

Dynamic Critical Behavior of the Swendsen–Wang Algorithm: The Two-Dimensional Three-State Potts Model Revisited

Jesús Salas¹ and Alan D. Sokal¹

Received May 20, 1996; final February 11, 1997

We have performed a high-precision Monte Carlo study of the dynamic critical behavior of the Swendsen–Wang algorithm for the two-dimensional three-state Potts model. We find that the Li–Sokal bound ($\tau_{\text{int},\sigma} \geq \text{const} \times C_H$) is almost but not quite sharp. The ratio $\tau_{\text{int},\sigma}/C_H$ seems to diverge either as a small power (≈ 0.08) or as a logarithm.

KEY WORDS: Potts model; Swendsen–Wang algorithm; cluster algorithm; Monte Carlo; Li–Sokal bound; dynamic critical phenomena.

1. INTRODUCTION

Monte Carlo (MC) simulations^(1–4) have become a standard and powerful tool for gaining new insights into statistical-mechanical systems and lattice field theories. However, their practical success is severely limited by critical slowing-down: the autocorrelation time τ —that is, roughly speaking, the time needed to produce one “statistically independent” configuration—diverges near a critical point. More precisely, for a finite system of linear size L at criticality, we expect a behavior $\tau \sim L^z$ for large L . The power z is a *dynamic critical exponent*, and it depends on both the system and the algorithm.

Single-site MC algorithms (such as single-site Metropolis or heat bath) have a dynamic critical exponent $z \geq 2$. This makes it very hard to get high-precision data very close to the critical point on large lattices.

In some cases, a much better dynamical behavior is obtained by allowing nonlocal moves, such as cluster flips.² The Swendsen–Wang (SW)

¹ Department of Physics, New York University, New York, New York 10003; e-mail: salas@mafalda.physics.nyu.edu, sokal@nyu.edu. Current address for Salas: Departamento de Física Teórica, Facultad de Ciencias, Universidad de Zaragoza, Zaragoza 50009, Spain; e-mail: jesus@jupiter.unizar.es.

² See refs. 4–6 for reviews of collective-mode Monte Carlo methods.

cluster algorithm⁽⁷⁾ for the q -state ferromagnetic Potts model achieves a significant reduction in z compared to the local algorithms: one has z between 0 and ≈ 1 , where the exact value depends on q and on the dimensionality of the lattice.⁽⁶⁾ The most favorable case is the two-dimensional (2D) Ising model ($q=2$), in which $z \lesssim 0.3$ (see below).⁽⁷⁻¹⁰⁾ In other cases, the performance of the SW algorithm is less impressive (though still quite good): e.g., $z = 0.55 \pm 0.03$ for the 2D three-state Potts model,⁽¹¹⁾ $z \approx 1$ for the 2D four-state Potts model,⁽¹¹⁻¹⁴⁾ and $z \approx 1$ for the 4D Ising model.^(15,16) Clearly, we would like to understand why this algorithm works so well in some cases and not in others. We hope in this way to obtain new insights into the dynamics of nonlocal Monte Carlo algorithms, with the ultimate aim of devising new and more efficient algorithms.

There is at present no adequate theory for predicting the dynamic critical behavior of an SW-type algorithm. However, there is one rigorous lower bound on z due to Li and Sokal.⁽¹¹⁾ The autocorrelation times of the standard (multicluseter) SW algorithm for the ferromagnetic q -state Potts model are bounded below by a multiple of the specific heat:

$$\tau_{\text{int}, \mathcal{N}}, \tau_{\text{int}, \mathcal{E}}, \tau_{\text{exp}} \geq \text{const} \times C_H \quad (1.1)$$

Here \mathcal{N} is the bond density in the SW algorithm, \mathcal{E} is the energy, and C_H is the specific heat; τ_{int} and τ_{exp} denote the integrated and exponential autocorrelation times, respectively.^(4,6) As a result one has

$$z_{\text{int}, \mathcal{N}}, z_{\text{int}, \mathcal{E}}, z_{\text{exp}} \geq \frac{\alpha}{\nu} \quad (1.2)$$

where α and ν are the standard *static* critical exponents. Thus, the SW algorithm for the q -state Potts model cannot *completely* eliminate the critical slowing-down in any model in which the specific heat is divergent at criticality. The bound (1.1)/(1.2) has also been proven to hold⁽¹³⁾ for the direct SW-type algorithm⁽¹⁷⁾ for the Ashkin–Teller (AT) model.^(18,19)

The important question is the following: Is the Li–Sokal bound (1.1)/(1.2) sharp or not? An affirmative answer would imply that we could use the bound to predict the dynamic critical exponent(s) z given only the static critical exponents of the system. There are three possibilities:

(i) The bound (1.1) is *sharp* (i.e., the ratio τ/C_H is bounded), so that (1.2) is an *equality*.

