Overcoming the slowing down of flat-histogram Monte Carlo simulations: Cluster updates and optimized broad-histogram ensembles

Yong Wu,[1] Mathias Körner,[2] Louis Colonna-Romano,[3] Simon Trebst,[2,4] Harvey Gould,[3] Jonathan Machta,[1] and Matthias Troyer[2]

[1]*Department of Physics, University of Massachusetts, Amherst, Massachusetts 01003-3720, USA*

[2]*Theoretische Physik, ETH Zürich, CH-8093 Zürich, Switzerland*

[3]*Department of Physics, Clark University, Worcester, Massachusetts 01610-1477, USA*

[4]*Computational Laboratory, Eidgenössische Technische Hochschule, Zürich, CH-8092 Zürich, Switzerland*

(Received 10 December 2004; published 4 October 2005)

We study the performance of Monte Carlo simulations that sample a broad histogram in energy by determining the mean first-passage time to span the entire energy space of d-dimensional ferromagnetic Ising/Potts models. We first show that flat-histogram Monte Carlo methods with single-spin flip updates such as the Wang-Landau algorithm or the multicanonical method perform suboptimally in comparison to an unbiased Markovian random walk in energy space. For the $d=1$, 2, 3 Ising model, the mean first-passage time τ scales with the number of spins $N=L^d$ as $\tau \propto N^2 L^z$. The exponent z is found to decrease as the dimensionality d is increased. In the mean-field limit of infinite dimensions we find that z vanishes up to logarithmic corrections. We then demonstrate how the slowdown characterized by $z>0$ for finite d can be overcome by two complementary approaches—cluster dynamics in connection with Wang-Landau sampling and the recently developed ensemble optimization technique. Both approaches are found to improve the random walk in energy space so that $\tau \propto N^2$ up to logarithmic corrections for the $d=1$, 2 Ising model.

DOI: 10.1103/PhysRevE.72.046704 PACS number(s): 02.70.Rr, 75.10.Hk, 64.60.Cn

I. INTRODUCTION

Recently, Wang and Landau [1,2] introduced a Monte Carlo (MC) algorithm that simulates a biased random walk in energy space, systematically estimates the density of states $g(E)$, and iteratively samples a flat histogram in energy. The bias depends on the total energy E of a configuration and is defined by a statistical ensemble with weights $w(E)=1/g(E)$. The idea is that the probability distribution of the energy $p(E)=w(E)g(E)$ will eventually become constant, producing a flat energy histogram. The algorithm has been successfully applied to a wide range of systems [3].

One measure of the performance of the Wang-Landau algorithm and other broad-histogram algorithms is the mean first-passage time τ, which we define to be the number of MC steps per spin for the system to go from the configuration of lowest energy to the configuration of highest energy. In earlier work, this time has been called the tunneling time [4–6] or the round-trip time [7]. The close connection of this time scale to sample the entire energy range to other time scales such as the autocorrelation times of measurements is discussed in Refs. [8–10]. Because the number of possible energy values in the Ising model scales linearly with the number of spins $N=L^d$, where d is the spatial dimension, we might expect from a simple random walk argument that $\tau \propto N^2$. However, it was recently shown that τ for flat-histogram algorithms scales as

$$\tau \sim N^2 L^z, \tag{1}$$

where the exponent z is a measure of the deviation of the random walk from unbiased Markovian behavior [6]. In analogy to the dynamical critical exponent for canonical simulations, the exponent z quantifies the slowing down for the simulated flat-histogram ensemble and the chosen local update dynamics [7]. For the Ising model on a square lattice it was shown in Ref. [6] that $z \approx 0.74$.

In this paper, we determine the performance of the single-spin flip Wang-Landau algorithm for the ferromagnetic Ising model as a function of the dimension d. We find that the exponent z is a decreasing function of d. By using Monte Carlo simulations, we determine that $z \approx 1.814$, 0.743, and 0.438 for $d=1$, 2, and 3, respectively (see Sec. II A). In the mean-field limit of infinite dimension we find that $z=0$ (up to logarithmic corrections) using Monte Carlo simulations and analytical methods (Sec. II B). To round out our examination of the single-spin flip algorithm, we also estimate z for the $q=10$ and $q=20$ Potts model in $d=2$ (see Sec. II C).

We then discuss two complementary approaches that overcome the slowing down of the single-spin flip flat-histogram algorithm. In Sec. III we demonstrate that flat-histogram MC simulations can be speeded up by *changing the spin dynamics*, that is, by using cluster updates instead of local updates. Alternatively, we apply the recently developed ensemble optimization technique [7] to *change the simulated statistical ensemble* and show that by sampling an optimized histogram instead of a flat histogram, the slowing down of the single-spin flip simulation also can be eliminated (Sec. IV). For the Ising model in one and two dimensions we find that both approaches result in optimal N^2 scaling of the random walk up to logarithmic corrections.

II. PERFORMANCE OF FLAT-HISTOGRAM ALGORITHMS WITH LOCAL UPDATES

In this section we study the performance of the single-spin flip flat-histogram algorithm by measuring the mean

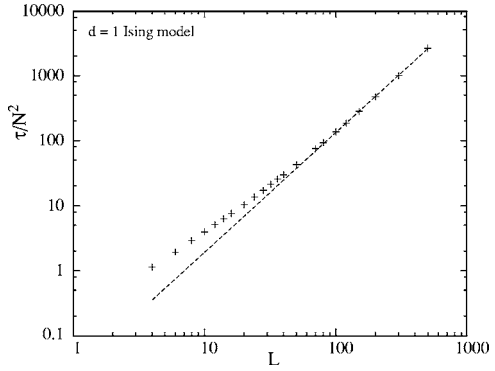

FIG. 1. Scaling of the mean first-passage time in the energy interval $[-N,+N]$ for the flat-histogram random walker of the $d=1$ Ising model. The dashed line corresponds to a power-law fit in the range $70 \leqslant N \leqslant 500$, which yields an exponent of $z=1.814\pm0.014$.

first-passage time for the $d=1$, 2, 3 and mean-field Ising models and the q-state Potts model in $d=2$.

A. The mean first-passage times for the ising model in $d=1$, 2, and 3

For the $d=1$ Ising model, the mean first-passage time was calculated using the exact density of states and 50 000 MC measurements for each value of N. A power-law fit in the range $70 \leqslant N \leqslant 500$ gives $z=1.814\pm0.014$ as shown in Fig. 1. However, there is some curvature in the data which indicates a deviation from this power-law scaling and the effective value of z seems to increase as larger system sizes are taken into account. Thus we cannot eliminate the possibility that $z=2$.

