

2000

# Monte Carlo Computation of Correlation Times of Independent Relaxation Modes at Criticality

M. P. Nightingale

*University of Rhode Island*, [nightingale@uri.edu](mailto:nightingale@uri.edu)

H. W.J. Blöte

Follow this and additional works at: [http://digitalcommons.uri.edu/phys\\_facpubs](http://digitalcommons.uri.edu/phys_facpubs)

Terms of Use

All rights reserved under copyright.

---

## Citation/Publisher Attribution

M. P. Nightingale and H. W. J. Blöte. (2000). "Monte Carlo computation of correlation times of independent relaxation modes at criticality." *Physical Review B*, 62(2), 1089. Available at: <http://dx.doi.org/10.1103/PhysRevB.62.1089>

This Article is brought to you for free and open access by the Physics at DigitalCommons@URI. It has been accepted for inclusion in Physics Faculty Publications by an authorized administrator of DigitalCommons@URI. For more information, please contact [digitalcommons@etal.uri.edu](mailto:digitalcommons@etal.uri.edu).

# Monte Carlo computation of correlation times of independent relaxation modes at criticality

M. P. Nightingale

*Department of Physics, University of Rhode Island, Kingston, Rhode Island 02881*

H. W. J. Blöte

*Faculty of Applied Sciences, Delft University, P.O. Box 5046, 2600 GA Delft, The Netherlands*

*Lorentz Institute, Leiden University, Niels Bohrweg 2, P.O. Box 9506, 2300 RA Leiden, The Netherlands*

(Received 13 January 2000)

We investigate aspects of the universality of Glauber critical dynamics in two dimensions. We compute the critical exponent  $z$  and numerically corroborate its universality for three different models in the static Ising universality class and for five independent relaxation modes. We also present evidence for universality of amplitude ratios, which shows that, as far as dynamic behavior is concerned, each model in a given universality class is characterized by a single nonuniversal metric factor which determines the overall time scale. This paper also discusses in detail the variational and projection methods that are used to compute relaxation times with high accuracy.

## I. INTRODUCTION

Critical-point behavior is a manifestation of power-law divergences of the correlation length and the correlation time. The power laws that describe the divergence of the correlation length on approach of the critical point are expressed by means of critical exponents that are dependent on the *direction* of this approach, which may, e.g., be ordering-field like or temperature like. The exponents describing the singularities in *thermodynamic* quantities can be expressed in terms of the same exponents. In addition to the exponents defining these power laws, another critical exponent, viz., the dynamic exponent  $z$ , is required for the singularities in the *dynamics*. This exponent  $z$  is defined by the relationship that holds between the correlation length  $\xi$  and the correlation time  $\tau$ , namely,  $\tau \propto \xi^z$ .

One of the directions along which one can approach the critical singularity is the finite-size direction; i.e., one increases the system size  $L$  while keeping the independent thermodynamic variables at their infinite-system critical values. In this case,  $\xi \propto L$  so that  $\tau \propto L^z$ . This relation has been used extensively to obtain the dynamic exponent  $z$  from finite-size calculations.

In this paper we deal with universality of dynamic critical-point behavior. One would not expect systems in different static universality classes to have the same dynamic exponents, and even within the same static universality class, different dynamics may have different exponents. For instance, in the case of the Ising model, Kawasaki dynamics which satisfies a local conservation law<sup>1</sup> has a larger value of  $z$  than Glauber dynamics,<sup>2</sup> in which such a conservation law is absent. Also the introduction of nonlocal spin updates, as realized, e.g., in cluster algorithms, is known to lead to a different dynamic universal behavior.<sup>3-5</sup>

Conservation laws and nonlocal updates tend to have a large effect on the numerical value of the dynamic exponents, but until fairly recently, numerical resolution of the expected differences of dynamic exponents of systems in different static universality classes for dynamics with local up-

dates has been elusive. This is caused by the difficulty of obtaining the required accuracy in estimates of the dynamic critical exponent. Under these circumstances it is evident that only a limited progress has been made with respect to the interesting questions regarding dynamic universality classes.

In this paper we present a detailed exposition of a method of computing dynamic exponents with high accuracy.<sup>6,7</sup> We consider single spin-flip Glauber dynamics. This is defined by a Markov matrix, and computation of the correlation time is viewed here as an eigenvalue problem, since correlation times can be obtained from the subdominant eigenvalues of the Markov matrix.

If a thermodynamic system is perturbed out of equilibrium, different thermodynamic quantities relax back at a different rates. More generally, there are infinitely many independent relaxation modes for a system in the thermodynamic limit. Let us label the models within a given universality class by means of  $\kappa$ , and denote by  $\tau_{Li\kappa}$  the autocorrelation time of relaxation mode  $i$  of a system of linear dimension  $L$ . In this paper we present strong numerical evidence that, as indeed renormalization group theory suggests, at criticality the relaxation times have the following factorization property:

$$\tau_{Li\kappa} \approx m_\kappa A_i L^z, \quad (1)$$

where  $m_\kappa$  is a *nonuniversal* metric factor, which differs for different representatives of the same universality class as indicated;  $A_i$  is a *universal* amplitude, which depends on the mode  $i$ ; and  $z$  is the *universal* dynamical exponent introduced above.

While the relaxation time of the slowest relaxation mode is obtained from the second-largest eigenvalue of the Markov matrix, lower-lying eigenvalues yield the relaxation times of faster modes. To compute these we construct, employing a Monte Carlo method, variational approximants for several eigenvectors. These approximants are called optimized trial vectors. The corresponding eigenvalues can then be estimated by evaluating with Monte Carlo techniques the

overlap of these trial vectors and the corresponding matrix elements of the Markov matrix in the truncated basis spanned by these optimized trial vectors. It should be noted that both the optimization scheme and the evaluation of these matrix elements critically depend on the fact that the Markov matrix is sparse. That is, the number of configurations accessible from any given configuration is equal to the number of sites only, rather than the number of possible spin configurations.

Given such fixed trial vectors, this approach has the advantage of simplicity and high statistical accuracy, but the disadvantage is that results are subject to systematic, variational errors, which only vanish in the ideal limit where the variational vectors become exact eigenvectors or span an invariant subspace of the Markov matrix. Since the condition is rarely satisfied in cases of practical interest, a projection Monte Carlo method is then used, to reduce the systematic error, but this is at the expense of an increase of the statistical errors. The method we use in this paper is a combination and generalization of the work of Umrigar *et al.*<sup>8</sup> and that of Ceperley and Bernu.<sup>9</sup>

To summarize, the Monte Carlo method discussed here consists of two phases. In the first phase, trial vectors are optimized. The ultimate, yet unattainable goal of this phase is to construct exact eigenvectors. In this phase of the computation, very small Monte Carlo samples are used, consisting typically of no more than a few thousand spin configurations. In the second phase, one performs a standard Monte Carlo computation in which one reduces statistical errors by increasing the length of the computation rather than the quality of the variational approximation.

The computed correlation times, derived from the partial solution of the eigenvalue problem as sketched above, are used in a finite-size analysis to compute the dynamic critical exponent  $z$ . We verify its universality for several models in the static universality class of the two-dimensional Ising model. We also address another manifestation of dynamic universality. As was mentioned, in addition to the usual static critical exponents, there is only one new exponent that governs the leading singularities of critical dynamics, viz.,  $z$ . Similarly, one would expect that, within the context of Glauber dynamics, the description of time-dependent critical-point amplitudes requires only a single nonuniversal metric factor to determine the time scale of each different model within a given universality class. Our results corroborate this idea, which is the immediate generalization to critical dynamics of work on static critical phenomena by Privman and Fisher.<sup>10</sup>

In this paper we apply the techniques outlined above to three different two-dimensional Ising models subject to Glauber-like spin dynamics. These models are defined on a simple quadratic lattice of size  $L \times L$  with periodic boundary conditions. The Hamiltonian  $\mathcal{H}$ , defined on a general spin configuration  $S = (s_1, s_2, \dots)$ , is given by

$$\frac{\mathcal{H}(S)}{kT} = -K \sum_{\langle ij \rangle} s_i s_j - K' \sum_{[kl]} s_k s_l, \quad (2)$$

where the first summation is on all nearest-neighbor pairs of sites of the square  $L \times L$  lattice, the second summation is on all next-nearest-neighbor pairs, and the Ising variables

$s_i, \dots, s_l$  assume values  $\pm 1$ . Periodic boundaries are used throughout. In particular, we focus on models described by three ratios  $\beta = K'/K$ , namely,  $\beta = -1/4$ , 0 (nearest-neighbor model) and 1 (equivalent-neighbor model). The nonplanar models for  $\beta \neq 0$  are not exactly solvable and their critical points are known only approximately. Yet it was demonstrated to a high degree of numerical accuracy that they belong to the static Ising universality class.<sup>11,12</sup> For the nearest-neighbor model the critical coupling is  $K = \frac{1}{2} \ln(1 + \sqrt{2})$ ; for the other two models estimates of the critical points are  $K = 0.190\,192\,680\,7(2)$  for  $\beta = 1$  and  $K = 0.697\,220\,7(2)$  for  $\beta = -1/4$ .<sup>12,13</sup>

We use the dynamics of the heat-bath algorithm with random site selection. The single-spin-flip dynamics is determined by the Markov matrix  $P$  defined as follows. The element  $P(S', S)$  is the transition probability of going from configuration  $S$  to  $S'$ . If  $S$  and  $S'$  differ by more than one spin,  $P(S', S) = 0$ . If both configurations differ by precisely one spin,

$$P(S', S) = \frac{1}{2L^2} \left\{ 1 - \tanh \left[ \frac{\mathcal{H}(S') - \mathcal{H}(S)}{2kT} \right] \right\}, \quad (3)$$

where  $L^2$  is the total number of spins. The diagonal elements  $P(S, S)$  follow from the conservation of probability,

$$\sum_{S'} P(S', S) = 1, \quad (4)$$

where  $S'$  runs over all possible  $2^{L^2}$  spin configurations.