(ii) The bound is *sharp modulo a logarithm* (i.e., $\tau/C_H \sim \log^p L$ for $p > 0$).

(iii) The bound is *not sharp* (i.e., $\tau/C_H \sim L^p$ for $p > 0$), so that (1.2) is a *strict inequality*.

Unfortunately, the empirical situation, even for the simplest cases, is far from clear. Let us review the status of this problem for the 2D Potts models. For the Ising case, the numerical data⁽⁷⁻¹⁰⁾ are consistent with a power-law behavior with $z_{\text{int},\delta}$ ranging from 0.35 ± 0.01 ⁽⁷⁾ to 0.25 ± 0.01 .^(9,10) However, the data are also consistent with a logarithmic behavior $z_{\text{int},\delta} = 0 \times \log$.⁽⁸⁾ In ref. 13 we reanalyzed the high-precision data of Baillie and Coddington^{(9,10),3} and found that the ratio $\tau_{\text{int},\delta}/C_H$ behaves most likely as a pure power law $\sim L^p$ with $p = 0.060 \pm 0.004$ (statistical error only) or as a logarithm $\sim A + B \log L$. This means that the bound (1.1) is either non-sharp by a small power or else it is sharp modulo a logarithm (in the latter case, the leading term would be $\tau_{\text{int},\delta} \sim \log^2 L$). It is extremely difficult to distinguish between these two scenarios with lattice sizes up to only $L = 512$.

The three-state Potts model was first considered under this perspective in ref. 11. The dynamic critical exponent was found to be $z_{\text{int},\delta} = 0.55 \pm 0.03$, which is significantly larger than the exact result $\alpha/\nu = 2/5 = 0.4$.⁽²¹⁾ So, it *seemed* that the bound (1.2) was not sharp at all.

The four-state Potts model is rather peculiar: the naive fit to the data, $z_{\text{int},\delta} = 0.89 \pm 0.05$,⁽¹¹⁾ is *smaller* than the (exactly known) value of $\alpha/\nu = 1$.⁽²²⁾ The explanation of this paradox is that the true leading term in the specific heat has a multiplicative logarithmic correction, $C_H \sim L \log^{-3/2} L$.^(14,23-25) Indeed, a naive power-law fit to the specific heat yielded $\alpha/\nu = 0.75 \pm 0.01$, consistent with the bound (1.2). A high-precision study of this model was carried out in refs. 12 and 13.⁴ Naive power-law fits to the data showed that the bound (1.2) was satisfied: $z_{\text{int},\delta} = 0.876 \pm 0.012$ and $\alpha/\nu = 0.768 \pm 0.009$. On the other hand, the behavior of the ratio $\tau_{\text{int},\delta}/C_H$ was consistent only with two scenarios: a power law $\sim L^p$ with $p = 0.118 \pm 0.012$ (statistical error only) or a logarithmic growth $\sim A + B \log L$. This means that (1.1) fails to be sharp either by a small power or by a logarithm. In conclusion, there are two likely behaviors for the autocorrelation time: either $\tau_{\text{int},\delta} \sim L^q \log^{-3/2} L$ with $q \approx 1.12$ or else $\tau_{\text{int},\delta} \sim L \log^{-1/2} L$. In both cases, we find multiplicative logarithmic corrections to the autocorrelation time, which make the numerical analysis extremely difficult. These two scenarios for the ratio $\tau_{\text{int},\delta}/C_H$ coincide with those obtained for the Ising case.

³ The exact value of the specific heat at finite L was taken from the paper by Ferdinand and Fisher.⁽²⁰⁾

⁴ In refs. 12 and 13 the “embedding” version of an SW-type algorithm for the AT model was used. This algorithm reduces to the standard SW algorithm at the Ising subspace, but not at the four-state Potts subspace. However, it was shown numerically that both algorithms belong (as expected) to the same dynamic universality class at the four-state Potts subspace.

In summary, for the 2D Potts models with $q=2$ and $q=4$, the Li–Sokal bound *might* be either sharp modulo a logarithm or else *not sharp* with a very small power $p \equiv z - \alpha/\nu$ ($0.04 \lesssim p \lesssim 0.12$). Moreover, if we interpolate between these models by following the self-dual (critical) curve of the AT model,^(18,19) we obtain^(12,13) the same two scenarios: the ratio $\tau_{\text{int},\delta}/C_H$ either grows like a power law with a very small power p (which increases slightly as we move from the Ising model to the four-state Potts model along the AT self-dual curve) or else grows like a logarithm. Thus, there is some kind of continuity along the AT self-dual curve for the dynamic critical behavior of the SW algorithm.⁵

These results have led us to reappraise the status of the Li–Sokal bound for the 2D three-state Potts model. In ref. 11 the bound was declared *not sharp* on the basis of numerical evidence suggesting that $p \equiv z - \alpha/\nu = 0.15 \pm 0.03$. However, this value is not much larger than that obtained^(12,13) for the four-state Potts model; this suggests that the data for the three-state model might also be consistent with a logarithm. Indeed, a closer look at the results of ref. 11 concerning the three-state model reveals that the lattices studied were not very large ($L \leq 256$) and the statistics were not very high (the number of iterations for the $L = 256$ lattice was at most $25 \times 10^3 \tau_{\text{int},\delta}$). This motivated us to reconsider this model and extend the results of ref. 11 to larger lattices with higher statistics.

