fine-lattice configuration. Hence, a long enough sequence of coarse-level realizations (generated by a sequence of coarsening-uncoarsening steps) is needed in order to decorrelate the system. Let us denote the "decorrelation time" (i.e., the number of coarsening-uncoarsening steps needed in such a sequence) by $\tau_d(b)$. When the sequence of coarse-level systems exceeds $\tau_d(b)$, we say that the coarse level represents the original system faithfully.

To understand the dependence of τ_d on b, consider a fine-spin cluster, roughly of size $b\times b$, in which all the spins point in the same direction. In most coarsening at-

tempts almost all spins in this cluster will be frozen to the same coarse spin, since typical frozen blocks of spins are of size b^2. $\tau_d(b)$ is of the order of the number of coarsening-uncoarsening steps needed in order to break the fine cluster into two (approximately equal) parts, each frozen to a different coarse spin. With this interpretation, it is clear that $\tau_d(b)$ increases with increasing length rescaling factor b. At this point we note that the SW procedure can be viewed as a single coarsening step, *with* $b=L$ (see Sec. II B). Also note that the SW procedure is completely local; the decision whether to freeze or delete a bond does not depend on L, and the Hamiltonian obtained after coarsening is completely trivial, and hence its relaxation time is independent of L. Nevertheless, SW do find a divergent relaxation time $\tau \sim L^{z_{\rm SW}}$. Since the only time scale that can possibly diverge in the SW algorithm is the decorrelation time, we identify $\tau_d(b=L)$ with the relaxation time τ. Thus we propose to interpret the growth of τ with L as arising from the need for *more* coarsening-uncoarsening steps with increasing b. From this identification, and the SW results, we infer that the decorrelation time $\tau_d(b)$ grows with b as

$$\tau_d(b) \sim b^{z_{\rm SW}} . \tag{5.1}$$

In the remainder of this section we assume validity of this relation, and present its substantiation, as well as various interesting consequences of its validity in Sec. VI.

We now turn to explain how the MGMC method does overcome the problem of CSD. Whereas in the SW algorithm we have $b=L$, the MGMC method allows for increasing L without changing b, simply by increasing the number of multigrid levels m. L is related to b and m by $L=b^{m-1}$. It is important to understand that in order for the coarsest level to represent the original system faithfully, each intermediate level should also be a faithful representation of the finest lattice. This arises from the fact that any phase-space restrictions present at level n, persist to the $(n+1)$st level. Now level n can be viewed as obtained from the fine lattice by coarsening with length rescaling factor b^{n-1}, and hence it will faithfully represent the fine lattice only if it has been reached $\tau_d(b^{n-1}) \sim (b^{z_{\rm SW}})^{n-1}$ times [where Eq. (5.1) for the decorrelation time was used]. That is, the coarser the level and higher the value of n, the more visits to the level are needed. This observation explains the manner in which MGMC overcomes CSD, whereas the SW method does not.

As explained in detail in Sec. II B, the SW procedure can be viewed either as a single coarsening step with rescaling factor $b=L$, or as a sequence of coarsening-uncoarsening steps. This sequence constitutes a $\gamma=1$ cycle with an arbitrary length rescaling factor, and no intermediate Metropolis sweeps. Therefore SW effectively visit every level the same number of times during the simulation. The number of these visits equals the required number of times the *coarsest* one has to be reached. We know, however, that for decorrelation fine levels need not be visited as often as coarse ones. Hence, one could save time and work by visiting levels the more often the coarser they are, and precisely this is accomplished by the multigrid method. In other words, since in a cycle with $\gamma>1$ coarse levels are reached more often than fine ones (see Fig. 2), we can improve on the SW algorithm by increasing γ.