A crude argument that $z \leqslant 2$ comes from considering the two lowest states of a spin chain. The ground state has all spins aligned, and the first excited state has a single misaligned domain bounded by two domain walls. The distance between the domain walls is typically the order of L, the linear dimension of the system. For single-spin flip dynamics the domain walls perform a random walk so that the typical number of sweeps required to make the transition from the first excited state to the ground state scales as L^2. If this diffusion time were to hold for all energies, the result would be $z=2$. However, for higher energy states the domain wall diffusion time becomes smaller, so $z=2$ is an upper bound for the $d=1$ Ising model.

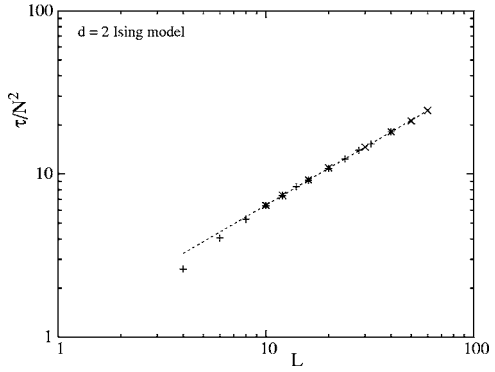

FIG. 2. Scaling of the mean first-passage time in the energy interval $[-2N,+2N]$ for the flat-histogram random walker of the $d=2$ Ising model. The + symbols are the Clark simulations; the $\times$ symbols refer to the ETH results. The dashed line corresponds to a power-law fit, which yields $z=0.743\pm0.002$.

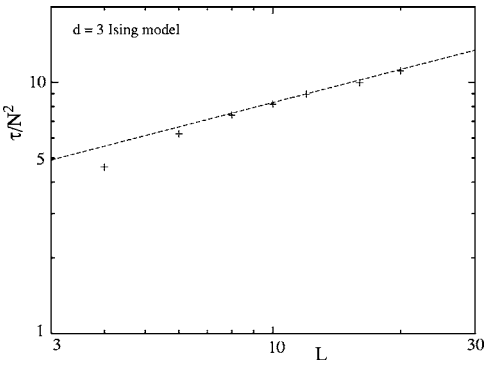

FIG. 3. Scaling of the mean first-passage time in the energy interval $[-3N,+3N]$ for the flat-histogram random walker of the $d=3$ Ising model. The dashed line corresponds to a power-law fit for $L \geqslant 8$, which yields $z=0.438\pm0.010$.

The exact density of states also was used for $d=2$ [11]. The combined results from Ref. [6] and this work are shown in Fig. 2 where each data point represents a total of 50 000 measurements. A power-law fit of the exponent z for $10 \leqslant L \leqslant 64$ gives $z=0.743\pm0.002$.

For the $d=3$ Ising model the exact Ising density of states cannot be calculated exactly, except for very small systems. We first used the Wang-Landau algorithm [1,2] to estimate the density of states. The criteria for the completion of each iteration in the determination of $g(E)$ was that each energy be visited at least five times and that the variance of the histogram be less than 10% of the mean. Because $g(E)$ is symmetric about $E=0$, the calculated values for $g(E)$ and $g(-E)$ were averaged. For each value of $N, g(E)$ was independently calculated ten times and then each $g(E)$ was used to obtain τ by averaging over 5000 measurements. The results for the first-passage time τ were then averaged over ten

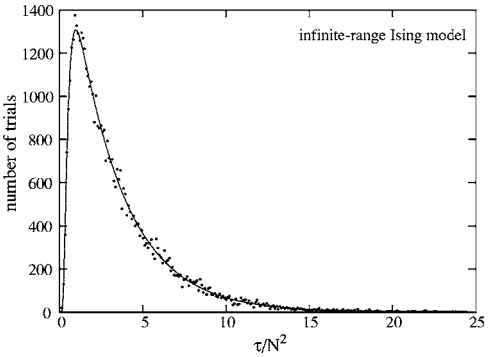

FIG. 4. Comparison of the distribution of first-passage times for the mean-field Ising model computed by iterating the master equation (line) and by direct simulation (points) for $N=32$. The master equation results were normalized to 50 000 MC estimates.

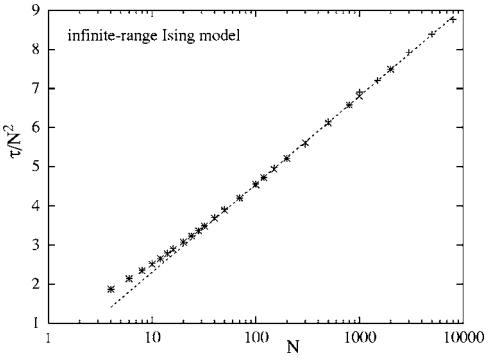

FIG. 5. Scaling of the mean first-passage time in the magnetization interval $[-N,+N]$ for the flat-histogram random walk of the mean-field Ising model. Note the linear scale on the y axis. The + symbol denotes the simulation results and × denotes results derived from the master equation. The dashed line corresponds to a logarithmic fit of the combined results for $L \geqslant 50$. The slope is found to be 0.977 ± 0.004.

independent runs. The results are shown in Fig. 3 and yield $z=0.438\pm0.010$.

B. Mean-field Ising model

The results of Sec. II A indicate that the value of z decreases with increasing dimension. This dependence suggests that it would be interesting to compute τ for the Ising model in the mean-field limit. We consider the "infinite-range" Ising model for which every spin interacts with every other spin with an interaction strength proportional to $1/N$. In this system the energy E and magnetization M are simply related, and it is convenient to express the density of states in terms of the latter. For each value of N we did 50 000 MC measurements of τ in the range $4 \leqslant N \leqslant 8000$.