We denote the probability of finding spin configuration  $S$  at time  $t$  by  $\rho_t(S)$ . By design, the stationary state of the Markov process is the equilibrium distribution

$$\rho_\infty(S) = \frac{\exp[-\mathcal{H}(S)/kT]}{Z} \equiv \frac{\psi_B(S)^2}{Z}, \quad (5)$$

where the normalization factor  $Z$  is the partition function.

The dynamical process defined by Eq. (3) is constructed so as to satisfy detailed balance, which is equivalent to the statement that the matrix  $\hat{P}$  with elements

$$\hat{P}(S', S) \equiv \frac{1}{\psi_B(S')} P(S', S) \psi_B(S) \quad (6)$$

is symmetric. Therefore the eigenvalues of  $P$  are real.

The Markov matrix determines the time evolution of  $\rho_t(S)$ , i.e.,

$$\rho_{t+1}(S) = \sum_{S'} P(S, S') \rho_t(S'). \quad (7)$$

The simultaneous probability distribution  $\rho_{t', t'+t}(S, S')$  that the system is in state  $S$  at time  $t'$  and in state  $S'$  at time  $t' + t$  is

$$\rho_{t', t'+t}(S, S') = P^t(S', S) \rho_{t'}(S), \quad (8)$$

where  $P^t(S', S)$  denotes the  $(S', S)$  element of the  $t$ th power of the matrix  $P$ . For sufficiently large times  $t'$ , one may take  $\rho_{t'}(S) = \rho_\infty(S)$  so that the autocorrelation function  $C_A(t)$  of an observable  $A$ , the average with respect to time  $t'$  of

$A(t')A(t+t')$ , can equivalently be written as the ensemble average  $\langle A(t')A(t'+t) \rangle$  for large  $t'$ . Thus

$$C_A(t) = \lim_{t' \rightarrow \infty} \sum_S \sum_{S'} A(S)A(S')\rho_{t',t'+t}(S,S'), \quad (9)$$

where  $A(S)$  denotes the value of  $A$  in a spin configuration  $S$ . After substitution of Eq. (9) and expansion of  $A(S)\rho_{t'}(S)$  in right-hand eigenvectors of the Markov matrix, it follows at once that the time-dependent correlation functions of a system of size  $L$  have the following form:

$$C_A(t) = \sum_i c_i \operatorname{sgn} \lambda_{Li}^t \exp\left[-\frac{t}{L^2 \tau_{Li}}\right], \quad (10)$$

where the dependence on the specific model  $\kappa$  has been suppressed in denoting by  $\tau_{Li}$  the relaxation times of the independent modes of the equilibration process. The  $\tau_{Li}$  are determined by the eigenvalues of the Markov matrix. We denote these eigenvalues  $\lambda_{Li}$  ( $i=0,1,2,\dots,2^{L^2}-1$ ), and order them so that  $1=\lambda_{L0} > |\lambda_{L1}| \geq |\lambda_{L2}| \geq \dots$ . Note that conservation of probability implies that  $\lambda_{L0}=1$ ; by construction, the corresponding right-hand eigenvector is the Boltzmann distribution.

The relaxation times are given by

$$\tau_{Li} = -\frac{1}{L^2 \ln|\lambda_{Li}|} \quad (i=1,2,\dots). \quad (11)$$

The factor  $L^2$  is inserted because, as usual, time is measured units of one flip per spin, which corresponds to  $L^2$  iterations of the process described by Eq. (7).

Note that the stochastic matrix  $P$  has the same symmetry properties as the Hamiltonian and the Boltzmann distribution. In addition to spin inversion, these symmetries include translations, reflections, and rotations of the  $L \times L$  lattice. It follows that each eigenvector of  $P$ , as well as its associated relaxation mode, has distinct symmetry properties that can be characterized by a set of ‘‘quantum numbers.’’ For instance, the eigenvector associated with the second-largest eigenvalue is antisymmetric under spin inversion and invariant under translations, reflections, and rotations. It describes the relaxation of the total magnetization; this process, with relaxation time  $\tau_{L1}$ , is thus the slowest relaxation mode contained in the stochastic matrix.

In this work, we restrict ourselves to relaxation modes that are invariant under geometric symmetries of the spin lattice. However, in addition to eigenvectors that are symmetric under spin inversion, we include antisymmetric ones, so as to obtain the longest relaxation time. As a consequence of this restriction to geometric invariance, spatially nonhomogeneous relaxation processes fall outside the scope of this work.

By design, the stationary state of the Markov process is the equilibrium state

$$\rho_\infty(S) = \frac{\exp[-\mathcal{H}(S)/kT]}{Z} \equiv \frac{\psi_B(S)^2}{Z}, \quad (12)$$

where the normalization factor  $Z$  is the partition function.

The dynamical process defined by Eq. (3) is constructed so as to satisfy detailed balance, which is equivalent to the statement that the matrix  $\hat{P}$  with elements

$$\hat{P}(S',S) \equiv \frac{1}{\psi_B(S')} P(S',S) \psi_B(S) \quad (13)$$

is symmetric.

The layout of the rest of this paper is as follows. In Sec. II we discuss the general principles of the method used in this paper. The expressions in this section feature exhaustive summation over all spin configurations, which renders them useless for practical computations. The Monte Carlo summation methods employed instead are discussed in Sec. III. Trial vectors, a vital ingredient of the method, are discussed in Sec. IV, and numerical results are discussed in Sec. V. Finally, Sec. VI contains a discussion of issues that remain to be addressed by future work.

## II. EXACT SUMMATION

### A. Variational approximation

#### *Single eigenvector*

In this section we discuss trial vector optimization, which is used to reduce the statistical errors in the Monte Carlo computation of eigenvalues of the Markov matrix. First, we review the case of a single eigenvalue,<sup>8,14,15</sup> and then we generalize to optimization of multiple trial vectors. In this section, we discuss the exact expressions involving summation over all possible spin configurations. In cases of practical interest, these expressions cannot be evaluated as written; for their approximate evaluation one uses the Monte Carlo methods discussed in Sec. III.

A powerful method of optimizing a single, many-parameter trial vector, say,  $|\psi_T\rangle$ , is minimization of the variance of the *configurational eigenvalue*, which in the context of quantum Monte Carlo is called the *local energy*. That is, define  $\psi_T(S) = \langle S | \psi_T \rangle$  for an arbitrary configuration  $S$ . We wish to satisfy the eigenvalue equation

$$\psi'_T(S) = \lambda \psi_T(S), \quad (14)$$

where the prime indicates matrix multiplication by  $\hat{P}$ , i.e.,  $f'(S) \equiv \sum_{S'} \hat{P}(S,S') f(S')$  for any function  $f$  defined on the spin configurations. Even if  $\psi_T$  is not an eigenvector, one can define the *configurational eigenvalue* by

$$\lambda_c(S) = \begin{cases} \frac{\psi'_T(S)}{\psi_T(S)}, & \text{if } \psi_T(S) \neq 0, \\ 0, & \text{otherwise.} \end{cases} \quad (15)$$

If  $\psi_T$  is not an eigenvector, Eq. (14) gives an overdetermined set of equations for  $\lambda$  for a given  $\psi_T$  and a sufficiently big set of configurations  $S$ . One can obtain a least-squares estimate of the eigenvalue  $\lambda$  by minimizing the squared residual of Eq. (14). This yields the usual variational estimate

$$\lambda(p) = \frac{\langle \psi_T | \hat{P} | \psi_T \rangle}{\langle \psi_T | \psi_T \rangle} = \frac{\sum_S \psi'_T(S) \psi_T(S)}{\sum_S \psi_T(S)^2} = \frac{\sum_S \lambda_c(S) \psi_T(S)^2}{\sum_S \psi_T(S)^2}, \quad (16)$$

which is the average of the configurational eigenvalue  $\lambda_c$ .

The standard Rayleigh-Ritz variational method, which can be used for the largest eigenvalue, consists in maximization of  $\bar{\lambda}(p)$  with respect to the parameters  $p$ . However, one can formulate a different optimization criterion as follows. The gradient of  $\bar{\lambda}(p)$  with respect to  $\psi_T(S)$  is

$$\frac{\partial \bar{\lambda}(p)}{\partial \psi_T(S)} = 2 \frac{\psi'_T(S) - \bar{\lambda}(p) \psi_T(S)}{\sum_{S'} \psi_T(S')^2}. \quad (17)$$

Clearly, this gradient vanishes for *any* eigenvector and this suggests as an alternative optimization criterion minimization of the magnitude of the gradient of a normalized trial vector  $\psi_T$ . With respect to Eq. (14), this corresponds to minimization of the normalized squared residual

$$\begin{aligned} \chi^2(p) &= \frac{\sum_S [\psi'_T(S) - \bar{\lambda}(p) \psi_T(S)]^2}{\sum_S \psi_T(S)^2} \\ &= \frac{\sum_S [\lambda_c(S) - \bar{\lambda}(p)]^2 \psi_T(S)^2}{\sum_S \psi_T(S)^2} = \frac{\langle \psi_T | (\hat{P} - \bar{\lambda})^2 | \psi_T \rangle}{\langle \psi_T | \psi_T \rangle}, \end{aligned} \quad (18)$$

which equals the variance of the configurational eigenvalue, as shown.

### B. Multiple eigenvectors

Minimization of  $\chi^2(p)$  is a valid criterion for any eigenvector, but if this is used without the equivalent of an orthogonalization procedure, one would in practice simply keep reproducing an approximation to the same eigenvector, the dominant one most of the time. Since orthogonalization is not easily implemented with Monte Carlo methods, we utilize straightforward generalizations of Eqs. (14) and (16) to deal with more than one eigenvalue and eigenvector. Equation (18) is a little problematic in this respect, as will become clear.