To see how this comes about, note that with the definition of γ, the number of times the nth level is created by coarsening during a cycle is $\mathcal{N}_\gamma(n)=\gamma^{n-1}$. Hence, the number of cycles needed in order to decorrelate the system is

$$N_{\rm cyc} \sim \max_n \frac{\tau_d(b^{n-1})}{\mathcal{N}_\gamma(n)} \approx \max_n (b^{z_{\rm SW}-\log_b\gamma})^{n-1} , \tag{5.2}$$

which implies that $N_{\rm cyc}$ diverges as $N_{\rm cyc} \sim L^{z_{\rm SW}-\log_b\gamma}$ if $z_{\rm SW} > \log_b\gamma$. On the other hand, for $z_{\rm SW} \le \log_b\gamma$ we get a finite $N_{\rm cyc}$. Taking into account the amount of work performed per cycle $W(\gamma,b)$ [see Eq. (4.1)], we have $\tau = W(\gamma,b) \times N_{\rm cyc}(\gamma,b)$, which yields, for $z_{\rm SW} \le d$,

$$\tau \sim \begin{cases} L^{z_{\rm SW}-\log_b\gamma} & \text{for } \log_b\gamma < z_{\rm SW} \\ 1 \quad \text{for } z_{\rm SW} \le \log_b\gamma < d \\ \log_b L \quad \text{for } \log_b\gamma = d \\ L^{\log_b\gamma - d} \quad \text{for } \log_b\gamma > d . \end{cases} \tag{5.3}$$

Hence we see that CSD is eliminated completely if

$$b^{z_{\rm SW}} \le \gamma < b^d . \tag{5.4}$$

For the Ising model, this condition translates to $b^{0.35} \le \gamma < b^2$ for $d=2$, and $b^{0.75} \le \gamma < b^3$ for $d=3$. For the three-state Potts model in two dimensions we show below that $z_{\rm SW} \approx 0.4$; hence γ and b must satisfy $b^{0.4} \le \gamma < b^2$. For values of b and γ that do not satisfy these bounds, Eq. (5.3) predicts CSD. In particular, for the three-dimensional Ising model, with $\gamma=2$ and $b=4$, we expect to get $z \approx 0.25$.

If, however, a model is found for which $z_{\rm SW} > d$, Eq. (5.4) is replaced by

$$\tau \sim \begin{cases} L^{z_{\rm SW}-\log_b\gamma} & \text{for } \log_b\gamma < d \\ L^{z_{\rm SW}-d}\log_b L \quad \text{for } \log_b\gamma = d \\ L^{z_{\rm SW}-d} \quad \text{for } d < \log_b\gamma < z_{\rm SW} \\ L^{\log_b\gamma - d} \quad \text{for } \log_b\gamma \ge z_{\rm SW} . \end{cases} \tag{5.5}$$

For a model with such high value of $z_{\rm SW}$ the MGMC method will reduce CSD significantly, but will not eliminate it completely.

VI. TESTS OF THE ARGUMENT

We now present evidence for the validity of Eq. (5.1),

$$\tau_d(b) \sim b^{z_{\rm SW}} \tag{6.1}$$

Surprisingly, we found that this expression for the decorrelation time has interesting consequences, which we now also present.

The main mechanism responsible for decorrelation in

the SW and MGMC algorithms is the deletion of bonds. We can think of the coarsening transformation as a series of attempts to delete bonds. A reasonable ansatz is that the decorrelation time is proportional to the average number of attempts needed to delete (rather than freeze) a bond. Let us denote by $\bar{P}_d(n)$ the average probability that a coarse bond of the nth level, which is killed during the coarsening of that level, is deleted. As long as b is small enough, we can consider attempts to delete bonds during coarsening as independent events. We may then approximate the average number of attempts needed to delete a given bond by $1/\bar{P}_d(n-1)$. According to our ansatz this implies that

$$1/\bar{P}_d(n) \sim \tau_d(b^n) ,$$

which together with Eq. (5.1) leads to

$$\bar{P}_d(n) \sim (b^{z_{SW}})^{-n} . \tag{6.2}$$

This is an important result. It means that rather than measuring the relaxation time, we can calculate z_{SW} from measurements of deletion probabilities at various levels.