This paper is organized as follows: Section 2 reviews the basics of the Swendsen–Wang algorithm for the Potts models as well as the proof of the Li–Sokal bound for these models. In Section 3 we describe our Monte Carlo simulations. In Section 4 we discuss in detail our methods of statistical data analysis. Finally, in Section 5 we present the analysis of our numerical (static and dynamic) data, culminating in a discussion of the sharpness of the Li–Sokal bound.

2. BASIC SETUP AND NOTATION

2.1. Potts Model and Swendsen–Wang Algorithm

The q -state Potts model assigns to each lattice site i a spin variable σ_i taking values in the set $\{1, 2, \dots, q\}$; these spins interact through the reduced Hamiltonian

$$\mathcal{H}_{\text{Potts}} = -\beta \sum_{\langle ij \rangle} (\delta_{\sigma_i, \sigma_j} - 1) \quad (2.1)$$

⁵ A similar study was carried out in ref. 17 for the single-cluster version of the algorithm, but the lattices were not very large ($L \leq 256$).

where the sum runs over all the nearest-neighbor pairs $\langle ij \rangle$. To simplify the notation we shall henceforth write $\delta_{\sigma_i, \sigma_j} \equiv \delta_{\sigma_b}$ for a bond $b = \langle ij \rangle$. The ferromagnetic case corresponds to $\beta > 0$. The partition function is defined as

$$Z = \sum_{\{\sigma\}} e^{-\mathcal{H}} = \sum_{\{\sigma\}} \exp \left[\beta \sum_b (\delta_{\sigma_b} - 1) \right] \quad (2.2)$$

Finally, the Boltzmann weight of a configuration $\{\sigma\}$ is given by

$$W_{\text{Potts}}(\{\sigma\}) = \frac{1}{Z} \prod_b \exp[\beta(\delta_{\sigma_b} - 1)] = \frac{1}{Z} \prod_b (1 - p + p\delta_{\sigma_b}) \quad (2.3)$$

where $p = 1 - e^{-\beta}$.

The idea behind the Swendsen–Wang algorithm^(4,6,26) is to decompose the Boltzmann weight by introducing new dynamical variables $n_b = 0, 1$ (living on the bonds of the lattice) and to simulate the joint model of old and new variables by alternately updating one set of variables conditional on the other set. The Boltzmann weight of the joint model is

$$W_{\text{joint}}(\{\sigma\}, \{n\}) = \frac{1}{Z} \prod_b [(1 - p) \delta_{n_b, 0} + p \delta_{\sigma_b} \delta_{n_b, 1}] \quad (2.4)$$

The marginal distribution of (2.4) with respect to the spin variables reproduces the Potts-model Boltzmann weight (2.3). The marginal distribution of (2.4) with respect to the bond variables is the Fortuin–Kasteleyn^(27–29) random-cluster model with parameter q :

$$W_{\text{RC}}(\{n\}) = \frac{1}{Z} \left[\prod_{b: n_b=1} p \right] \left[\prod_{b: n_b=0} (1 - p) \right] q^{\mathcal{C}(\{n\})} \quad (2.5)$$

where $\mathcal{C}(\{n\})$ is the number of connected components (including one-site components) in the graph whose edges are the bonds with $n_b = 1$.

We can also consider the conditional probabilities of the joint distribution (2.4). The conditional distribution of the $\{n\}$ given the $\{\sigma\}$ is as follows: Independently for each bond $b = \langle ij \rangle$, one sets $n_b = 0$ when $\sigma_i \neq \sigma_j$ and sets $n_b = 0$ and 1 with probabilities $1 - p$ and p , respectively, when $\sigma_i = \sigma_j$. Finally, the conditional distribution of the $\{\sigma\}$ given the $\{n\}$ is as follows: Independently for each connected cluster, one sets all the spins σ_i in that cluster equal to the same value, chosen with uniform probability from the set $\{1, 2, \dots, q\}$.