Suppose we start from a set of  $n$  trial vectors  $\psi_{Ti}$  ( $i = 0, 1, \dots, n-1$ ). We can then write Eq. (14) in matrix form

$$\psi'_{Ti}(S) = \sum_{j=0}^{n-1} \hat{\Lambda}_{ij} \psi_{Tj}(S). \quad (19)$$

As before, the prime on the left-hand side indicates matrix multiplication by  $\hat{P}$ , and again Eq. (19) for all  $i$  and  $S$  form an overdetermined set of equations for the matrix  $\hat{\Lambda}_{ij}$ . These equations have no solution, unless the  $n$  basis vectors  $\psi_{Ti}$

span an invariant subspace of the matrix  $\hat{P}$ , which in non-trivial applications of course is never the case. Again, however, one can solve for the matrix elements  $\hat{\Lambda}_{ij}$  in a least-squares sense. This yields

$$\hat{\Lambda} = \hat{\mathcal{P}} \hat{N}^{-1}, \quad (20)$$

where

$$\hat{N}_{ij} = \sum_S \psi_{Ti}(S) \psi_{Tj}(S) = \langle \psi_{Ti} | \psi_{Tj} \rangle \quad (21)$$

and

$$\hat{\mathcal{P}}_{ij} = \sum_S \psi'_{Ti}(S) \psi_{Tj}(S) = \langle \psi_{Ti} | \hat{P} | \psi_{Tj} \rangle. \quad (22)$$

Note that although these matrix elements depend on the normalization of the  $|\psi_{Ti}\rangle$ , the matrix  $\hat{\Lambda}$  is invariant under an overall change of normalization.

By diagonalization of the  $n \times n$  matrix  $\hat{\Lambda}$  one obtains an approximate, partial eigensystem of the Markov matrix. More specifically, suppose that

$$\hat{\Lambda} = D^{-1} \text{diag}(\tilde{\lambda}_0, \dots, \tilde{\lambda}_{n-1}) D. \quad (23)$$

The eigenvalues  $\tilde{\lambda}_0 > \tilde{\lambda}_1 \geq \dots \geq \tilde{\lambda}_{n-1}$  of  $\hat{\Lambda}$  are variational lower bounds for the exact eigenvalues of the Markov matrix  $P$ , in the sense that  $\tilde{\lambda}_i \leq \lambda_i$ ,<sup>16</sup> if the exact eigenvalues are numbered such that  $\tilde{\lambda}_0 > \tilde{\lambda}_1 \geq \dots$ , in contrast with the convention used in the discussion following Eq. (10). This property is a consequence of the interlacing property of the eigenvalues of symmetric matrices and their submatrices, also known as the separation theorem.<sup>17</sup> Note that in denoting the eigenvalues we omit the index  $L$  indicating system size, where this is not confusing. The approximate eigenvectors  $\tilde{\psi}_i$  are given by

$$\tilde{\psi}_i = \sum_{j=0}^{n-1} D_{ij} \psi_{Tj}, \quad (24)$$

which can be verified as follows: multiply Eq. (19) through by  $D_{ki}$ , and sum on  $i$  to verify that  $\tilde{\psi}'_k$  proportional to  $\tilde{\psi}_k$ . The expressions derived above are usually<sup>9</sup> derived by starting from the linear combinations given in this last equation. The  $D_{ij}$  then are treated as variational parameters, and are determined by requiring stationarity of the Rayleigh quotient. This yields the following equation for the  $D_{ij}$ :

$$\sum_j D_{ij} \hat{P}_{jk} = \tilde{\lambda}_i \sum_j D_{ij} \hat{N}_{jk}, \quad (25)$$

a generalized eigenvalue problem equivalent to the eigenvalue problem defined by  $\hat{\Lambda}$  defined in Eq. (20).

Next we discuss the generalization to more than one trial vector of minimization of the variance as given by Eq. (18). In this context it is important to keep in mind that the variational approximation is invariant under replacement of the basis vectors by a nonsingular linear superposition. This yields a similarity transformation of  $\hat{\Lambda}$ , and leaves invariant the approximate eigenvalues and eigenvectors. Of course,

one would like to have an optimization criterion that shares this invariance. The squared residual of Eq. (19) fails in this respect. This sum is not even invariant under a simple rescaling of the basis functions  $\psi_{T_i}$ , and there is no obvious normalization comparable to the one used in Eq. (18).

One way to perform the optimization in an invariant way is for each choice of the optimization parameters to compute linear combinations  $\sum_j V_{ij} |\psi_{T_j}\rangle$ , where  $V$  is the  $n \times n$  matrix, such that  $V\hat{\Lambda}V^{-1}$  is diagonal. Each of these linear combinations defines a  $\chi^2$  via Eq. (18), and the parameters in the basis functions can then be optimized by minimization of these  $n$  sums of squares. One may define a convex sum of these  $\chi_i^2$  and optimize all parameters for all basis functions simultaneously with respect to this combined object function, or, as we did in our computations, one can perform the optimization iteratively one vector at a time for eigenvalues with increasing distance from the top of the spectrum.

Another approach that also yields invariant results is to perform the optimization by dividing the set of configurations into several subsets, and computing a matrix  $\hat{\Lambda}$  for each subset. One can then minimize the variance of the eigenvalues over these subsets. This is the procedure we followed to produce the results reported in this paper.

We have not investigated which of the two procedures described above is superior. Both do have a problem in common, namely, that for a wide class of variational basis vectors, they give rise to a singular or nearly singular optimization problem. This is a consequence of the fact that the basis states are not unique, even if the eigenvalue problem has a unique solution. For the optimization problem this means that there are many almost equivalent solutions, a problem commonly encountered when one performs (nonlinear) least-squares parameter fits.

More specifically, if the basis vectors are such that a linear combination of trial vectors can be expressed exactly (or to good approximation) in the same functional form as the trial vectors themselves, then there is a gauge symmetry (or an approximate gauge symmetry) that yields a class of equivalent (or almost equivalent) solutions of the minimization problem. That is, if  $|\psi_{T_i}\rangle$  is a solution,  $\sum_j V_{ij} |\psi_{T_j}\rangle$  is an equivalent (or almost equivalent) solution for any  $V$ . This problem can be solved straightforwardly by fixing the gauge and performing the optimization subject to the constraint that  $\sum_j V_{ij} |\psi_{T_j}\rangle \propto |\psi_{T_i}\rangle$ . If the gauge symmetry holds only approximately, this additional constraint may produce a sub-optimal solution.

### C. Beyond the variational approximation

The eigenvalues obtained by the variational scheme discussed in the previous sections have a bias caused by admixture of eigenvectors in that part of the spectrum that is being ignored. This variational bias can be reduced in principle arbitrarily as follows.<sup>9</sup>

Let us introduce generalized matrices with elements

$$\hat{N}_{ij}(t) = \langle \psi_{T_i} | \hat{P}^t | \psi_{T_j} \rangle \quad (26)$$

and

$$\hat{P}_{ij}(t) = \langle \psi_{T_i} | \hat{P}^{t+1} | \psi_{T_j} \rangle. \quad (27)$$

For  $t=0$  these expressions reduce to Eqs. (21) and (22). One can view the matrix elements for  $t>0$  as having been obtained by the substitution  $|\psi_{T_i}\rangle \rightarrow \hat{P}^{t/2} |\psi_{T_i}\rangle$ . Expansion in the exact eigenvectors immediately shows that the spectral weights are reduced of ‘‘undesirable’’ eigenvectors with less dominant eigenvalues, so that the vectors  $\hat{P}^{t/2} |\psi_{T_i}\rangle$  span a more nearly invariant subspace of  $\hat{P}$  than the original states. This process, however, becomes numerically unstable as  $t \rightarrow \infty$ , since in that case all basis vectors of the same symmetry collapse onto the corresponding dominant state.

### III. MONTE CARLO SUMMATION

Obviously, the summation over all spin configurations used in the expressions in the previous section can, in general, be done only for small systems. In this section, we discuss the Monte Carlo estimators of the expressions presented above. In principle, matrix multiplication involves summation over all configurations and therefore is not practically feasible. However, for the dynamics we consider in this paper the summation required for the matrix multiplication by  $P$  in  $\langle S | \hat{P} | \psi_T \rangle$  is an exception, since for a given  $S$  there are only  $L^2$  configurations  $S'$  from which  $S$  can be reached with one or fewer spin flips, and these are the only configurations for which  $P(S, S')$  does not vanish. For all other configuration sums a Monte Carlo method is used.

To produce a Monte Carlo estimate of  $\chi^2(p)$  as given in Eq. (18), sample  $M_c$  spin configurations  $S_\alpha$  with  $\alpha = 1, \dots, M_c$  from the Boltzmann distribution  $\psi_B(S_\alpha)^2$ . This yields a Monte Carlo estimate of  $\bar{\lambda}(p)$ :

$$\bar{\lambda}(p) \approx \frac{\sum_\alpha \hat{\psi}'_T(S_\alpha) \hat{\psi}_T(S_\alpha)}{\sum_\alpha \hat{\psi}_T(S_\alpha)^2}, \quad (28)$$

where

$$\hat{\psi}_T(S_\alpha) = \frac{\psi_T(S_\alpha)}{\psi_B(S_\alpha)}, \quad (29)$$

$$\hat{\psi}'_T(S_\alpha) = \frac{\psi'_T(S_\alpha)}{\psi_B(S_\alpha)}. \quad (30)$$

Similarly,

$$\chi^2(p) \approx \frac{\sum_\alpha [\hat{\psi}'_T(S_\alpha) - \bar{\lambda} \hat{\psi}_T(S_\alpha)]^2}{\sum_\alpha \hat{\psi}_T(S_\alpha)^2}. \quad (31)$$

Parameter optimization for a single vector is done by generating a sample of a few thousand configurations and subsequently varying the parameters  $p$  while keeping this sample fixed. The same applies to the optimization of more than one vector, in which case estimates of the required matrix elements  $\hat{N}_{ij}$  and  $\hat{P}_{ij}$  are computed by

$$\hat{N}_{ij} \approx \sum_{\alpha} \hat{\psi}_{T_i}(S_{\alpha}) \hat{\psi}_{T_j}(S_{\alpha}) \equiv \tilde{N}_{ij} \quad (32)$$

and

$$\hat{P}_{ij} \approx \sum_{\alpha} \hat{\psi}'_{T_i}(S_{\alpha}) \hat{\psi}_{T_j}(S_{\alpha}) \equiv \tilde{P}_{ij}. \quad (33)$$

We attached tildes to the symbols on the right-hand side of Eqs. (32) and (33) to indicate that the corresponding quantities are stochastic variables, which is important to keep in mind for the following discussion.