We also calculated τ from a master equation. Because flipping a spin changes the magnetization by ±2, the master equation takes the form (suitably modified near the extremes of $M=\pm N$)

$$\begin{aligned} P(M,t+1) = {} & P(M,t) + T(M+2,M)P(M+2,t) \\ & - [T(M,M+2) + T(M,M-2)]P(M,t) \\ & + T(M-2,M)P(M-2,t), \end{aligned} \tag{2}$$

where $P(M,t)$ is the probability that the system has magnetization M at time t and $T(M_1,M_2)$ is the probability of a transition from a state with magnetization M_1 to one with M_2. The transition probabilities T are products of the probability of choosing a spin in the desired direction times the probability of accepting the flip. The latter is the Wang-Landau probability, $W(M_1,M_2)$, which is given by

$$W(M_1,M_2) = \min\left[1, \frac{g(M_1)}{g(M_2)}\right]. \tag{3}$$

The probability of choosing a spin in the desired direction is determined as follows. In a transition in which the magnetization increases, a down spin must be chosen. The probability of choosing a down spin is the number of down spins $N_\downarrow$ divided by the total number of spins

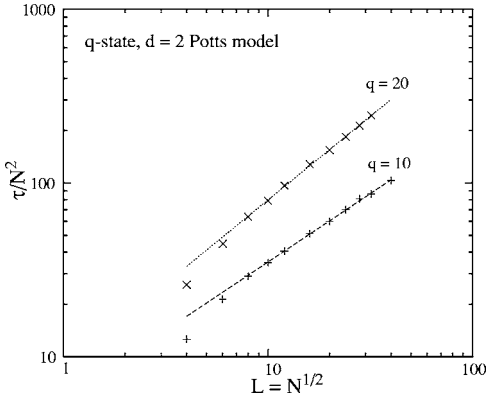

FIG. 6. Scaling of the mean first-passage time in the energy interval $[-2N,0]$ for the flat-histogram random walk of the $d=2$, $q=10$, and $q=20$ Potts models. The dashed lines correspond to power-law fits for $L \geqslant 8$ and yield $z=0.786\pm0.010$ and $z=0.961\pm0.008$ for the $q=10$ and $q=20$ Potts models, respectively.

$$\frac{N_\downarrow(M)}{N} = \frac{N-M}{2N}. \tag{4a}$$

Similarly, in a transition in which the magnetization decreases, an up spin must be chosen. The probability is

$$\frac{N_\uparrow(M)}{N} = \frac{N+M}{2N}. \tag{4b}$$

The transition probabilities can now be written as

$$T(M+2,M) = \frac{N_\uparrow(M+2)}{N} W(M+2,M), \tag{5a}$$

$$T(M,M+2) = \frac{N_\downarrow(M)}{N} W(M,M+2), \tag{5b}$$

$$T(M-2,M) = \frac{N_\downarrow(M-2)}{N} W(M-2,M), \tag{5c}$$

$$T(M,M-2) = \frac{N_\uparrow(M)}{N} W(M,M-2). \tag{5d}$$

The Wang-Landau probability $W(M_1,M_2)$ can be simplified for this system in the following way: The density of states in terms of the magnetization is

$$g(M) = \binom{N}{N_\uparrow}. \tag{6}$$

If $|M|$ increases in the transition, that is, $|M_1|<|M_2|$, then $W(M_1,M_2)=1$ because the ratio of the density of states exceeds unity. If $|M|$ decreases, there are two cases. For $M<0$, $N_\uparrow$ increases by 1 and M increases by 2, and we have

$$W(M,M+2)=\frac{g(M)}{g(M+2)}=\frac{\binom{N}{N_\uparrow}}{\binom{N}{N_\uparrow+1}}=\frac{N+M+2}{N-M}. \quad (7a)$$

If $M>0$, $N_\uparrow$ decreases by 1 and M decreases by 2, and

$$W(M,M-2)=\frac{g(M)}{g(M-2)}=\frac{\binom{N}{N_\uparrow}}{\binom{N}{N_\uparrow-1}}=\frac{N-M+2}{N+M}. \quad (7b)$$

Equations (3), (7a), and (7b) can be combined into the following simple expression:

$$W(M,M\pm 2)=\min\left[1,\frac{N\pm M+2}{N+|M|}\right]. \quad (8)$$

To compute the mean first-passage time, we take the state with magnetization N (the state with the highest magnetization) to be absorbing. The initial condition is

$$P(M,0)=\begin{cases}1 & M=-N\\ 0 & \text{otherwise}.\end{cases} \quad (9)$$

To compute τ, we iterate Eq. (2) using Eqs. (5a)–(5d), (7a), and (7b) and compute $\Delta P(N,t)$, the change in the value of $P(N,t)$ after each iteration t. The mean first-passage time is then given by

$$\tau=\sum_t t\Delta P(N,t). \quad (10)$$

A comparison of the distribution of first-passage times calculated from the master equation and from a direct simulation is shown in Fig. 4. As shown in Fig. 5, the scaling of the mean first-passage time with N is consistent with $\tau/N^2\sim\ln N$. The fit includes both the master equation and simulation results.

The logarithmic behavior of τ can be determined analytically from the transition rates using the first-passage time methods described in Ref. [12]. Let $t(m)$ be the mean first-passage time starting from magnetization $m=M/N$ and ending at the absorbing boundary $m=1$ with a reflecting boundary at $m=-1$. We seek $\tau=t(-1)$. The mean first-passage time satisfies

$$t(m)=[t(m-\delta m)+\delta t]T_-(m)+[t(m+\delta m)+\delta t]T_+(m)+[t(m)+\delta t]T_0(m), \quad (11)$$

where $T_-(m)=T(M,M-2)$, $T_+(m)=T(M,M+2)$ are the transition probabilities, $T_0=1-T_--T_+$ is the waiting probability, $\delta m=2/N$ is the magnetization step size, and $\delta t=1$ is the time unit for a step. Equation (11) can be cast in differential form by expanding $t(m)$ to second order in δm

$$1+h(m)t'(m)+f(m)t''(m)=0, \quad (12)$$

where

$$f(m)=(\delta m^2/2\delta t)[T_+(m)+T_-(m)], \quad (13)$$

and $h(m)=(\delta m/\delta t)[T_+(m)-T_-(m)]$. The absorbing boundary requires that

$$t(1)=0. \quad (14)$$

At the reflecting boundary, the transition to the left vanishes, $T_-(-1)=0$, and we have

$$t(-1)=[t(-1+\delta m)+\delta t]T_+(-1)+[t(-1)+\delta t][1-T_+(-1)]. \quad (15)$$