The Swendsen–Wang algorithm simulates the joint probability distribution (2.4) by alternately applying the two conditional distributions just

described. That is, we first erase the current $\{n\}$ configuration and generate a new $\{n\}$ configuration from the conditional distribution given $\{\sigma\}$; we then erase the current $\{\sigma\}$ configuration and generate a new $\{\sigma\}$ configuration from the conditional distribution given $\{n\}$.

2.2. Li–Sokal Bound

To prove the Li–Sokal bound we first notice that the transition matrix P_{SW} of the Swendsen–Wang algorithm can be written as a product

$$P_{\text{SW}} = P_{\text{bond}} P_{\text{spin}} \quad (2.6)$$

where P_{bond} (the update of the bond variables) and P_{spin} (the update of the spin variables) are given by the conditional expectation operators $E(\cdot | \{\sigma\})$ and $E(\cdot | \{n\})$, respectively.

The strategy behind the proof is to compute explicitly the autocorrelation function at time lags 0 and 1 for a suitable observable \mathcal{O} (which we will choose to be a “slow mode” of the algorithm). The unnormalized autocorrelation function is defined as

$$C_{\mathcal{O}}(t) \equiv \langle \mathcal{O}(s) \mathcal{O}(s+t) \rangle - \langle \mathcal{O} \rangle^2 \quad (2.7)$$

(where the expectations are taken *in equilibrium*), and the normalized autocorrelation function as

$$\rho_{\mathcal{O}}(t) \equiv \frac{C_{\mathcal{O}}(t)}{C_{\mathcal{O}}(0)} = \frac{C_{\mathcal{O}}(t)}{\text{var}(\mathcal{O})} \quad (2.8)$$

Then, using some general properties of reversible Markov chains, we will deduce lower bounds for the autocorrelation times $\tau_{\text{int},\mathcal{O}}$ and $\tau_{\text{exp},\mathcal{O}}$. These will in turn imply lower bounds on the dynamic critical exponents $z_{\text{int},\mathcal{O}}$ and $z_{\text{exp},\mathcal{O}}$.

For the observable \mathcal{O} we shall use the bond occupation

$$\mathcal{N} \equiv \sum_b n_b \quad (2.9)$$

From the joint Boltzmann weight (2.4) it is easy to compute the following bond expectation values conditional on the spin configuration $\{\sigma\}$:

$$E(n_b | \{\sigma\}) = p \delta_{\sigma_b} \quad (2.10a)$$

$$E(n_b n_{b'} | \{\sigma\}) = \begin{cases} p^2 \delta_{\sigma_b} \delta_{\sigma_{b'}} & \text{for } b \neq b' \\ p \delta_{\sigma_b} & \text{for } b = b' \end{cases} \quad (2.10b)$$

From these equations it is easy to compute the mean values $\langle \mathcal{N} \rangle$ and $\langle \mathcal{N}^2 \rangle$ and thus $C_{\mathcal{N},\mathcal{N}}(0) \equiv \text{var}(\mathcal{N}) \equiv \langle \mathcal{N}^2 \rangle - \langle \mathcal{N} \rangle^2$:

$$\langle \mathcal{N} \rangle = p \langle \mathcal{E} \rangle \quad (2.11a)$$

$$\langle \mathcal{N}^2 \rangle = p^2 \langle \mathcal{E}^2 \rangle + p(1-p) \langle \mathcal{E} \rangle \quad (2.11b)$$

$$C_{\mathcal{N},\mathcal{N}}(0) \equiv \text{var}(\mathcal{N}) = p^2 \text{var}(\mathcal{E}) + p(1-p) \langle \mathcal{E} \rangle \quad (2.11c)$$

where the energy is defined as

$$\mathcal{E} \equiv \sum_b \delta_{\sigma_b} \quad (2.12)$$

The unnormalized autocorrelation function at time lag 1 is given by

$$C_{\mathcal{N},\mathcal{N}}(1) \equiv \langle \mathcal{N}(0) \mathcal{N}(1) \rangle - \langle \mathcal{N} \rangle^2 = \text{var}(P_{\text{bond}} \mathcal{N}) = \text{var}(E(\mathcal{N} | \{\sigma\})) \quad (2.13)$$

Now, $P_{\text{bond}} \mathcal{N}$ is equal to

$$P_{\text{bond}} \mathcal{N} \equiv E(\mathcal{N} | \{\sigma\}) = p \sum_b \delta_{\sigma_b} = p \mathcal{E} \quad (2.14)$$

Therefore,

$$C_{\mathcal{N},\mathcal{N}}(1) = p^2 \text{var}(\mathcal{E}) \quad (2.15)$$

and

$$\rho_{\mathcal{N},\mathcal{N}}(1) \equiv \frac{C_{\mathcal{N},\mathcal{N}}(1)}{C_{\mathcal{N},\mathcal{N}}(0)} = \frac{p C_H}{p C_H + (1-p) E} = 1 - \frac{(1-p) E}{p C_H + (1-p) E} \quad (2.16)$$

where the energy density E and the specific heat C_H have been defined as

$$E = \frac{1}{V} \langle \mathcal{E} \rangle \quad (2.17)$$

$$C_H = \frac{1}{V} \text{var}(\mathcal{E}) \equiv \frac{1}{V} [\langle \mathcal{E}^2 \rangle - \langle \mathcal{E} \rangle^2] \quad (2.18)$$

and V is the number of lattice sites.