Since the matrix  $\hat{P}$  is symmetric, one might be inclined to symmetrize its estimator  $\tilde{P}_{ij}$  with respect to  $i$  and  $j$ . This symmetrization, however, destroys the zero-variance principle satisfied by the expressions as written. As mentioned before, the eigensystem of  $\hat{\Lambda}$  is obtained exactly and without statistical noise, if the basis vectors  $\psi_{T_i}$  are linear combinations of  $n$  exact eigenvectors. In that ideal case,  $\hat{\Lambda}$  is uniquely determined by Eq. (19) even if it is applied only to an subset of configurations  $S$ . The same holds for a weighted subset as represented by a Monte Carlo sample. Even though the matrices  $\hat{P}$  and  $\hat{N}$  themselves depend on the weights and the subset, factors responsible for statistical noise cancel in the product  $\hat{P}\hat{N}^{-1}$ . To demonstrate this, we write the estimator in matrix form

$$\tilde{N} = \tilde{\Psi} \tilde{\Psi}^{\dagger} \quad (34)$$

and

$$\tilde{P} = \tilde{\Psi}' \tilde{\Psi}^{\dagger}, \quad (35)$$

where  $\tilde{\Psi}$  is a rectangular matrix with elements  $\tilde{\Psi}_{i\alpha} = \hat{\psi}_{T_i}(S_{\alpha})$  and  $\tilde{\Psi}'_{i\alpha} = \hat{\psi}'_{T_i}(S_{\alpha})$ . Equation (19) in matrix form becomes

$$\tilde{\Psi}' = \hat{\Lambda} \tilde{\Psi}. \quad (36)$$

Clearly, if this last equation holds,  $\tilde{P}\tilde{N}^{-1} = \tilde{\Psi}' \tilde{\Psi}^{\dagger} (\tilde{\Psi} \tilde{\Psi}^{\dagger})^{-1} = \hat{\Lambda}$  without statistical noise, as announced.

We have assumed that one matrix multiplication by the Markov matrix can be done exactly; repeated multiplications rapidly become intractable. This is a problem for the computation of the matrix elements given in Eqs. (26) and (27). To obtain a statistical estimate of these matrix elements, one generates a time series with the Markov matrix  $P$ . One then exploits the fact that in the steady state of the Markov process, the relative probability of finding configurations  $S_1, S_2, \dots, S_{t+1}$  in immediate succession is given by

$$P(S_{t+1}|S_t) \cdots P(S_2|S_1) \psi_B(S_1)^2. \quad (37)$$

For a Monte Carlo run of length  $M_c$ , this property allows us to write

$$\begin{aligned} \hat{N}_{ij}^{(t)} &= \sum_{S_0, \dots, S_t} \psi_{T_i}(S_t) \hat{P}(S_t|S_{t-1}) \cdots \hat{P}(S_1|S_0) \psi_{T_j}(S_0) \\ &= \sum_{S_0, \dots, S_t} \hat{\psi}_{T_i}(S_t) \hat{\psi}_{T_j}(S_0) P(S_t|S_{t-1}) \cdots P(S_1|S_0) \\ &\quad \times \psi_B(S_0)^2 \end{aligned} \quad (38)$$

$$\approx M_c^{-1} \sum_{\sigma=1}^{M_c} \hat{\psi}_{T_i}(S_{\sigma+t}) \hat{\psi}_{T_j}(S_{\sigma}) \quad (39)$$

Similarly,

$$\begin{aligned} \hat{P}_{ij}^{(t)} &= \sum_{S_0, \dots, S_{t+1}} \psi_{T_i}(S_{t+1}) \hat{P}(S_{t+1}|S_t) \cdots \hat{P}(S_1|S_0) \psi_{T_j}(S_0) \\ &= \sum_{S_0, \dots, S_t} \hat{\psi}'_{T_i}(S_t) \hat{\psi}_{T_j}(S_0) \hat{P}(S_t|S_{t-1}) \cdots \hat{P}(S_1|S_0) \\ &\quad \times \psi_B(S_0)^2 \end{aligned} \quad (40)$$

$$\begin{aligned} &\approx (2M_c)^{-1} \sum_{\sigma=1}^{M_c} [\hat{\psi}'_{T_i}(S_{\sigma+t}) \hat{\psi}_{T_j}(S_{\sigma}) + \hat{\psi}'_{T_i}(S_{\sigma}) \\ &\quad \times \hat{\psi}_{T_j}(S_{\sigma+t})]. \end{aligned} \quad (41)$$

The first term in expression (41) follows immediately from expression (40); to obtain the second term one has to use the time reversal symmetry of a stochastic process that satisfies detailed balance, viz.,

$$\begin{aligned} &P(S_{t+1}|S_t) \cdots P(S_2|S_1) \psi_B(S_1)^2 \\ &= P(S_1|S_2) \cdots P(S_t|S_{t+1}) \psi_B(S_{t+1})^2. \end{aligned} \quad (42)$$

Again, these estimators satisfy the zero-variance principle mentioned above, as long as the expressions are used as written, i.e., without symmetrization with respect to  $i$  and  $j$ .

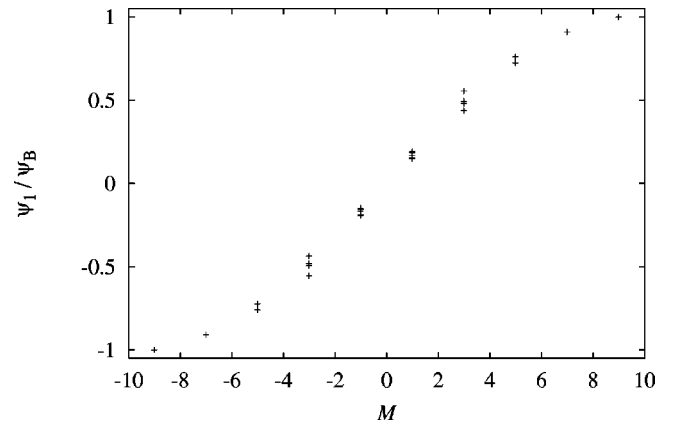


FIG. 1. Prefactor  $\psi_1 / \psi_B$  of the first subdominant eigenvector of the Markov matrix vs total magnetization  $M$  for a  $3 \times 3$  nearest-neighbor Ising ( $\beta=0$ ) lattice. For each value of  $M$   $\psi_1(S) / \psi_B(S)$  is plotted for all configurations  $S$  with  $M(S)=M$ . All 512 points are plotted, but because of symmetries, many coincide.

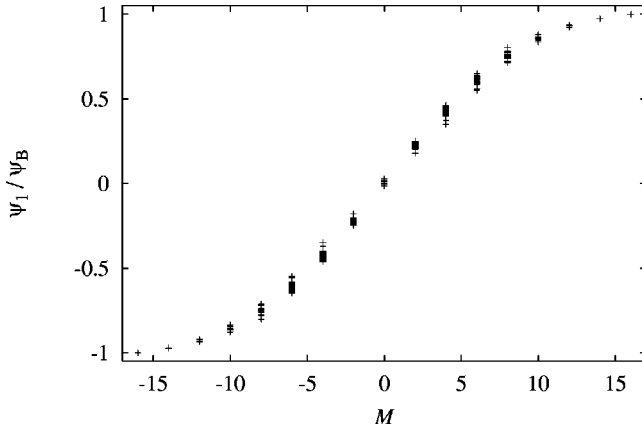


FIG. 2. Prefactor  $\psi_1/\psi_B$  of the first subdominant eigenvector of the Markov matrix vs total magnetization  $M$  for a  $4 \times 4$  nearest-neighbor Ising ( $\beta=0$ ) lattice.

#### IV. TRIAL VECTORS

As we mentioned, the form of the trial vectors used in these calculations is a major factor determining the statistical accuracy of the results. It not too difficult to make an initial guess for the form of the eigenvector corresponding to the second largest eigenvalue of the Markov matrix. Numerically exact calculations for small systems show that this eigenvector is antisymmetric under spin inversion, which is a manifestation of the longevity of fluctuations of the magnetization and not a peculiarity of small systems.

This suggests the following initial approximation of the eigenvector belonging to the second largest eigenvalue, the first subdominant eigenvector of the symmetrized Markov matrix  $\hat{P}$ :

$$\psi_{T1}(S) = m \psi_B(S), \quad (43)$$

where  $m$  is the average magnetization. Figures 1–3 are plots of  $\psi_1(S)/\psi_B(S)$  versus the total magnetization  $M = L^2 m$  for the exact eigenvector  $\psi_1$  computed for  $3 \times 3$ ,  $4 \times 4$ , and  $5 \times 5$  nearest-neighbor Ising systems. For all three, the prefactor  $m$  in Eq. (43) clearly captures a significant part of the

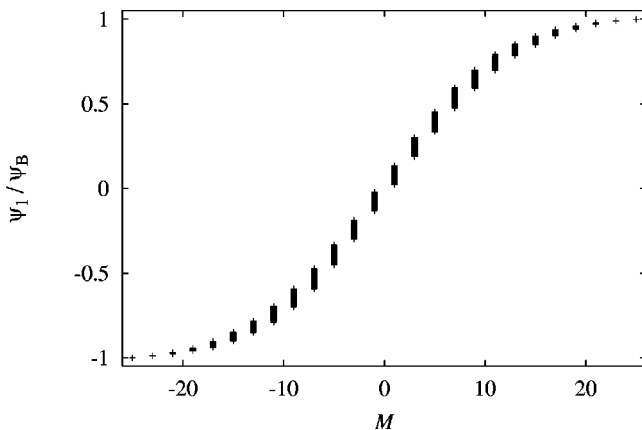


FIG. 3. Prefactor  $\psi_1/\psi_B$  of the first subdominant eigenvector of the Markov matrix vs total magnetization  $M$  for a  $5 \times 5$  nearest-neighbor Ising ( $\beta=0$ ) lattice. Although up to 400 configuration collapse onto a single point, crowding prevents individual resolution of most data points.