If we expand Eq. (15) to first order in δm, we obtain the boundary condition

$$t'(-1)=-\frac{\delta m}{2f(-1)}, \quad (16)$$

where $f(-1)=(\delta m^2/2\delta t)T_+(m)$

The flat-histogram feature of the Wang-Landau method means that the random walk satisfies the simple detailed balance condition,

$$T_+(m)=T_-(m+\delta m), \quad (17a)$$

$$T_-(m)=T_+(m-\delta m). \quad (17b)$$

We expand the detailed balance conditions to first order in δm and find a relation between the coefficients in Eq. (12)

$$h(m)=f'(m), \quad (18)$$

with the result that Eq. (12) reduces to

$$1+f'(m)t'(m)+f(m)t''(m)=0. \quad (19)$$

If we take into account the reflecting boundary condition, the solution to the differential equation is

$$t(m)=\tau-\int_{-1}^{m}\frac{x+1+\delta m/2}{f(x)}dx, \quad (20)$$

where $\tau=t(-1)$ is the first-passage time we are seeking. We use the absorbing boundary condition to find that

$$\tau=\int_{-1}^{1}\frac{x+1+\delta m/2}{f(x)}dx. \quad (21)$$

Because $f(m)$ is an even function, the first term in the integrand vanishes. We substitute $\delta m=2/N$ and use Eq. (13) for $f(m)$ to obtain the leading behavior of τ

$$\tau=2N^2\int_{-1}^{1}\frac{1}{1-T_0(x)}dx. \quad (22)$$

From Eqs. (4a), (4b), (5a)–(5d), and (8) we have that $T_0(m)=|m|-1/N$, which yields the desired result

$$\tau=2N^2\int_{-1}^{1}\frac{dm}{1-|m|+1/N}\sim N^2\log N. \quad (23)$$

Note that the integral can be interpreted as the average waiting time and that this average is dominated by the two boundaries.

C. Potts model

In Ref. [7] it was shown that the nonvanishing scaling exponent z quantifies the slowing down of the flat-histogram simulation near a critical point. Most strikingly, this slowing down can be observed as a suppression of the local diffusivity of the flat-histogram random walker (see Sec. IV). The Ising model undergoes a second-order phase transition. Here we extend our analysis to first-order phase transitions as is observed in q-state Potts models with $q>4$ on a square lattice.

Because the exact density of states is not known for the q-state Potts model, we follow the same procedure as for the Ising model in $d=3$ making use of the Wang-Landau algorithm. Based on these estimates for the density of states, we measured the mean first-passage time as a function of the linear system size L as is illustrated in Fig. 6. For $q=10$ we find that the exponent z is $z=0.786\pm0.010$, which is close to our result for the $d=2$ Ising model, thereby showing the close proximity of this weak first-order phase transition to a continuous phase transition. For the $q=20$ Potts model we find $z=0.961\pm0.008$, implying increasing slowing down as the strength of the first-order transition increases. We cannot rule out exponential slowing down for even larger values of q as was observed for multicanonical simulations of droplet condensation in the $d=2$ Ising model [13].

Recently, an algorithm [14] was introduced that makes possible the calculation of the partition function for Potts models with arbitrary values of q. The algorithm relies on an efficient cluster algorithm [15] to measure the number of connected components in the corresponding Fortuin-Kasteleyn representation [16]. For small values of q where the transition changes from second to first order, this algorithm can accurately estimate the density of states for significantly larger systems than the flat-histogram algorithms described here.

III. CHANGING THE DYNAMICS: CLUSTER UPDATES

In Sec. II we found that single-spin flip flat-histogram Monte Carlo simulations suffer from slowing down, that is, $z>0$, for the Ising model in $d=1$, 2, and 3. The critical slowing down associated with single-spin flip algorithms near critical points in the canonical ensemble has been reduced by the introduction of cluster algorithms such as the Swendsen-Wang [17] and the Wolff algorithms [18]. In this section we follow a similar approach and combine efficient cluster updates with flat-histogram simulations.

In conventional cluster algorithms [17,18] that simulate the canonical ensemble, clusters of parallel spins are built up by adding aligned spins to the cluster with a probability $p(\beta)$, which explicitly depends on the temperature $\beta=1/T$. When simulating a broad-histogram ensemble in energy there is no explicit notion of temperature, and thus no straightforward analog of a cluster algorithm. Recently, Reynal and Diep suggested a solution of this problem by using an estimate of the microcanonical temperature $\beta(E)=dS(E)/dE$ [19].

A complementary approach is to sample a broad-histogram ensemble using an alternative representation of the system's partition function for which it is possible to genuinely introduce cluster updates. An example is the multibondic method introduced by Janke and Kappler [5] and Carroll *et al.* [8], which uses the Fortuin-Kasteleyn representation [16] in the context of multicanonical sampling. Although the multibondic method performs *local* updates of the graph in the Fortuin-Kasteleyn representation and cluster updates of the spin configurations, Yamaguchi and Kawashima were able to show that a *global* and *rejection free* update of the graph in the Fortuin-Kasteleyn representation is possible [20]. We will discuss methods based on graph representations in Sec. III A.

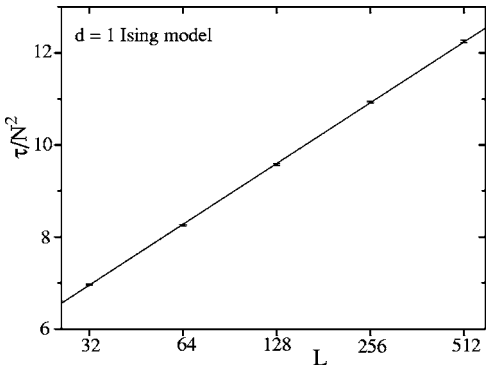

FIG. 7. Scaling of the mean first-passage time applying cluster updates in the bond representation for the $d=1$ Ising model. Note the linear scale of the y axis.

The above methods can only be applied to Ising and Potts models. In Sec. III B we introduce a representation using multigraphs that allows cluster updates for continuous spin models in the context of broad-histogram ensembles.

In the following we assume that the Hamiltonian of the system has the form:

$$H[\sigma] = -\sum_{\langle ij\rangle} h_{ij}[\sigma_i,\sigma_j], \tag{24}$$

where σ denotes the configuration of the system, σ_i the spin of site i, and the sum is over all bonds $\langle ij\rangle$ that define the lattice of the system.

A. Graph representation

We first review the cluster methods based on an additional graph variable and present results showing that $z=0$ for multibondic flat-histogram simulations of Ising models. In the spin representation the canonical partition function of the Ising-Potts model is given by $Z=\Sigma_\sigma W_\sigma(\sigma)$, where the summation is over all spin configurations, and the weight of a spin configuration is $W_\sigma(\sigma)=e^{-\beta H[\sigma]}$.