The correlation functions of \mathcal{N} under $P_{\text{SW}} \equiv P_{\text{bond}} P_{\text{spin}}$ are the same as under the positive-semidefinite self-adjoint operator $P'_{\text{SW}} \equiv P_{\text{spin}} P_{\text{bond}} P_{\text{spin}}$. This implies (see, e.g., ref. 4) that we have a spectral representation

$$\rho_{\cdot, t, \cdot}(t) = \int_0^1 \lambda^{|t|} d\nu(\lambda) \quad (2.19)$$

with a positive measure $d\nu$. From this spectral representation we conclude that

$$\rho_{\cdot, t, \cdot}(t) \geq \rho_{\cdot, t, \cdot}(1)^{|t|} \quad (2.20)$$

If we now recall the definitions of the integrated and exponential autocorrelation times

$$\tau_{\text{int.}, \cdot} = \frac{1}{2} \sum_{t=-\infty}^{\infty} \rho_{\cdot, t, \cdot}(t) \quad (2.21)$$

$$\tau_{\text{exp.}, \cdot} = \lim_{|t| \rightarrow \infty} \frac{-|t|}{\log \rho_{\cdot, t, \cdot}(t)} \quad (2.22)$$

we conclude from (2.16)/(2.20) that

$$\tau_{\text{int.}, \cdot} \geq \frac{1}{2} \times \frac{1 + \rho_{\cdot, t, \cdot}(1)}{1 - \rho_{\cdot, t, \cdot}(1)} \geq \text{const} \times C_H \quad (2.23)$$

$$\tau_{\text{exp.}, \cdot} \geq \frac{-1}{\log \rho_{\cdot, t, \cdot}(1)} \geq \text{const} \times C_H \quad (2.24)$$

These are precisely the bounds (1.1). If we take into account the expected behavior close to the critical point of the specific heat and the autocorrelation times, we deduce immediately the bounds (1.2).

Similar bounds hold for the autocorrelation times of the energy \mathcal{E} . This can be seen from the fact that $P_{\text{bond}} \mathcal{N} = p \mathcal{E}$, which implies that

$$C_{\mathcal{E}\mathcal{E}}(t) = p^{-2} C_{\cdot, t, \cdot}(t+1) \quad (2.25)$$

and hence

$$\rho_{\mathcal{E}\mathcal{E}}(t) = \frac{\rho_{\cdot, t, \cdot}(t+1)}{\rho_{\cdot, t, \cdot}(1)} \geq \rho_{\cdot, t, \cdot}(t) \quad (2.26)$$

This equation proves that

$$\tau_{\text{exp.}, \mathcal{E}} = \tau_{\text{exp.}, \cdot} \quad (2.27)$$

and furthermore allows us to bound $\tau_{\text{int},\sigma}$ above and below in terms of $\tau_{\text{int},\tau}$:

$$\tau_{\text{int},\tau} \leq \tau_{\text{int},\sigma} = \frac{\tau_{\text{int},\tau} - 1/2}{\rho_{\tau,\tau}(1)} - \frac{1}{2} \leq \frac{\tau_{\text{int},\tau}}{\rho_{\tau,\tau}(1)} \quad (2.28)$$

If the critical slowing-down is not completely eliminated, we expect the factor $\rho_{\tau,\tau}(1)$ to approach 1 (from below) as $L \rightarrow \infty$. Moreover, irrespective of the presence or absence of critical slowing-down, we expect $\rho_{\tau,\tau}(1)$ to be *bounded away from zero* as $L \rightarrow \infty$. Modulo this very weak hypothesis, (2.28) implies the equality of the dynamic critical exponents for the energy and bond-occupation observables:

$$z_{\text{int},\sigma} = z_{\text{int},\tau} \quad (2.29)$$

Unfortunately, we do not know how to *rigorously* rule out “exotic” behaviors, in which $\rho_{\tau,\tau}(1)$ tends to *zero* as $L \rightarrow \infty$ and yet $\tau_{\text{int},\tau}$ diverges because the autocorrelation function has an extremely long tail.