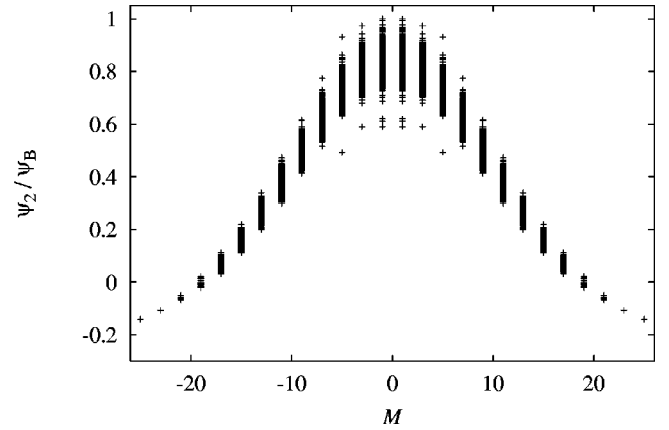


FIG. 4. Prefactor  $\psi_2/\psi_B$  of the second subdominant eigenvector of the Markov matrix vs total magnetization  $M$  for a  $5 \times 5$  nearest-neighbor Ising ( $\beta=0$ ) lattice.

truth, but there are two shortcomings. First of all, there is scatter, which indicates that  $\psi_1(S)/\psi_B(S)$  is a function of more than just the magnetization. Second, the ‘‘curve’’ is nonlinear. The latter problem can be cured quite easily by replacing  $m$  in Eq. (43) by an odd polynomial  $m(1 + a_2 m^2 + \dots)$  with coefficients  $a_k$  to be determined variationally.

Similarly, computations for small systems (see Sec. V A for further details) suggest that the second largest eigenvalue is associated with an eigenvector that is even under spin inversion, as illustrated in Fig. 4. A trial vector of this form is readily constructed by replacing  $m$  on the right-hand side of Eq. (43) by a polynomial even in  $m$ . It turns out that the general picture as just described is largely independent of  $L$ .

More in general, the plots shown in Figs. 1–7 strongly suggest that the subdominant eigenvectors of the Markov matrix  $P$ , subject to the imposed spin, rotation, and translation symmetries, are reasonably approximated by the Boltzmann distribution multiplied by a mode-dependent function of the magnetization. As can be seen in Figs. 1–7, the number of nodes of this prefactor increases by 1 as one steps down the spectrum, but it is also clear that, especially for the less dominant eigenvectors, the residual variance is significant.

To begin to address the problem of the scatter and to improve the trial vector systematically, it is necessary to

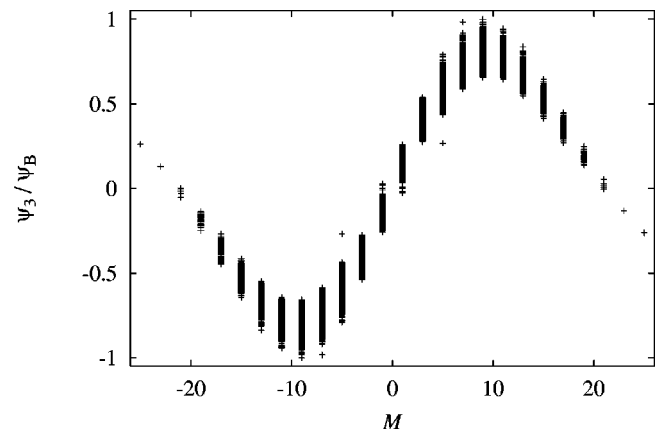


FIG. 5. Prefactor  $\psi_3/\psi_B$  of the third subdominant eigenvector of the Markov matrix vs total magnetization  $M$  for a  $5 \times 5$  nearest-neighbor Ising ( $\beta=0$ ) lattice.



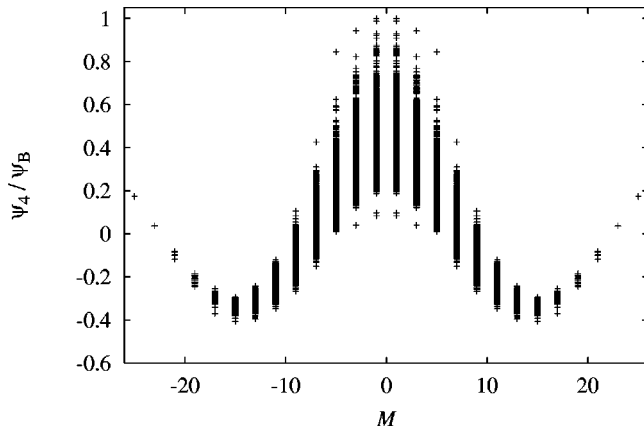


FIG. 6. Prefactor  $\psi_4/\psi_B$  of the fourth subdominant eigenvector of the Markov matrix vs total magnetization  $M$  for a  $5 \times 5$  nearest-neighbor Ising ( $\beta=0$ ) lattice.

identify other important variables besides the magnetization and to incorporate them in the trial vector. We tried multi-spin correlations involving nearby spins but after considerable failed experimentation we established that long-wavelength fluctuations of the magnetization are the suitable variables. This is reasonable when one compares, e.g., the eigenvalue equations for  $\psi_0(S')$  and  $\psi_1(S')$  and realizes that the eigenvalues differ only very little from unity except for very small systems. We therefore used the Fourier components of the spin configuration, which are defined by

$$m_{\mathbf{k}} = L^{-2} \sum_{l_1=1}^L \sum_{l_2=1}^L \exp\left[\frac{2\pi i}{L}(k_1 l_1 + k_2 l_2)\right] s_{l_1 l_2}, \quad (44)$$

where  $\mathbf{k}=(k_1, k_2)$  with  $0 \leq k_1, k_2 < L$ , and  $s_{l_1 l_2}$  denotes the spin at lattice site  $(l_1, l_2)$ . Note that  $m \equiv m_{0,0}$ . If we restrict ourselves to eigenvectors that are translationally invariant, the arguments presented in the previous paragraph yield the following trial odd or even vectors:

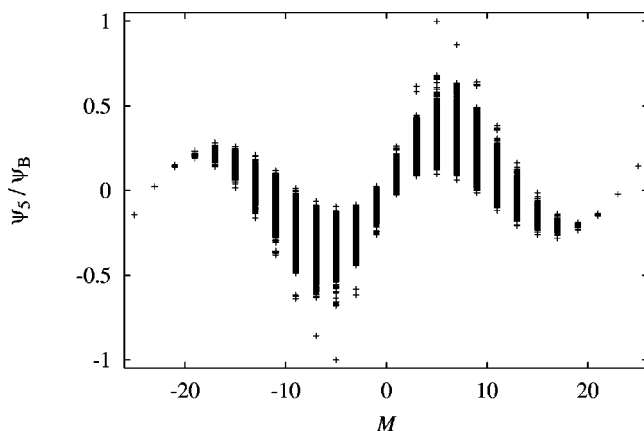


FIG. 7. Prefactor  $\psi_5/\psi_B$  of the fifth subdominant eigenvector of the Markov matrix vs total magnetization  $M$  for a  $5 \times 5$  nearest-neighbor Ising ( $\beta=0$ ) lattice.

$$\begin{aligned} \psi_T(S) = \psi_B(S) & \left[ a^\pm(m) + \sum'_{\mathbf{k}_1, \mathbf{k}_2} a_{\mathbf{k}_1 \mathbf{k}_2}^\pm(m) m_{\mathbf{k}_1} m_{\mathbf{k}_2} \delta_{\mathbf{k}_1 + \mathbf{k}_2, 0} \right. \\ & + \sum'_{\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3} a_{\mathbf{k}_1 \mathbf{k}_2 \mathbf{k}_3}^\mp(m) m_{\mathbf{k}_1} m_{\mathbf{k}_2} m_{\mathbf{k}_3} \delta_{\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3, 0} \\ & \left. + \dots \right]. \end{aligned} \quad (45)$$

The primes attached to the summation signs indicate that terms with  $\mathbf{k}_i=(0,0)$  are excluded. The coefficients  $a^\pm$ ,  $a_{\mathbf{k}_1 \mathbf{k}_2}^\pm$ ,  $\dots$  are polynomials in  $m$ , which are either odd or even under spin inversion and are to be chosen according to the desired symmetry. Rotation and reflection symmetries of the lattice are imposed by equating coefficients of the appropriate monomials in  $m$ .