The q-state Ising-Potts system also can be represented by a sum over all graphs ω that can be embedded into the lattice. The partition function is then

$$Z=\sum_\omega W_\omega(\omega). \tag{25}$$

The weight of a graph ω is given by

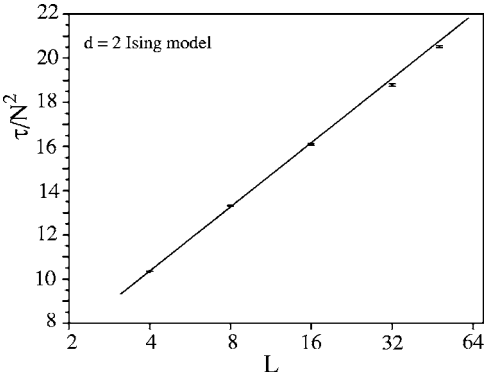

FIG. 8. Scaling of the mean first-passage time using cluster updates in the spin-bond representation of the $d=2$ Ising model. Note the linear scale of the y axis.

$$W_\omega(\omega) = p^{n(\omega)}(1-p)^{n_b-n(\omega)}q^{c(\omega)}, \tag{26}$$

where $n(\omega)$ is the number of of bonds in the graph $\omega, c(\omega)$ is the number of connected components (counting isolated sites), and n_b is the total number of bonds in the lattice. Because the graph ω can be viewed as a subgraph of the lattice, its bonds are sometimes referred to as "occupied bonds," and the bond probability of the ferromagnetic Ising-Potts model with $h_{ij}[\sigma_i,\sigma_j]=\delta_{\sigma_i,\sigma_j}$ is given by [16]

$$p(\beta) = 1 - e^{-\beta}. \tag{27}$$

A third representation of the Ising-Potts model is the spin-bond representation, which is employed in the Swendsen-Wang algorithm [17]. In the spin-bond representation the system is characterized by both spins and bonds, with the requirement that a bond can be occupied only if it is satisfied, that is, if the two spins connected by a bond in ω have the same value. In this representation the partition function is

$$Z = \sum_{\sigma,\omega} W_{\sigma\omega}(\sigma,\omega), \tag{28}$$

and the weight is given by

$$W_{\sigma\omega}(\sigma,\omega) = p^{n(\omega)}(1-p)^{n_b-n(\omega)}\Delta(\sigma,\omega), \tag{29}$$

where $\Delta(\sigma,\omega)=1$ if all occupied bonds are satisfied and zero otherwise.

It is natural to introduce a density of states for each of the three representations: For the spin representation in terms of the energy E, we write

$$g_\sigma(E) = \sum_\sigma \delta(H[\sigma]-E). \tag{30}$$

For the bond representation in terms of the number of occupied bonds n and number of clusters c, we have

$$g_\omega(n,c) = \sum_\omega \delta(n(\omega)-n)\,\delta(c(\omega)-c). \tag{31}$$

And for the spin-bond representation in terms of the number of occupied bonds n, we write

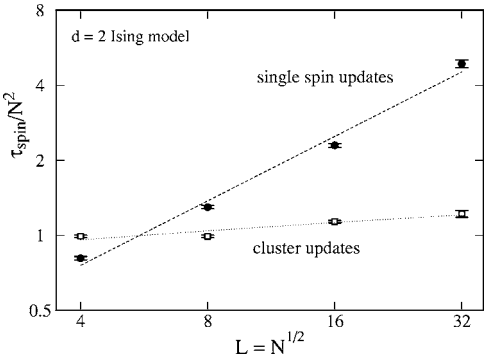

FIG. 9. Scaling of the mean first-passage time for the $d=2$ Ising model using the simple high temperature representation with single spin updates (•) and using the spin-multigraph representation with Swendsen-Wang cluster updates (□). The lines are drawn as a guide to the eye.

$$g_{\sigma\omega}(n) = \sum_{\omega,\sigma} \delta(n(\omega)-n)\Delta(\sigma,\omega). \tag{32}$$

The corresponding forms of the partition function are

$$Z = \sum_E g_\sigma(E)e^{-\beta E} \text{ (spin)} \tag{33}$$

$$Z = \sum_{n,c} g_\omega(n,c)p^n(1-p)^{n_b-n}q^c \text{ (bond)} \tag{34}$$

$$Z = \sum_n g_{\sigma\omega}(n)p^n(1-p)^{n_b-n}. \text{ (spin bond)} \tag{35}$$

The Wang-Landau algorithm can be applied in all three representations. As discussed in Sec. II, the spin representation has a relatively large value of z.

The bond representation generally requires a two-dimensional histogram, but for the $d=1$ Ising-Potts model, the number of clusters c is completely determined by the number of occupied bonds n and a one-dimensional histogram is sufficient. We simulated the $d=1$ Ising model in the bond representation and found that the mean first-passage time scales as $N^2 \log L$, as shown in Fig. 7. It is noteworthy that the domain wall arguments used to explain the large value of z in one dimension do not apply in the bond representation.

In higher dimensions n does not determine c, and it is simpler to use the spin-bond representation. The resulting multibondic algorithm [20] is very similar to the Swendsen-Wang algorithm, and can be summarized as follows:

(1) Choose a bond at random.

(2) If the bond is satisfied and occupied (unoccupied), make it unoccupied (occupied) with probability $\min[g_{\omega\sigma}(n)/g_{\omega\sigma}(n'),1]$, where n and n' are the number of occupied bonds before and after the change, respectively. Update the observables.

(3) If the bond is unsatisfied, the system stays in its original state and the observables are updated.

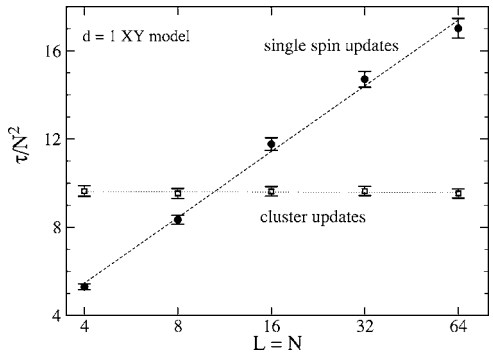

FIG. 10. Scaling of the mean first-passage time to reach a low energy state for the $d=1$ *XY* model using single spin updates (•) and Swendsen-Wang cluster updates (□) in the spin multigraph representation. Lines are drawn as a guide to the eye.