To obtain deletion probabilities we first measured bond-strength distributions. The statistics were obtained from simulations of a 128×128 lattice, using a cycle of $\gamma=2$ and $b=2$. Starting from a fully magnetized state we first performed 30 equilibration cycles and then started the measurement. The number of bonds of any given strength was counted separately for each level (each time it was reached during the simulation). We also separated satisfied from broken bonds (i.e., counted the satisfied and broken bonds of the same strength separately). The average normalized distributions of levels 2–7 at criticality are plotted in Fig. 14. In the positive J side of each plot one finds the part of the satisfied bonds in the distribution, while the part of the broken bonds is plotted in the negative J side. Both broken and satisfied bond distributions seem to decay exponentially as the strength of the bond is increased. While the broken bond distribution is essentially independent of the level, the distribution of satisfied bonds flattens at coarser levels.

Remembering that the deletion probability of a broken bond is 1, and that of a satisfied bond of strength K is $\exp(-2K)$, we can calculate (numerically) the average probability to delete a bond at any level, assuming that the value of the bond is picked randomly from the distributions of Fig. 14. The result is shown in Fig. 15; the numerically determined average probability to delete a bond is plotted versus level number. This deletion probability is seen to decrease exponentially with the level number:

$$\bar{P}_d(n) \sim 1.27^{-n} . \tag{6.3}$$

This numerically found exponential decay is in perfect agreement with (6.2), which was derived from the expected form of $\tau_d(b)$ as given by Eq. (5.1). The agreement can be viewed as confirmation of the assumptions on which Eq. (5.1) was based, as well as the ansatz that related the decorrelation time to deletion probability. Moreover, using Eqs. (6.3) and (6.2), we get for the two-dimensional Ising model $2^{z_{SW}} \approx 1.27$, or

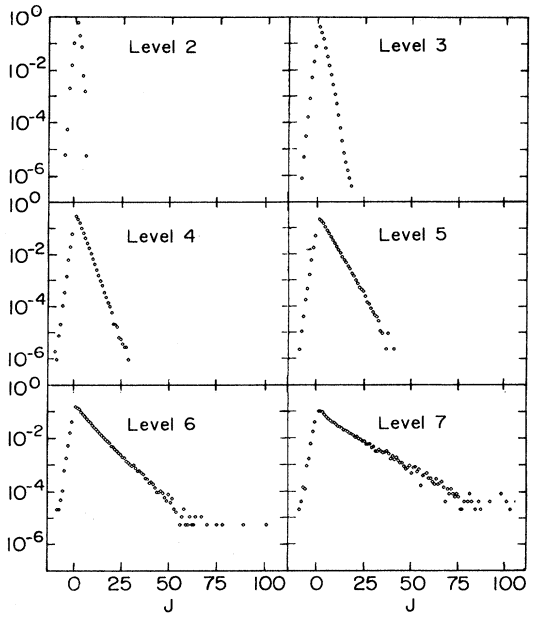

FIG. 14. Bond-strength distributions for levels 2–7 of a multigrid cycle $\gamma=2$ for a 128×128 Ising model. The measurement was performed at the critical point of the infinite lattice. J denotes bond strengths in units of k_BT. The left part of each graph ($J<0$) corresponds to unsatisfied bonds (see text for details). This graph was calculated from the numerically obtained data of Fig. 12; see text [above Eq. (6.3)] for details.

$$z_{SW} \approx 0.34 ,$$

in excellent agreement with the SW result.[6]

A further consistency check of our assumptions and results is provided by measuring the deletion probability for noncritical systems. We have established a direct connection of slowing down to the exponential dependence of the deletion probability on the level index. Since slowing down is observed only at criticality, we expect that once the temperature is changed from T_c, we should

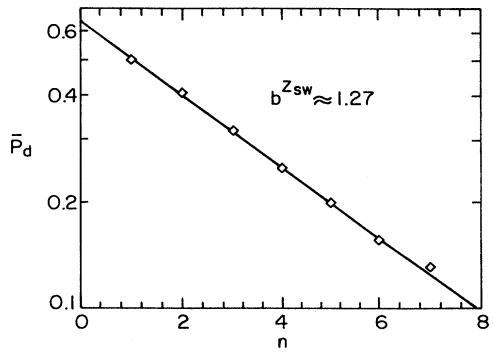

FIG. 15. Log plot of the bond deletion probability $\bar{P}_d$ vs level index n for the Ising model. The exponential decay is consistent with Eq. (6.2), and yields an estimate for the SW dynamic exponent $z_{SW} \approx 0.34$.