Finally, we can define a new (“energy-like”) bond observable: the nearest-neighbor connectivity

$$\mathcal{E}' \equiv \sum_b \gamma_b \quad (2.30)$$

where γ_b equals 1 if both ends of the bond b belong to the same cluster, and 0 otherwise. More generally, the connectivity γ_{ij} can be defined for an arbitrary pair i, j of sites:

$$\gamma_{ij}(\{n\}) = \begin{cases} 1 & \text{if } i \text{ is connected to } j \\ 0 & \text{if } i \text{ is not connected to } j \end{cases} \quad (2.31)$$

The interest in \mathcal{E}' stems from the fact that the conditional expectation of δ_{σ_n} given the bond configuration $\{n\}$ is essentially γ_b :

$$E(1 - \delta_{\sigma_n} | \{b\}) = \frac{q-1}{q} (1 - \gamma_b) \quad (2.32)$$

This implies the following relation between the energy and the connectivity densities:

$$\frac{k}{2} - E = \frac{q-1}{q} \left(\frac{k}{2} - E' \right) \quad (2.33)$$

where k is the coordination number of the lattice (i.e., the number of nearest neighbors of any given site) and the connectivity density E' is defined as

$$E' = \frac{1}{V} \langle \mathcal{E}' \rangle \quad (2.34)$$

Furthermore, (2.32) tells us that

$$\begin{aligned} P_{\text{spin}} \left(\frac{kV}{2} - \mathcal{E} \right) &= P_{\text{spin}} \sum_b (1 - \delta_{\sigma_b}) \\ &= \frac{q-1}{q} \sum_b (1 - \gamma_b) = \frac{q-1}{q} \left(\frac{kV}{2} - \mathcal{E}' \right) \end{aligned} \quad (2.35)$$

This implies that

$$C_{\mathcal{E}'\mathcal{E}'}(t) = \left(\frac{q}{q-1} \right)^2 C_{\mathcal{E}\mathcal{E}}(t+1) \quad (2.36)$$

and hence

$$\rho_{\mathcal{E}'\mathcal{E}'}(t) = \frac{\rho_{\mathcal{E}\mathcal{E}}(t+1)}{\rho_{\mathcal{E}\mathcal{E}}(1)} \geq \rho_{\mathcal{E}\mathcal{E}}(t) \quad (2.37)$$

Thus, the bounds (1.1)/(1.2) hold also for the autocorrelation times of \mathcal{E}' . Furthermore, we can obtain bounds analogous to (2.27)–(2.29) for the observable \mathcal{E}' :

$$\tau_{\text{exp},\mathcal{E}} = \tau_{\text{exp},\mathcal{E}'} \quad (2.38)$$

and

$$\tau_{\text{int},\mathcal{E}} \leq \tau_{\text{int},\mathcal{E}'} \leq \frac{\tau_{\text{int},\mathcal{E}}}{\rho_{\mathcal{E}\mathcal{E}}(1)} \quad (2.39)$$

Using again (2.28) and (2.26), we arrive at

$$\tau_{\text{int},\mathcal{E}'} \leq \tau_{\text{int},\mathcal{E}'} \leq \frac{\tau_{\text{int},\mathcal{E}'}}{\rho_{\mathcal{E}'\mathcal{E}'}(2)} \quad (2.40)$$

If $\rho_{\mathcal{E}'\mathcal{E}'}(2)$ is bounded away from zero as $L \rightarrow \infty$, we can conclude that the dynamic critical exponents of \mathcal{N} , \mathcal{E} , and \mathcal{E}' are all equal:

$$z_{\text{int},\mathcal{E}'} = z_{\text{int},\mathcal{E}} = z_{\text{int},\mathcal{N}} \quad (2.41)$$

3. DESCRIPTION OF THE SIMULATIONS

3.1. Observables To Be Measured

We made simulations of the two-dimensional three-state Potts model at criticality,

$$\beta = \beta_c = \log(1 + \sqrt{3}) \approx 1.00505254 \quad (3.1)$$

on a periodic square lattice of linear size L .

We measured five basic observables. Three of them have been already defined: the energy \mathcal{E} , (2.12); the bond occupation \mathcal{N} , (2.9); and the nearest-neighbor connectivity \mathcal{E}' , (2.30). The other two are

$$\mathcal{M}^2 = \left(\sum_x \sigma_x \right)^2 \quad (3.2a)$$

$$= \frac{q}{q-1} \sum_{\alpha=1}^q \left(\sum_x \delta_{\sigma_x, \alpha} \right)^2 - \frac{V^2}{q-1} \quad (3.2b)$$

and

$$\mathcal{F} = \frac{1}{2} \left(\left| \sum_x \sigma_x e^{2\pi i x_1/L} \right|^2 + \left| \sum_x \sigma_x e^{2\pi i x_2/L} \right|^2 \right) \quad (3.3a)$$