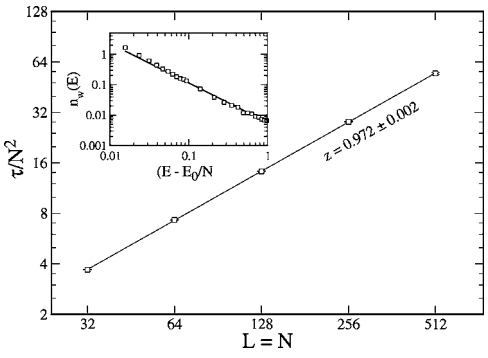

FIG. 11. Scaling of the mean first-passage time of the optimized ensemble for the $d=1$ Ising model using Metropolis updates. The solid line corresponds to a power-law fit with exponent 0.972 ± 0.002. The inset shows the optimized histogram $n_w(E)$, which exhibits a power-law divergence $[(E-E_0)/N]^{1.30}$.

(4) After one sweep of the lattice, identify clusters (connected components) and assign spin values to them with equal probability.

The algorithm is ergodic and satisfies detailed balance. The ergodicity is obvious from the fact that the Swendsen-Wang algorithm is ergodic. In general, the detailed balance relation can be written as

$$\pi(\sigma,\omega)p(\sigma\to\sigma'|\omega)p(\omega\to\omega'|\sigma') = \pi(\sigma',\omega')p(\omega'\to\omega|\sigma')p(\sigma'\to\sigma|\omega), \tag{36}$$

where $\pi(\sigma,\omega)$ is the equilibrium probability of microstate (σ,ω), and $p(\sigma\to\sigma'|\omega)$ is the transition probability from σ to σ' given that the bond configuration is fixed at ω.

The algorithm is designed to sample the distribution,

$$\pi(\sigma,\omega) = \frac{1}{(n_b+1)g_{\sigma\omega}(n(\omega))}\Delta(\sigma,\omega). \tag{37}$$

The transition probability of a spin flip is

$$p(\sigma\to\sigma'|\omega) = \frac{1}{q^{c(\omega)}}\Delta(\sigma,\omega), \tag{38}$$

and the transition probability of a bond change $(\omega'\neq\omega)$ is given by

$$p(\omega\to\omega'|\sigma) = \frac{1}{n_b}\min\left[\frac{g_{\sigma\omega}(n(\omega))}{g_{\sigma\omega}(n(\omega'))},1\right], \tag{39}$$

because the probability of choosing a bond to change is $1/n_b$, and the probability of accepting the change is $\min[g_{\sigma\omega}(n(\omega))/g_{\sigma\omega}(n(\omega')),1]$. It is easy to verify that the detailed balance relation Eq. (36) is satisfied.

We have applied this algorithm to the $d=2$ Ising model and measured the mean first-passage time, which also scales as $N^2\log L$, as shown in Fig. 8. At least 10^5 first-passage times were measured for each value of L.

B. Spin-multigraph representation

We now consider a representation of the partition function that can be used to implement cluster updates in flat-histogram simulations of *continuous* spin models such as $O(n)$ models. Like the spin-bond representation, it has a configuration and a graph variable. In contrast to the spin-bond representation, the graph is a multigraph and can therefore have multiple bonds between two vertices. Before introducing this representation, we briefly review how a simple high-temperature series representation can be sampled in a flat-histogram simulation.

We write the partition function as a series in the inverse temperature β

$$Z = \sum_n g_\beta(n)\beta^n, \tag{40}$$

with a corresponding density of states given by

$$g_\beta(n) = \sum_\sigma \frac{(-H[\sigma])^n}{n!}. \tag{41}$$

If $H[\sigma]\leqslant 0$, then the weight of a configuration-order pair

$$W_\beta(\sigma,n) = \frac{(-H[\sigma])^n}{n!}, \tag{42}$$

is always positive, and we can perform a flat-histogram simulation in the extended phase space (σ,n) using σ and n updates. If the σ and n updates are performed independently of each other, an update $n\to n'$ (for example $n'=n\pm1$) is accepted with probability

$$P_{n\to n'} = \min\left[(-H[\sigma])^{n'-n}\frac{n!}{n'!}\frac{g_\beta(n)}{g_\beta(n')},1\right]. \tag{43}$$

A configuration update $\sigma\to\sigma'$ (for example, the change of the configuration of a single site $\sigma_i\to\sigma_i'$) at fixed n is accepted with probability

$$P_{\sigma\to\sigma'} = \min\left[\left(\frac{H[\sigma']}{H[\sigma]}\right)^n,1\right]. \tag{44}$$

In a practical computation n is truncated at some cutoff Λ, restricting the largest inverse temperature that can be reached to $\beta_{\max}\sim\Lambda$ [21].

Because the weight of a configuration as defined in Eq. (42) is not a product of bond weights on the lattice, the (σ,n) representation cannot be directly used to implement cluster updates. To obtain such a representation, we start from

$$e^{-\beta H[\sigma]} = \prod_{\langle ij\rangle} e^{\beta h_{ij}[\sigma_i,\sigma_j]}, \tag{45}$$

and expand each exponential term in β with a separate integer variable ω_{ij}

$$e^{\beta h_{ij}[\sigma_i,\sigma_j]} = \sum_{\omega_{ij}=0}^{\infty} \frac{h_{ij}[\sigma_i,\sigma_j]^{\omega_{ij}}\beta^{\omega_{ij}}}{\omega_{ij}!}. \tag{46}$$

We identify each set of ω_{ij} on the lattice with a multigraph ω that has ω_{ij} bonds between sites i and j. If we write $n(\omega)=\Sigma\omega_{ij}$ for the total number of bonds in the multigraph $\omega, g_\beta(n)$ as defined in Eq. (41) is given by

$$g_\beta(n) = \sum_{\sigma,\omega} \delta(n(\omega)-n)\prod_{\langle ij\rangle} \frac{h_{ij}[\sigma_i,\sigma_j]^{\omega_{ij}}}{\omega_{ij}!}, \tag{47}$$

where the sum is over all spin configurations σ and all multigraphs ω on the lattice. The corresponding partition function is given by Eq. (40), and the weight of a configuration-multigraph pair is given by

$$W_\beta(\sigma,\omega) = \prod_{\langle ij\rangle} W_{ij}(\sigma_i,\sigma_j,\omega_{ij}), \tag{48}$$

where $W_{ij}(\sigma_i,\sigma_j,\omega_{ij})=h_{ij}[\sigma_i,\sigma_j]^{\omega_{ij}}/\omega_{ij}!$ is the weight of the bond $\langle ij\rangle$ in the lattice. The energies in the system have to be shifted so that $h_{ij}[\sigma_i,\sigma_j]\geqslant 0$ to ensure that all weights are positive. Again two types of updates, one in the multigraph ω and one in the configuration σ are needed. Before discussing them in detail, we mention some differences of the spin-multigraph representation compared to the representations used in Sec. III A. In the spin-multigraph representation we cannot, in general, integrate explicitly over all configurations to obtain a multigraph-only representation similar to the bond representation except for Ising or Potts models. It also is not possible to apply the global graph update of Yamaguchi and Kawashima [20] to the multigraph so that local updates have to be used. Finally, the representation can be viewed as the classical limit of the stochastic series expansion method for quantum systems [22]. The order of the operators in the operator string used in the series expansion method is unimportant for a classical system, so that one only has to count how often an operator $h_{ij}[\sigma_i,\sigma_j]$ occurs (ω_{ij}).