The results reported in Ref. 6 were obtained using a more complicated version of Eq. (45), namely,

$$\begin{aligned} \psi_T(S) = \psi_B(S) & \left[ a^-(m) + \sum'_{\mathbf{k}_1, \mathbf{k}_2} a_{\mathbf{k}_1 \mathbf{k}_2}^-(m) m_{\mathbf{k}_1} m_{\mathbf{k}_2} \delta_{\mathbf{k}_1 + \mathbf{k}_2, 0} \right. \\ & \left. + \sum'_{\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3} a_{\mathbf{k}_1 \mathbf{k}_2 \mathbf{k}_3}^+(m) m_{\mathbf{k}_1} m_{\mathbf{k}_2} m_{\mathbf{k}_3} \delta_{\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3, 0} + \dots \right] \\ & \times \left[ a^+(m) + \sum'_{\mathbf{k}_1, \mathbf{k}_2} a_{\mathbf{k}_1 \mathbf{k}_2}^+(m) m_{\mathbf{k}_1} m_{\mathbf{k}_2} \delta_{\mathbf{k}_1 + \mathbf{k}_2, 0} \right. \\ & + \sum'_{\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3} a_{\mathbf{k}_1 \mathbf{k}_2 \mathbf{k}_3}^-(m) m_{\mathbf{k}_1} m_{\mathbf{k}_2} m_{\mathbf{k}_3} \delta_{\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3, 0} \\ & \left. + \dots \right]. \end{aligned} \quad (46)$$

In those calculations also the coupling constant appearing in the Boltzmann factor was treated as a variational parameter, but it turned out that the optimal value of this parameter was indistinguishable from the critical coupling. It does not seem that the more complicated form of expression (46) resulted in a major improvement, but we did not perform a systematic comparison of these trial vectors.

The coefficients in the trial vector are treated as variational parameters. As in all nonlinear fitting problems it is important to use parameters parsimoniously, and to do so one has to establish a hierarchy among these parameters. The scheme we used was to iterate the following step: (a) systematically add terms of increasing degree in  $m$ ; (b) when this saturates, increase the degree of terms with products of  $m_{\mathbf{k}}$  with  $m_{\mathbf{k}} \neq (0,0)$ .

The effectivity of this variational approach using low-momentum Fourier components, as described here, becomes apparent when one compares the variational eigenvalues with the exact numerical ones. For instance, the difference in the case of the second eigenvalue of the  $L=5$  nearest-neighbor model was only  $2 \times 10^{-7}$ .

## V. NUMERICAL RESULTS

### A. Exact eigenvectors for small systems

The full, symmetric Markov matrix  $\hat{P}$  for an  $L \times L$  Ising model is a  $2^{L^2} \times 2^{L^2}$  matrix, so that exact numerical calcula-

tions are possible only for very small systems; see, e.g., results for  $L \leq 4$  in Ref. 18. In the present work, we performed such exact computations for systems up to  $L = 5$ . In order to restrict the numerical task, we chose representations of  $\hat{P}$  in subspaces with the appropriate symmetries. Two distinct symmetries were chosen, both of which impose invariance of the eigenvectors of  $\hat{P}$  with respect to geometric translation, rotation, and mirror inversion. The vectors were chosen to be either even or odd under spin inversion. This reduced by almost a factor 400 the dimensionality of  $\hat{P}$ . In this way, the computation of a restricted set of eigenvectors became feasible for the resulting matrices of order 86056 for the  $L = 5$  cases. For the diagonalization we made use of sparse-matrix methods and the conjugate-gradient method (see, e.g., Refs. 19 and 20) which computes the eigenvector with the largest eigenvalue. Subsequent orthogonalization with respect to this eigenvector yields the eigenvector with the second largest eigenvalue, and further eigenvectors can be obtained similarly. Thus we obtained exact numerical solutions for six eigenvalues  $\lambda_{L_i}$  and their corresponding eigenvectors  $\psi_i(S)$  ( $i=0, \dots, 5$ ) of the eigenvalue equation

$$\sum_{S'} \hat{P}(S, S') \psi_i(S') = \lambda_{L_i} \psi_i(S). \quad (47)$$

The largest eigenvalue  $\lambda_{L_0}$  is equal to 1, in accordance with the conservation of probability; its corresponding eigenvector satisfies  $\psi_0(S) = \psi_B(S)$ , as follows from detailed balance. It is even with respect to spin inversion:  $\psi_0(S) = \psi_0(-S)$  where  $-S$  is obtained from  $S$  by inverting all spins. For all system sizes and models included here, we observed that the six leading eigenvectors, ordered according to magnitude of their eigenvalues, alternate between the odd and even subspaces: the first eigenvector is even, the second one is odd, the third one is even subspace, and so on, with the caveat that for  $L = 2$ , e.g., the odd subspace contains only two independent states. As we discussed above, the resulting eigenvectors provide useful information on how to construct trial vectors; moreover, knowledge of accurate eigenvalues for  $L \leq 5$  provided an powerful test of the Monte Carlo method, the results of which are presented in the following subsection.

### B. Monte Carlo calculations

All simulations took place at the respective critical points of the models considered. This point is known exactly in the case of the nearest-neighbor model [ $K_c = \ln(1 + \sqrt{2})/2$ ], and was determined numerically<sup>12</sup> for the other two models:  $K_c = 0.190\,192\,680\,7(2)$  for the equivalent-neighbor model and  $K_c = 0.697\,220\,7(2)$  for the model with antiferromagnetic next-nearest-neighbor interactions. The finite-size scaling analysis presented in Ref. 12 showed that, to the extent they are compatible with the numerical results, deviations from Ising universal behavior are extremely small. The raw simulation data used in this current paper include the data on which were based the numerical results for the largest relaxation time of the nearest-neighbor model, reported in Ref. 6. The latter results were obtained from  $8 \times 10^8$  Monte Carlo samples for systems with finite sizes up to  $L = 15$ . The trial vector used for these computations consisted was of the form

given in expression (46) and used up to 36 variational parameters. Also included in the present analysis are the simulations reported in Ref. 7, which contained  $1.2 \times 10^8$  Monte Carlo samples for all three models with system sizes up to  $L = 20$ .

In addition, new simulations for each of the three models were performed, with a length of  $2 \times 10^8$  Monte Carlo samples for system sizes up to  $L = 20$  and of  $1.6 \times 10^8$  Monte Carlo samples for system size  $L = 21$ . These new simulations used up to 89 variational parameters in the trial functions for each eigenvector of each model.

In order to suppress biases due to deviations of randomness, we made use of a random number generator which combines two different binary shift registers such as described and discussed in Ref. 21.

The required Fourier components of the spatial magnetization distribution were sampled at intervals of one sweep for the smallest systems up to about 15 sweeps for the largest ones. The Monte Carlo calculation of the autocorrelation times (actually the eigenvalues of the Markov matrix) was performed for each run as a whole as well as separately for a number of up to 1024 blocks into which the run was split. This blocking procedure enabled us to estimate the statistical errors. Furthermore, the calculation of the eigenvalues according to  $\hat{P}(t) \hat{N}(t)^{-1} = \hat{\Lambda}^{(t)}$  still depends on the time displacements  $t$  [see Eqs. (39) and (41)]. The calculation of  $\hat{\Lambda}^{(t)}$  was performed for time displacements  $tL^{-2} = 0, 1, 2, \dots$  up to 10 or 20 of the above-mentioned intervals. For small  $t$  these eigenvalue estimates reflect variational bias due to the residual contributions of relaxation modes decaying faster than the mode for which the trial vector was constructed. If the relaxation times of these faster modes are considerably shorter than that of the mode under investigation, one can clearly see a fast convergence of the eigenvalue estimate as a function of  $t$ . Convergence, however, occurs to a level that is only approximately constant because of the correlated statistical noise whose effect still depends on  $t$ . With increasing  $t$ , one can also observe that the statistical errors increase. The latter effect, which is as slow as the pertinent relaxation mode, occurs when the autocorrelations of the Monte Carlo sample are decreasing significantly with  $t$ . This situation was indeed observed for the largest eigenvalues; the data converged well with  $t$  before the coherence of the sampled data was lost. It was thus rather simple to select a ‘‘best estimate’’ of those eigenvalues. However, the situation for the smaller eigenvalues investigated here was much more difficult, because the relative differences between subsequent autocorrelation times are much smaller. The numerical results for the eigenvalues are listed in Ref. 22

### C. Determination of the dynamic exponent

In two-dimensional Ising models, finite-size corrections are known that decay with finite size as  $L^{-2}$ , and integral powers thereof may also be expected. In the absence of information on possible additional finite-size corrections of a different type that could occur in dynamic phenomena, we try to describe the finite-size data for the various autocorrelation times, as given in Eq. (11), by the formula

TABLE I. Best estimates for the dynamic exponent  $z$  for five relaxation modes in three Ising-like models. These results were selected from a much larger set of least-squares fits, obtained for different choices of the minimum system size and of the number of corrections taken into account (see Ref. 22). The error estimate in the last decimal place of each entry is listed in parentheses, and is taken to be two standard deviations in the best fit.

Mode	$\beta = -1/4$	$\beta = 0$	$\beta = 1$
1	2.164 (3)	2.1660 (10)	2.1667 (5)
2	2.166 (3)	2.167 (1)	2.167 (1)
3	2.164 (5)	2.170 (2)	2.167 (1)
4	2.17 (1)	2.162 (4)	2.170 (8)
5	2.15 (2)	2.17 (1)	2.17 (1)

$$\tau_L \approx L^z \sum_{k=0}^{n_c} \alpha_k L^{-2k}. \quad (48)$$

Here  $z$  is the dynamic exponent,  $\alpha_k$  the finite-size amplitude, and  $n_c$  is the number of correction-to-scaling terms included. Not explicitly shown in this notation is that the autocorrelation times depend on the relaxation mode and the model.

On the basis of Eq. (48), a considerable number of least-squares fits were applied to the numerical results for the autocorrelation times. For each model and relaxation mode, one may vary both  $n_c$ , the number of correction terms, and the low- $L$  cutoff specifying the minimal system size included in the fit. The smaller the number of corrections, the larger the low- $L$  cutoff must be chosen in order to obtain an acceptable squared residual  $\chi^2$ . A selection of fits that display the numerical trends is presented in Ref. 22

The ‘‘best fits’’ were chosen on the basis of the  $\chi^2$  criterion, the dependence on the low- $L$  cutoff, and the mutual consistency of fits with different  $n_c$ . The fits are summarized in Table I. Since the errors are not only of a statistical nature, but also depend on residual bias in the autocorrelation times and subjective choices made in the selection of the best fits, we quote error bars equal to two standard deviations as obtained from statistical considerations only. We believe that these  $2\sigma$  error estimates are conservative in the case of the analysis of the second and third largest eigenvalues of the  $\beta \geq 0$  models. The  $\beta = -1/4$  model was found to be numerically less well behaved: the statistical errors, as well as the corrections to scaling, appear to be larger. Also, the construction of trial vectors was somewhat less successful than in the cases of the  $\beta \geq 0$  models.