$$= \frac{q}{q-1} \times \frac{1}{2} \sum_{\alpha=1}^q \left(\left| \sum_x \delta_{\sigma_x, \alpha} e^{2\pi i x_1/L} \right|^2 + \left| \sum_x \delta_{\sigma_x, \alpha} e^{2\pi i x_2/L} \right|^2 \right) \quad (3.3b)$$

where $\sigma_x \in \mathbb{R}^{q-1}$ is the Potts spin in the hypertetrahedral representation [for $q=3$ this means $\sigma_x = (\cos(2\pi\sigma_x/3), \sin(2\pi\sigma_x/3))$], $V=L^2$ is the number of lattice sites, and (x_1, x_2) are the Cartesian coordinates of point x . The observable \mathcal{F} can be regarded as the square of the Fourier transform of the spin variable at the smallest allowed nonzero momenta [i.e., $(\pm 2\pi/L, 0)$ and $(0, \pm 2\pi/L)$ for the square lattice]. It is normalized to be comparable to its zero-momentum analogue \mathcal{M}^2 .

From these observables we compute the following expectation values: the energy density E , (2.17); the specific heat C_H , (2.18); the connectivity density E' , (2.34); the bond density

$$N = \frac{1}{V} \langle \mathcal{N} \rangle; \quad (3.4)$$

the magnetic susceptibility

$$\chi = \frac{1}{V} \langle \mathcal{M}^2 \rangle; \quad (3.5)$$

the correlation function at momentum $(2\pi/L, 0)$

$$F = \frac{1}{V} \langle \mathcal{F} \rangle; \quad (3.6)$$

and the second-moment correlation length

$$\xi = \frac{(\chi/F - 1)^{1/2}}{2 \sin(\pi/L)} \quad (3.7)$$

This definition of the correlation length is *not* equal to the exponential correlation length ($=1/\text{mass gap}$); but it is expected that both correlation lengths scale in the same way as we approach the critical point.

Remark. As a check we have also computed the mean-square size of the clusters,

$$\mathcal{S}_2 = \sum_{\mathcal{C}} \#(\mathcal{C})^2 \quad (3.8)$$

where the sum is over all the clusters \mathcal{C} of activated bonds (i.e., with $n_b = 1$) and $\#(\mathcal{C})$ is the number of sites of the cluster \mathcal{C} . Using the Fortuin–Kasteleyn identities,^(4,27–29) it is not difficult to show that

$$\langle \mathcal{S}_2 \rangle = \langle \mathcal{M}^2 \rangle \quad (3.9)$$

For each observable \mathcal{O} discussed above we have measured its autocorrelation function $\rho_{\mathcal{O}}(t)$, (2.8); and from this function we have estimated the corresponding integrated autocorrelation time $\tau_{\text{int}, \mathcal{O}}$, (2.21). In Section 4 we explain in detail how we derived estimates of the mean values and the error bars for both static and dynamic quantities.

3.2. Summary of the Simulations

We ran our Monte Carlo program on lattices with L ranging from 4 to 1024. In all cases the initial configuration was random, and we discarded the first 10^5 iterations to allow the system to reach equilibrium; this discard interval is always greater than $10^3 \tau_{\text{int}, \delta}$.⁶ For $4 \leq L \leq 256$ the total run length was approximately $10^6 \tau_{\text{int}, \delta}$; for $L = 512$, it was $2.2 \times 10^5 \tau_{\text{int}, \delta}$; and for $L = 1024$, it was $6.7 \times 10^4 \tau_{\text{int}, \delta}$. In all cases, the statistics are high enough to permit a high accuracy in our estimates of the static (error $\approx 0.1\text{--}0.5\%$) and dynamic (error $\approx 0.5\text{--}2\%$) quantities.

For $4 \leq L \leq 128$ our data were obtained from a single long run. For $L = 256$ we made two independent runs (with different random-number-

⁶ We expect that $\tau_{\text{int}, \delta} / \tau_{\text{exp}, \delta} \approx 0.96$ and $\tau_{\text{exp}, \delta} = \tau_{\text{exp}}$ for this algorithm⁽¹³⁾ (see Section 5.5). So the discard interval is always greater than $10^3 \tau_{\text{exp}}$, which is more than sufficient.

generator seeds); for $L=512$ we made three independent runs; and for $L=1024$ we made four independent runs. In each case we discarded the first 10^5 iterations of each run. The individual runs are all of length $\gtrsim 10^4 \tau_{\text{int},\delta}$, which is long enough to allow a good determination of the dynamic quantities.

To test the program, we computed the static results on the 4×4 lattice to the exact solution (obtained by enumerating all the possible configurations on this lattice). The agreement was excellent (see Table I). In addition, for all lattice sizes we checked the relations (2.11a), (2.33), and (3.9) between the mean values of static observables; the relations (2.25)/(2.36) among the autocorrelation functions $C_{\cdot,1,1}$, $C_{\delta,\delta}$, and $C_{\delta',\delta'}$; and the relation (2.16) between the autocorrelation function $\rho_{\cdot,1,1}(1)$ and the static observables E and C_H . In all cases the agreement was also good.