If we use a flat-histogram algorithm in order n, a multigraph update $\omega\rightarrow\omega'$ (as before $\omega'_{ij}=\omega_{ij}\pm 1$ for one randomly chosen bond $\langle ij\rangle$ is a good choice) will be accepted with probability

$$P_{\omega\rightarrow\omega'} = \min\left[\frac{g_\beta(n(\omega))}{g_\beta(n(\omega'))}\prod_{\langle ij\rangle}\frac{W_{ij}(\sigma_i,\sigma_j,\omega'_{ij})}{W_{ij}(\sigma_i,\sigma_j,\omega_{ij})},1\right]. \tag{49}$$

Because the weight of a spin-multigraph configuration is a product of bond weights, a cluster update scheme such as the Swendsen-Wang [17] or Wolff algorithm [18] can be used to update the spin configuration. After choosing a flip operation for the spins, the only difference with the canonical cluster algorithms is the probability of adding a site to a Swendsen-Wang cluster: If W_{ij} is the weight if either none or both sites are flipped and $\tilde{W}_{ij}$ the weight after a flip of one of the sites [note that this requires $W_{ij}(\sigma_i,\sigma_j)=W_{ij}(\sigma'_i,\sigma'_j)$ and $W_{ij}(\sigma'_i,\sigma_j)=W_{ij}(\sigma_i,\sigma'_j)$], the probability of adding a site to a cluster is given by

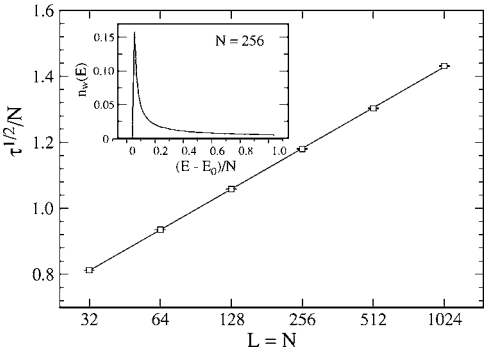

FIG. 12. Scaling of the mean first-passage time of the optimized ensemble for the $d=1$ Ising model using N-fold way updates (note the linear scale on the y axis). The solid line corresponds to a logarithmic fit. The inset shows the optimized histogram $n_w(E)$.

$$P_{\text{connect}} = 1 - \min\left[\frac{\tilde{W}_{ij}}{W_{ij}},1\right]. \tag{50}$$

For the canonical ensemble simulation of the Ising model with $W(\uparrow,\uparrow,\beta)=e^\beta$ and $W(\uparrow,\downarrow,\beta)=1$, Eq. (50) reduces to the well-known $1-e^{-\beta}$ for parallel spins as discussed earlier. In the spin-multigraph representation the probability to add a site to a cluster is

$$P_{\text{connect}} = 1 - \min\left[\left(\frac{h_{ij}[\sigma_i,\sigma'_j]}{h_{ij}[\sigma_i,\sigma_j]}\right)^{\omega_{ij}},1\right]. \tag{51}$$

Thus only sites that are connected by at least one bond of the multigraph can be added to the cluster. For the Ising model with $h_{ij}[\sigma_i,\sigma_j]=\delta_{\sigma_i,\sigma_j}$, the weight of a configuration (σ,ω) as defined in Eq. (48) is

$$W_\beta(\sigma,\omega) = \Delta(\sigma,\omega)\frac{\beta^{n(\omega)}}{\prod_{\langle ij\rangle}\omega_{ij}!}, \tag{52}$$

where $\Delta(\sigma,\omega)=1$ if all bonds in the multigraph ω are satisfied and 0 otherwise, and P_{connect} simplifies to 1 if two sites are connected by one or more bonds in the multigraph and to 0 if they are not. $O(n)$ models can be simulated by performing the spin update with respect to a mirror plane that is randomly chosen for one cluster update [18].

In a local update scheme, a configuration update $\sigma\rightarrow\sigma'$ at fixed ω will be accepted with probability

$$P_{\sigma\rightarrow\sigma'} = \min\left[\prod_{\langle ij\rangle}\left(\frac{h_{ij}[\sigma'_i,\sigma'_j]}{h_{ij}[\sigma_i,\sigma_j]}\right)^{\omega_{ij}},1\right]. \tag{53}$$

We have applied a flat-histogram sampling of this representation to the $d=2$ Ising model with periodic boundary

conditions and to the $d=1$ XY model with open boundary conditions. To compare the effectiveness of the cluster updates to the single-spin flip updates, we measure the mean first-passage time τ_{spin} from a random state σ at $n=0$, which is equivalent to $\omega_{ij}=0$ for all $\langle ij \rangle$, to an ordered state. For the Ising model this state is chosen to be the ground state and for the $d=1$ model any state that is within 2.5% of the width of the spectrum to the ground state. We perform a fixed ratio of spin to multigraph updates and measure the mean first-passage time in spin updates for different ratios of spin to multigraph updates. The time τ_{spin} converges to a fixed number with an increasing fraction of multigraph updates, and for sufficiently many multigraph updates we can assume the multigraph is equilibrated for a particular configuration. The cutoff Λ is chosen so that the average n for which the ground state is reached is much smaller than Λ.

Figure 9 shows the scaling of τ_{spin} for the $d=2$ Ising model for single-spin flips in the simple representation of Eq. (42) and for Swendsen-Wang cluster updates in the spin-multigraph representation. The single-spin flip updates show a power-law behavior with $z=0.85\pm0.06$, while the cluster updates can be fitted by a logarithm or a power law with $z=0.10\pm0.02$. Figure 10 shows the scaling of τ_{spin} for the $d=1$ XY model using single spin updates and Swendsen-Wang cluster updates in the spin-multigraph representation.