The new data are somewhat more accurate than and consistent with our previous work.<sup>7</sup> They are also consistent with the results of Wang and Hu<sup>23</sup> for the slowest relaxation mode of a different set of Ising-like models. They provide a clear confirmation of universality of the dynamic exponent, with regard to relaxation modes as well as models. Our best estimate  $z = 2.1667(5)$  applies to the slowest (odd) relaxation mode of the equivalent-neighbor ( $\beta = 1$ ) model.

This result for  $z$  is consistent with most of the recently published values. This agreement includes the results of Stauffer<sup>24</sup> on damage spreading in the Ising model. (The value listed in Ref. 6 was incorrectly quoted.) It is slightly larger than the value 2.14 derived by Alexandrowicz<sup>25</sup> on the basis of a scaling argument.

TABLE II. Comparison between fits to the autocorrelation times, with and without a logarithmic finite-size dependence. These fits apply to the slowest relaxation mode. The first column shows the minimum system size included, the second the number of correction terms included. The fourth column displays the squared residual  $\chi_z^2$  obtained when the dynamic exponent  $z$  was left free and the amplitude  $b$  of the logarithm was fixed at zero. The fifth column shows the squared residual  $\chi_b^2$  when  $z$  was fixed at value 2, while  $b$  was left free. The sixth column lists the number of degrees of freedom of the fit for comparison.

$L \geq$	Model	$n_c$	$\chi_z^2$	$\chi_b^2$	$d_f$
5	$\beta = 0$	1	239.	274.	76
6	$\beta = 0$	1	98.	187.	72
7	$\beta = 0$	1	83.5	127.	67
8	$\beta = 0$	1	68.7	87.9	62
9	$\beta = 0$	1	63.8	71.9	57
10	$\beta = 0$	1	55.9	57.5	52
8	$\beta = 1$	1	56.0	71.1	51
9	$\beta = 1$	1	50.1	57.4	47
10	$\beta = 1$	1	49.4	49.7	43
4	$\beta = 0$	2	89.2	341.	76
5	$\beta = 0$	2	85.0	136.	75
6	$\beta = 0$	2	83.3	97.2	71

Next we address the question whether the finite-size divergence of the autocorrelation times at criticality can be described by a dynamic exponent  $z = 2$  when a logarithmic factor is included. This possibility was suggested by Domany,<sup>26</sup> and pursued by Swendsen<sup>27</sup> and Stauffer<sup>28</sup>, who used very large lattices and found that this possibility seems inconsistent with one way of simulation, but not with a different one. Further references concerning this question are given in Ref. 29.

Although the present work is restricted to very small system sizes, the data are relatively accurate. Thus we tried to fit the following form to the finite-size data for the slowest relaxation mode:

$$\tau_L \approx L^z (1 + b \ln L) \left( \sum_{k=0}^{n_c} \alpha_k L^{-2k} \right). \quad (49)$$

Fixing  $z = 2$  and taking  $b$  as a variable parameter, we found that this form could well describe the data for large enough  $n_c$ . The quality, as determined by the  $\chi^2$  criterion, of a number of such fits is shown in Table II. For comparison we include fits in which  $z$  is a variable parameter without a logarithmic correction.

The results in Table II indicate that the fits with a variable exponent are usually better than those which include a logarithmic term; i.e., the residual  $\chi^2$  decreases faster when the low- $L$  cutoff is increased. This is especially apparent for small  $n_c$  and for the  $\beta = 0$  and  $\beta = 1$  models, where the statistical accuracy is optimal.

Finally we tried a fit according to Eq. (49) with both  $z$  and  $b$  as free parameters. The resolution of both parameters simultaneously is quite hard, and lies near the limit of what can be gleaned from the present data. For the nearest-neighbor model we find, using system sizes  $L = 8$  and larger,

TABLE III. Best estimates for the finite-size amplitudes of five relaxation modes in three Ising-like models. These results were selected from a much larger set of least-squares fits obtained for different choices of the minimum system size and of the number of corrections taken into account (see Ref. 22). The error estimate in the last decimal place of each entry is listed in parentheses, and is taken to be two standard deviations of the best fit. The amplitudes can be written as the product of mode-dependent and model-dependent constants (see text); the difference in the last decimal place between the amplitudes and this product is shown between square brackets.

Mode/model	$\kappa=1$			$\kappa=2$			$\kappa=3$		
1 (odd)	6.763	(6)	[-10]	4.4089	(13)	[8]	2.8312	(5)	[0]
2 (even)	0.2516	(2)	[2]	0.16364	(5)	[0]	0.10510	(2)	[0]
3 (odd)	0.1188	(1)	[1]	0.07727	(3)	[0]	0.04963	(1)	[0]
4 (even)	0.07195	(7)	[3]	0.04677	(4)	[-4]	0.03008	(3)	[2]
5 (odd)	0.0466	(1)	[-1]	0.03041	(3)	[-1]	0.01956	(2)	[2]

and  $n_c=1$ , that  $z=2.13\pm 0.07$  and  $b=0.05\pm 0.09$ , which again fails to support the presence of a logarithmic term.

#### D. Universality of finite-size amplitudes

In order to determine the leading amplitudes  $\alpha_0$  more precisely, we repeated the fits as used for the determination of the dynamic exponent, but with the value of the latter fixed at  $z=13/6$ . We note that the combined results for  $z$  are consistent with this fraction. A considerable number of fits were made, and ‘‘best estimates’’ of the amplitudes are presented in Table III.

As mentioned in Ref. 7, according to a modest generalization of accepted ideas on universality, the finite-size amplitudes of the autocorrelation times should satisfy  $\alpha_0 = A_i m_\kappa$ , where the  $m_\kappa$  are nonuniversal, model-dependent constants; the subscript  $\kappa$  refers to the specific model. We use the notation  $\kappa=2+\text{sgn}(\beta)$  so that  $\kappa=1$  refers to the model with ferromagnetic nearest-neighbor and antiferromagnetic next-nearest-neighbor couplings,  $\kappa=2$  denotes the nearest-neighbor model, and  $\kappa=3$  refers to the equivalent-neighbor model. The  $A_i$ ,  $i=1, \dots, 5$ , are mode-dependent constants, whose ratios are universal. Since only the product matters, we are free to choose an arbitrary value for one of these constants. We chose to fix  $A_1=1$ , so that all  $A_i$  should become universal constants.

The remaining constants  $A_i$  ( $i=2, \dots, 5$ ) and  $m_\kappa$  were fitted accordingly to the amplitudes listed in Table III. The result of this least-squares fit is  $m_1=6.773\pm 0.003$ ,  $m_2=4.4081\pm 0.0009$ ,  $m_3=2.8312\pm 0.0005$ ,  $A_2=0.037\ 123\pm 0.000\ 008$ ,  $A_3=0.017\ 530\pm 0.000\ 004$ ,  $A_4=0.010\ 618\pm 0.000\ 006$ , and  $A_5=0.006\ 901\pm 0.000\ 005$ . This fit has eight degrees of freedom and  $\chi^2=9.0$ , in good agreement with the assumptions of dynamic universality, and suggesting that our  $2\sigma$  error estimates are not unrealistic. The differences between our amplitude estimates and the fitted product  $A_i m_\kappa$  are included in Table III, in units of the last decimal place listed.

## VI. DISCUSSION

The analysis presented above produces an apparently highly accurate estimate of the dynamic critical exponent  $z=2.1667\pm 0.0005$  for the case of the equivalent-neighbor model, where the statistical accuracy and the convergence

are best. However, even in this case it is not possible to rule out a divergence of the relaxation time of the form  $L^2(1+b\ln L)$ . Nevertheless, our results viewed in their totality make this behavior rather unlikely. First, the assumption of this logarithmic form yields fits that converge less rapidly, as mentioned above. Second, one would have to have  $b\approx 1/6$  universally, independent of model and relaxation mode, since our results for the range of system size we studied are consistent with a divergence of the form  $\tau_L \propto L^z \propto L^{13/6} \approx L^2(1+1/6\ln L)$ . Universality of the amplitude of a logarithmic correction would be quite unusual and does not fit into any theoretical framework of which we are aware.

Some open questions remain regarding the optimized variational vectors to which this computation owes its accuracy. Variational basis vectors of the general form given in Eq. (45) are special in the sense that all parameters enter linearly. The method outlined here does not require this feature and in fact it was not present in previous computations, reported in Ref. 6.

For most of the results presented here we used trial vectors with linear parameters optimized by minimization of the variance of the configurational eigenvalues. As an apparently equivalent alternative, the full set of symmetrized monomials in the Fourier coefficients of the spin configuration could be chosen as basis vectors rather than the linear combinations defined in Eq. (45). With this choice, the basis vectors would not have contained any parameters, but employing this bigger truncated basis, the same linear parameters would have been reintroduced by computing the matrix elements  $\hat{N}_{ij}$  and  $\hat{P}_{ij}$  and solving the generalized eigenvalue problem defined by Eq. (25). In this way, we would have obtained the coefficients for which the Rayleigh quotient is stationary, at least if the summation over configurations could have been done exactly. Proceeding in this way, we could have altogether skipped the optimization scheme based on minimization of the variance of the configurational eigenvalues [cf. Eq. (18) and Sec. II B].