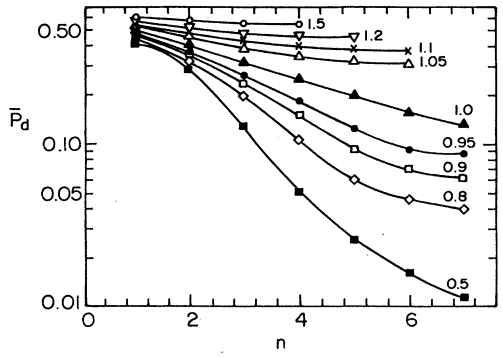

FIG. 16. Log plot of the deletion probabilities $\bar{P}_d$ vs level index n for the Ising model at various temperatures. Numerals near the solid lines denote the temperature in units of T_c. Exponential dependence occurs only at criticality. Note that deletion probabilities get higher as the temperature is raised [see Eq. (2.3)]. Therefore at coarse levels the spins are disconnected and all the bonds vanish. This is the reason for the disappearance of the last levels at high temperatures.

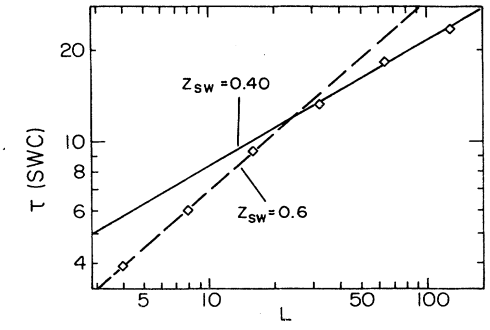

FIG. 18. Log-log plot of the energy relaxation time τ_E [in units of SW cycles (SWC)], vs linear system size for the three-state Potts model. These results were obtained from simulations with the original SW algorithm. The behavior at small L (dashed line) is consistent with the SW result ($z_{SW} \approx 0.6$), while the large L behavior (solid line) is more consistent with our prediction ($z_{SW} \approx 0.4$).

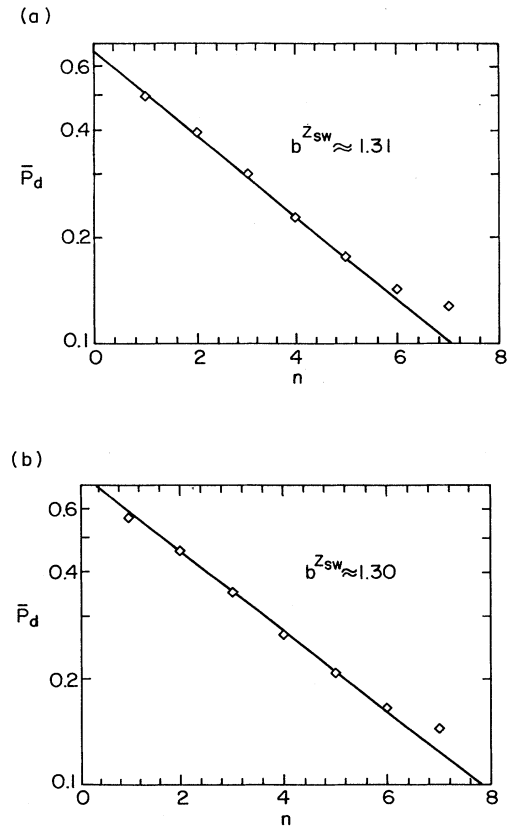

FIG. 17. (a) Log plot of the deletion probability $\bar{P}_d$ vs level index n for the three-state Potts model. Statistics were taken from simulations of a 128×128 lattice at the critical point of the infinite lattice. The exponential decay is consistent with Eq. (6.2), and yields an estimate for the SW dynamic exponent $z_{SW} \approx 0.39$. (b) The same as (a) for the four-state Potts model. The corresponding SW dynamic exponent is $z_{SW} \approx 0.38$.

no longer find an exponential function. Indeed, in Fig. 16 average deletion probabilities are plotted versus level number for a range of temperatures, and we see that only at the critical temperature is exponential decay seen.