IV. CHANGING THE ENSEMBLE: OPTIMIZING THE SAMPLED HISTOGRAM

An alternative to changing the dynamics from local to cluster updates for flat-histogram sampling is to optimize the simulated statistical ensemble and retain local spin-flip updates. When simulating a flat-histogram ensemble, the random walker is slowed down close to a critical point, which is reflected in the suppressed local diffusivity. To overcome this slowing down, we can use this information and define a new statistical ensemble $w(E)$ by feeding back the local diffusivity $D(E)$ [7]. After the feedback additional weight is shifted toward the critical point. The histogram is no longer flat, but exhibits a peak at the critical energy, that is, resources in the form of local updates are shifted toward the critical point. Ultimately, this feedback procedure results in an optimal histogram $n_w(E)$ that is proportional to the inverse of the square root of the local diffusivity $D(E)$

$$n_w(E) \propto \frac{1}{\sqrt{D(E)}}. \tag{54}$$

For the $d=2$ Ising model it was shown in Ref. [7] that the mean first-passage time scales as $\tau \propto (N \ln N)^2$ for the optimized ensemble and satisfies the scaling of an unbiased Markovian random walk up to logarithmic corrections. It also was found that the distribution of statistical errors in the calculated density of states $g(E)=n_w(E)/w(E)$ is uniformly distributed in energy, in contrast to calculations based on flat-histogram sampling.

Here we show results for the optimized ensemble of the $d=1$ Ising model for both Metropolis and N-fold way updates [23]. For single-spin flip Metropolis updates, we find that the histogram is shifted toward the ground-state energy $E_0=-N$ and follows a power-law divergence, $n_w(E) \propto [(E-E_0)/N]^{1.30\pm0.01}$ (see the inset of Fig. 11). However, the mean first-passage time of the random walk in the energy interval $[-N,0]$ still exhibits a power-law slowdown with $z=0.972\pm0.002$ as illustrated in the main panel of Fig. 11. This remaining slowdown may originate from the slow dynamics that occur whenever two domain walls reside on neighboring bonds, and a spin flip of the intermediate spin is suggested with a probability of $1/N$. This argument suggests that changing the dynamics from simple Metropolis updates to N-fold way updates would strongly increase the probability of annihilating two neighboring domain walls.

We optimized the ensemble for N-fold way updates and found that the mean first-passage time is further reduced and scales as $\tau \propto (N \ln N)^2$ as shown in Fig. 12. This scaling behavior corresponds to the results for the optimized ensemble for the $d=2$ Ising model [7]. However, for the $d=2$ model the optimized ensembles for both Metropolis and N-fold way updates resulted in the same scaling behavior with $\tau \propto (N \ln N)^2$.

V. CONCLUSIONS

We have studied the performance of flat-histogram simulations for Ising models in different dimensions. By measuring the mean first-passage time of a random walk in energy space, we find that flat-histogram simulations exhibit slowing down in one, two, and three dimensions with an exponent $z>0$ that decreases as a function of the spatial dimension d. For the infinite range Ising model the slowdown vanishes (up to logarithmic corrections).

We demonstrated that the slowing down of the flat-histogram simulations can be overcome by either introducing cluster updates in connection with flat-histogram sampling or by optimizing the simulated broad-histogram ensemble and keeping local updates. In order to apply cluster updates to continuous spin models in connection with broad-histogram simulations, we introduced a spin-multigraph-based representation. The broad-histogram simulations that take advantage of either cluster updates or an optimized statistical ensemble do not suffer from slowing down. The scaling of the mean first-passage time for these simulations is reduced to that of an unbiased Markovian random walk and therefore is optimal.

ACKNOWLEDGMENTS

We acknowledge financial support of the Swiss National Science Foundation and the U.S. National Science Foundation DMR-0242402 (Machta), and NSF DBI-0320875 (Colonna-Romano and Gould). We thank Sid Redner for useful discussions.

[1] Fugao Wang and D. P. Landau, Phys. Rev. E **64**, 056101 (2001).
[2] Fugao Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001).
[3] D. P. Landau, in *The Monte Carlo Method in the Physical Sciences*, edited by J. E. Gubernatis, AIP Conf. Proc. No. 690, (AIP, New York, 2003), p. 134.
[4] B. A. Berg and T. Neuhaus, Phys. Rev. Lett. **68**, 9 (1992).
[5] W. Janke and S. Kappler, Phys. Rev. Lett. **74**, 212 (1995).
[6] P. Dayal, S. Trebst, S. Wessel, D. Würtz, M. Troyer, S. Sabhapandit, and S. N. Coppersmith, Phys. Rev. Lett. **92**, 097201 (2004).
[7] S. Trebst, D. A. Huse, and M. Troyer, Phys. Rev. E **70**, 046701 (2004).
[8] M. S. Carroll, W. Janke, and S. Kappler, J. Stat. Phys. **90**, 1277 (1998).
[9] W. Janke and T. Sauer, Phys. Rev. E **49**, 3475 (1994).
[10] W. Janke and T. Sauer, J. Stat. Phys. **78**, 759 (1995).
[11] In Ref. [6] the exact density of states was found using the MATHEMATICA program of P. D. Beale, Phys. Rev. Lett. **76**, 78 (1996). Here we used a C program and an infinite precision integer arithmetic package that also implements Beale's approach. The latter approach allowed most of the calculation of the density of states to be spread over many processors.
[12] S. Redner, *A Guide to First-Passage Processes* (Cambridge Universiy, Cambridge, England, 2001).
[13] T. Neuhaus and J. S. Hager, J. Stat. Phys. **113**, 47 (2003).
[14] A. K. Hartmann, Phys. Rev. Lett. **94**, 050601 (2005).
[15] L. Chayes and J. Machta, Physica A **254**, 477 (1998).
[16] C. M. Fortuin and P. M. Kasteleyn, Physica (Amsterdam) **57**, 536 (1972).
[17] R. H. Swendsen and J.-S. Wang, Phys. Rev. Lett. **58**, 86 (1987).
[18] U. Wolff, Phys. Rev. Lett. **62**, 361 (1989).
[19] S. Reynal and H. T. Diep, cond-mat/0409529 (unpublished).
[20] C. Yamaguchi and N. Kawashima, Phys. Rev. E **65**, 056710 (2002).
[21] M. Troyer, S. Wessel, and F. Alet, Phys. Rev. Lett. **90**, 120201 (2003).
[22] A. W. Sandvik and J. Kurkijärvi, Phys. Rev. B **43**, 5950 (1991).
[23] A. B. Bortz, M. H. Kalos, and J. L. Lebowitz, J. Comput. Phys. **17**, 10 (1975).