The obvious question is what is accomplished by the nonlinear minimization of the variance of the configurational eigenvalues. We do not yet have a convincing answer to this question. On the one hand, it is not difficult to show that the zero-variance principle holds for individual eigenstates. That is, if an eigenvector can be represented exactly as a linear combination of the basis vectors, the variational ( $t=0$ ) case will

already produce the exact result even if the eigenvalue problem contains eigenvectors that cannot be represented exactly in the truncated basis. This is in fact precisely what happens for the dominant even eigenvector: this vector is represented exactly even in the truncated basis we use, and indeed its eigenvalue is reproduced exactly. On the other hand, our tentative numerical experiments show that the optimization method produces more accurate results, which is not surprising when one considers that the optimized basis functions give rise to much smaller truncated basis sets. This in turn yields a generalized eigenvalue problem involving much smaller matrices that are numerically and statistically much more robust.

A related problem with a large basis set is that the matrix  $\hat{N}$  in the generalized eigenvalue problem of Eq. (25) becomes numerically singular. In fact, the only way in which we were able to obtain meaningful results at all is by performing the inversion in the usual regularized fashion as follows. Use the fact that  $\hat{N}$  is symmetric and non-negative definite to write it in the form

$$\hat{N} = W \text{diag}(\mu_1^2, \dots, \mu_n^2) W^\dagger. \quad (50)$$

Then define a regularized inverse of  $\hat{N}^{1/2}$  as follows:

$$\bar{N}^{-1/2} = W \text{diag}(\bar{\mu}_1^{-1}, \dots, \bar{\mu}_n^{-1}) W^\dagger, \quad (51)$$

where  $\bar{\mu}_i^{-1} = \mu_i^{-1}$  if  $\mu_i$  exceeds a suitable chosen threshold, e.g., the square root of the machine accuracy, and  $\bar{\mu}_i^{-1} = 0$  otherwise. The nonvanishing eigenvalues of  $\bar{N}^{-1/2} \hat{P} \bar{N}^{-1/2}$  then yield a subset of the eigenvalues of  $\hat{P}$  that are least affected by the numerical singularity of  $\hat{N}$ .

Although further modifications of the computational procedures may lead to additional improvements of our technique, the numerical results obtained thus far are already quite promising and the question arises what further applications are obvious in the field of dynamics of Monte Carlo methods.

For instance, it seems well possible to apply the present techniques in three dimensions and to spin-conserving Kawasaki dynamics<sup>1</sup> although it is clear that the construction of trial vectors will have to be modified. In the same context, we note that direct application of the method used in this paper to the dynamics of cluster algorithms<sup>3-5</sup> is frustrated by the requirement that one should be able to compute  $\langle S | \hat{P} | \psi_T \rangle$  numerically exactly. An additional problem for such dynamics from the perspective of our approach is that the concept *correlation time* has to be handled carefully in this context.

Let us demonstrate this point by means of the following thought experiment: the application of the Wolff algorithm to the ferromagnetic, critical Ising model. As usual, we define the autocorrelation times in terms of the eigenvalues of

the stochastic matrix. In order to enable a comparison with other types of dynamics, we choose our unit of time as  $L^{2d-2y_h}$  Wolff steps ( $L$  is the linear system size,  $d$  the dimensionality, and  $y_h$  the magnetic renormalization exponent). Since the average Wolff cluster consists of a number of sites proportional to  $L^{2y_h-d}$ , this choice guarantees that, under equilibrium conditions, an average number of order  $L^d$  spins is processed per unit of time.

Because of the efficiency of the Wolff algorithm, only a few units of time are needed to generate an independent spin configuration under practical circumstances. However, if the fully ordered antiferromagnetic state is chosen as the initial spin configuration, a number of Wolff steps of order  $L^d$  is required to remove the antiferromagnetic order; i.e., its relaxation to equilibrium is anomalously slow. A less extreme but related phenomenon is observed under practical Wolff simulation conditions at equilibrium: from time to time large critical fluctuations occur that bring the system into a state of relatively large disorder and small magnetization. These configurations are relatively long lived. In the time autocorrelation function of the magnetization, this phenomenon translates into a slower-than-exponential decay,<sup>30</sup> at least on the numerically accessible time scale. In the language of Eq. (10) such a situation follows if one assumes the existence of anomalously large autocorrelation times  $\tau_{Li}$  associated with anomalously small amplitudes  $c_i$ . Under these circumstances we cannot exclude the possibility that the longest relaxation time following from the Markov matrix for the Wolff simulation of a finite system corresponds with an extremely unlikely deviation from equilibrium. Since this kind of fluctuations may have too low a probability to be of practical significance, these considerations suggest the possibility that the time needed to generate an ‘‘independent configuration’’ is not simply related to the second largest eigenvalue of the Markov matrix, but rather to some intricate average, possibly involving the complete spectrum.

## ACKNOWLEDGMENTS

It is our pleasure to acknowledge stimulating discussions with Bob Swendsen. This research was supported by the U.S. National Science Foundation (NSF) through Grants DMR-9725080. This research was conducted in part using the resources of the Cornell Theory Center, which received major funding from the NSF and New York State, with additional support from the Advanced Research Projects Agency (ARPA), the National Center for Research Resources at the National Institutes of Health (NIH), the IBM Corporation, and other members of the center’s Corporate Research Institute. This research is supported in part by the Dutch FOM foundation (‘‘Stichting voor Fundamenteel Onderzoek der Materie’’) which is financially supported by the NWO (‘‘Nederlandse Organisatie voor Wetenschappelijk Onderzoek’’).

- <sup>1</sup>K. Kawasaki, in *Phase Transitions and Critical Phenomena*, edited by C. Domb and M.S. Green (Academic, New York, 1972), Vol. 2.
- <sup>2</sup>R. J. Glauber, *J. Math. Phys.* **4**, 294 (1963).
- <sup>3</sup>R. H. Swendsen and J. S. Wang, *Phys. Rev. Lett.* **58**, 86 (1987).
- <sup>4</sup>U. Wolff, *Phys. Rev. Lett.* **60**, 1461 (1988).
- <sup>5</sup>J. R. Heringa and H. W. J. Blöte, *Phys. Rev. E* **57**, 4976 (1998).
- <sup>6</sup>M. P. Nightingale and H. W. J. Blöte, *Phys. Rev. Lett.* **76**, 4548 (1996).
- <sup>7</sup>M. P. Nightingale and H. W. J. Blöte, *Phys. Rev. Lett.* **80**, 1007 (1997).
- <sup>8</sup>C. J. Umrigar, K. G. Wilson, and J. W. Wilkins, *Phys. Rev. Lett.* **60**, 1719 (1988); C. J. Umrigar, K. G. Wilson, and J. W. Wilkins, in *Computer Simulation Studies in Condensed Matter Physics, Recent Developments*, edited by D. P. Landau, K. K. Mon, and H. B. Schüttler, Springer Proceedings in Physics (Springer, Berlin, 1988).
- <sup>9</sup>D. M. Ceperley and B. Bernu, *J. Chem. Phys.* **89**, 6316 (1988). Also see B. Bernu, D. M. Ceperley, and W. A. Lester, Jr., *ibid.* **93**, 552 (1990); W. R. Brown, W. A. Glauser, and W. A. Lester, Jr., *ibid.* **103**, 9721 (1995).
- <sup>10</sup>V. Privman and M. E. Fisher, *Phys. Rev. B* **30**, 322 (1984).
- <sup>11</sup>H. W. J. Blöte and M. P. Nightingale, *Physica A* **134**, 274 (1985).
- <sup>12</sup>M. P. Nightingale and H. W. J. Blöte, *Physica A* **251**, 211 (1998).
- <sup>13</sup>M. P. Nightingale and H. W. J. Blöte, *J. Phys. A* **15**, L33 (1982).
- <sup>14</sup>M. P. Nightingale, in *Computer Simulation Studies in Condensed Matter Physics*, edited by D. P. Landau, K. K. Mon, and H. B. Schüttler, Springer Proceedings in Physics (Springer, Berlin, 1997).
- <sup>15</sup>M. P. Nightingale and C. J. Umrigar, in *Recent Advances in Quantum Monte Carlo Methods*, edited by W. A. Lester, Jr. (World Scientific, Singapore, 1997).
- <sup>16</sup>J. K. L. MacDonald, *Phys. Rev.* **43**, 830 (1933).
- <sup>17</sup>J. H. Wilkinson, *The Algebraic Eigenvalue Problem* (Clarendon Press, Oxford, 1965), p. 103; G. H. Golub and C. F. van Loan, *Matrix Computations* (The John Hopkins University Press, Baltimore, 1989), p. 411.
- <sup>18</sup>M. P. Nightingale and H. W. J. Blöte, *Physica A* **104**, 352 (1980).
- <sup>19</sup>M. P. Nightingale, *Proc. K. Ned. Akad. Wet., Ser. B: Phys. Sci.* **82**, 235 (1979).
- <sup>20</sup>M. P. Nightingale, V. S. Viswanath, and G. Müller, *Phys. Rev. B* **48**, 7696 (1993).
- <sup>21</sup>H. W. J. Blöte, E. Luijten, and J. R. Heringa, *J. Phys. A* **28**, 6289 (1995).
- <sup>22</sup>See EPAPS Document No. E-PRBMDO-62-042026 for appendices. This document may be retrieved via the EPAPS homepage (<http://www.aip.org/pubservs/epaps.html>) or from <ftp.aip.org> in the directory /epaps/. See the EPAPS homepage for more information.
- <sup>23</sup>F.-G. Wang and C.-K. Hu, *Phys. Rev. E* **56**, 2310 (1997).
- <sup>24</sup>D. Stauffer, *J. Phys. A* **26**, L599 (1993).
- <sup>25</sup>Z. Alexandrowicz, *Physica A* **189**, 148 (1992).
- <sup>26</sup>E. Domany, *Phys. Rev. Lett.* **52**, 871 (1984).
- <sup>27</sup>R. H. Swendsen (private communication).
- <sup>28</sup>D. Stauffer, *Int. J. Mod. Phys. C* **10**, 931 (1999).
- <sup>29</sup>D. Stauffer, *Physica A* **244**, 344 (1997).
- <sup>30</sup>W. Kerler, *Phys. Rev. D* **47**, R1285 (1993).