To further check the validity of Eq. (6.2) we repeated the measurement of deletion probabilities for the three- and four-state Potts models in two dimensions. We used the same lattice size and the same type of multigrid cycle as in the Ising model, and obtained bond-strength distributions. From them we calculated the average deletion probabilities which are plotted as a function of level number in Figs. 17(a) and 17(b), for the three- and four-state Potts models, respectively. As in the Ising model, the deletion probabilities decay exponentially with the level number, except at very coarse levels, where finite size effects are important. For the three-state model we deduced that $\bar{P}_d(n) \sim 1.31^{-n}$, while for the four-state model $\bar{P}_d(n) \sim 1.30^{-n}$. Using Eq. (6.2) we predict the values of z_{SW} to be 0.39 and 0.38 for the three- and four-state models, respectively.

SW reported[6] $z_{SW} \approx 0.6$ for the three-state Potts model in contradiction with our prediction. To clarify the situation we performed simulations with the SW algorithm, and measured the decay of the energy with time, from a fully magnetized state. The relaxation times are plotted versus lattice size in Fig. 18. We see that for small lattices ($L < 32$) the SW result is indeed reproduced. A clear crossover is observed, however, and the slope becomes smaller for larger lattices. For the largest sizes simulated the slope is consistent with $z_{SW} \approx 0.40$, in agreement with our prediction.

VII. SUMMARY AND DISCUSSION

In this work we have discussed a new multigrid Monte Carlo simulation method, which is especially efficient near critical points. For Ising and Potts systems, it performs moves on all length scales, as produced by a stochastic coarsening process. Coarsening introduces, how-

ever, correlations, and our procedure is designed so that the amount of work performed at each length scale (level) is of the order of the amount of work needed to eliminate these correlations; hence, CSD is eliminated completely. We have demonstrated elimination of CSD for the Ising model on the square lattice, and obtained an extremely small upper bound for the dynamic exponent $z < 0.02$. Significant improvements were achieved for the three-state Potts model in two dimensions as well, but due to large finite size effects the upper bound for z is higher. We do believe that CSD is eliminated completely in this case too.

We have devised a scaling argument that explains how the MGMC technique eliminates CSD while the SW method does not. The main point is that in order to decorrelate the system one needs to create a sequence of lattices and Hamiltonians at the coarsest level. In the SW algorithm the average amount of work needed to create a coarse system in this sequence is proportional to the number of spins in the original (finest) lattice (L^d). The MGMC method succeeds in reducing the amount of work without need for a longer sequence. The average amount of work performed, in the MGMC procedure, while creating a coarse lattice in this sequence scales as $L^{d-z_{SW}}$ (if γ is chosen appropriately). Using bond-strength statistics on various levels, we were able to obtain predictions of z_{SW} (based on our scaling argument), which were used to test the validity of the argument. Our predictions agree very well with the SW results for the Ising model in two dimensions, but disagree in the case of the three-state Potts model. We attribute this disagreement to finite size effects that reduced the accuracy of the SW measurement.

In both SW and MGMC algorithms the lattice is transformed into a new lattice with fewer degrees of freedom. A spin of the new lattice represents a block of spins of the original problem. For the Ising and Potts models each of the blocks created in this way is a cluster of spins found in one of the ground states of the model. This property of the method may be essential for elimination of CSD. If the method is applied to models with frustration or competing interactions, the clusters are not necessarily in a ground state, which may reduce the efficiency of the simulation. Thus the MGMC method will not necessarily eliminate, without modifications, slowing down in spin glasses. We intend to explore the influence of frustration on the effectiveness of our method, aiming at a modified technique that will simulate spin glasses efficiently.

ACKNOWLEDGMENTS

We wish to thank E. Y. Loh for fruitful collaboration. We thank R. Benav for stimulating discussions and R. Swendsen for helpful discussions and correspondence. This research was supported by the Air Force Office of Scientific Research, United States Air Force under Grant No. AFOSR-86-0127 (A.B.) and by the US-Israel Binational Science Foundation (E.D.).

APPENDIX

In order to perform on coarse levels, measurements of quantities defined on the finest level, we carry with each block of spins its internal magnetization and energy on the finest levels. These values are not changed until we come back to the level where the block was created. We also carry with each bond of the coarse level its "finest value" (what would have been its value had no bonds been deleted). The energy of the finest lattice is evaluated at the coarse level by summing the "finest values" of all satisfied bonds, subtracting from the result the "finest values" of all broken bonds, and adding the sum of all internal energies of all blocks (including disconnected ones). Note that in this procedure fine bonds that connected disconnected blocks are disregarded. This is justified, since we are interested only in averages, and such bonds do not contribute to the average energy. The reason for this is that each disconnected block has the same probability of being up or down. Hence, each of the fine bonds that connected such a block to other blocks is satisfied or broken with equal probabilities, and its contribution to the average energy vanishes.

The square of the magnetization is evaluated by summing the internal magnetizations of all coarse spins, squaring the result, and adding to it the sum of squares of the internal magnetizations of all disconnected blocks. Here, as in the case of the average energy, we took into account the fact that disconnected blocks are up or down with equal probabilities.

[1]See *Applications of Monte Carlo Method in Statistical Physics*, 2nd ed., edited by K. Binder (Springer-Verlag, Berlin, 1987).
[2]I. Montvay, Rev. Mod. Phys. **59**, 263 (1987).
[3]See, for example, P. C. Hohenberg and B. I. Halperin, Rev. Mod. Phys. **49**, 435 (1977).
[4]J. Goodman and A. Sokal, Phys. Rev. Lett. **56**, 1015 (1986).
[5]E. Dagotto and J. Kogut, Phys. Rev. Lett. **58**, 299 (1987).
[6]R. H. Swendsen and J. S. Wang, Phys. Rev. Lett. **58**, 86 (1987).
[7]M. Sweeny, Phys. Rev. B **27**, 4445 (1983).
[8]W. L. Briggs, *A Multigrid Tutorial* (Society for Industrial and Applied Mathematics, Philadelphia, PA, 1987).
[9]A. Brandt, in *Proceedings of the Third Copper Mountain Conference on Multigrid Methods, Copper Mountain, Colorado, 1987*, edited by S. McCormick (Dekker, New York, 1988).
[10]A. Brandt, D. Ron, and D. J. Amit, in *Multigrid Methods II; Proceedings, Cologne*, Vol. 1228 of *Lecture Notes in Mathematics*, edited by W. Hackbusch (Springer-Verlag, Berlin, 1985), p. 66.
[11]P. W. Kasteleyn and C. M. Fortuin, J. Phys. Soc. Jpn. Suppl. **26s**, 11 (1969); C. M. Fortuin and P. W. Kasteleyn, Physica (Utrecht) **57**, 536 (1972).
[12]S. Miyashita and H. Takano, Prog. Theor. Phys. **73**, 1122 (1985).
[13]J. K. Williams, J. Phys. A **18**, 49 (1985); S. Tang and D. P. Landau, Phys. Rev. B **36**, 567 (1987).
[14]D. Kandel, E. Domany, D. Ron, A. Brandt, and E. Loh, Phys.

Rev. Lett. **60**, 1591 (1988).
[15]U. Wolff, Phys. Rev. Lett. **60**, 1461 (1988); F. Niedermayer *ibid.* **61**, 2026 (1988).
[16]N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[17]D. Kandel and E. Domany (unpublished).
[18]The order by which we visit the bonds during coarsening does not seem to have any effect on the results.
[19]It may happen that a block of fine spins that contains a boxed spin, and hence is represented by a spin on the coarse lattice, is disconnected from the rest of the coarse lattice. In principle, such a disconnected block may be left out of the coarse lattice and treated as all the other disconnected blocks. However, in order not to destroy the geometry of a square lattice (which will complicate programming), we consider such disconnected spins as part of the coarse lattice.
[20]A. E. Ferdinand and M. E. Fisher, Phys. Rev. **185**, 832 (1969).
[21]See, for example, M. N. Barber, in *Phase Transitions and Critical Phenomena,* edited by C. Domb and J. L. Lebowitz (Academic, London, 1983), Vol. 8, p. 145.
[22]Z. Racz, Phys. Rev. B **13**, 263 (1976).