The CPU time required by our program is approximately $7.2L^2 \mu\text{sec/iteration}$ on an IBM RS-6000/370, and $3.6L^2 \mu\text{sec/iteration}$ on a single processor of an IBM RS/6000 SP2. The total CPU time used in the project was approximately 1.5 years on an IBM RS/600 SP2. The smallest lattices were run on an IBM RS-6000/370 at NYU, while the largest ones ($L \geq 128$) were run on the IBM SP2 cluster at the Cornell Theory Center.

4. STATISTICAL ANALYSIS OF THE MONTE CARLO DATA

In this paper we are aiming at extremely high precision for both static and dynamic quantities; furthermore, we need to disentangle the effects of statistical errors from the effects of systematic errors due to corrections to scaling. For this, it is essential to obtain accurate estimates not only of the static and dynamic quantities of interest, but also of their *error bars*: in this way we will be able (see Section 5) to perform χ^2 tests which provide an objective measure of the goodness of fit in each scaling Ansatz.

In this section we discuss in some detail how we performed the statistical analysis of our raw Monte Carlo data. In particular, we describe how to compute the estimators for the mean value and the variance of both static and dynamic quantities. These methods are based on well-known results of time-series analysis,^(30,31) which we review briefly in Section 4.1.⁷ Then (Section 4.2) we describe an alternative analysis method, based on independent ‘‘bunches,’’ and report the results of detailed cross-checks that confirm (with one slight exception) the validity and reliability of the standard time-series-analysis method.

⁷ A review of time-series-analysis methods as applied to MC simulations can be found in ref. 32, Appendix C, and in ref. 4.

4.1. "Standard" Time-Series-Analysis Method

Let us consider a generic observable \mathcal{O} whose mean is equal to μ_c . Its corresponding unnormalized and normalized autocorrelation functions are denoted by $C_{cc}(t) \equiv \langle \mathcal{O}(0) \mathcal{O}(t) \rangle - \langle \mathcal{O} \rangle^2$ and $\rho_{cc}(t) \equiv C_{cc}(t)/C_{cc}(0)$, respectively. We also define the integrated autocorrelation time

$$\tau_{\text{int},c} = \frac{1}{2} \sum_{t=-\infty}^{\infty} \rho_{cc}(t) \quad (4.1)$$

Given a sequence of n Monte Carlo measurements of the observable \mathcal{O} —call them $\{\mathcal{O}_1, \dots, \mathcal{O}_n\}$ —the natural estimator of the mean μ_c is the sample mean

$$\bar{\mathcal{O}} \equiv \frac{1}{n} \sum_{i=1}^n \mathcal{O}_i \quad (4.2)$$

This estimator is unbiased and has a variance

$$\text{var}(\bar{\mathcal{O}}) = \frac{1}{n^2} \sum_{r,s=1}^n C_{cc}(r-s) \quad (4.3a)$$

$$= \frac{1}{n} \sum_{t=-(n-1)}^{n-1} \left(1 - \frac{|t|}{n}\right) C_{cc}(t) \quad (4.3b)$$

$$\approx \frac{1}{n} 2\tau_{\text{int},c} C_{cc}(0) \quad \text{for } n \gg \tau_{\text{int},c} \quad (4.3c)$$

This means that the variance is a factor $2\tau_{\text{int},c}$ larger than it would be if the measurements were uncorrelated. It is therefore very important to estimate the autocorrelation time $\tau_{\text{int},c}$ in order to ensure a correct determination of the error bar on the (static) quantity μ_c .

The natural estimator for the unnormalized autocorrelation function $C_{cc}(t)$ is

$$\hat{C}_{cc}(t) \equiv \frac{1}{n-|t|} \sum_{i=1}^{n-|t|} (\mathcal{O}_i - \mu_c)(\mathcal{O}_{i+|t|} - \mu_c) \quad (4.4)$$

if the mean μ_c is known, and

$$\hat{\hat{C}}_{cc}(t) \equiv \frac{1}{n-|t|} \sum_{i=1}^{n-|t|} (\mathcal{O}_i - \bar{\mathcal{O}})(\mathcal{O}_{i+|t|} - \bar{\mathcal{O}}) \quad (4.5)$$

if the mean μ_c is unknown. We emphasize that, for each t , the estimators $\hat{C}_{c,c}(t)$ and $\hat{\tilde{C}}_{c,c}(t)$ are *random variables* [in contrast to $C_{c,c}(t)$, which is a *number*]. The estimator $\hat{C}_{c,c}(t)$ is unbiased, and $\hat{\tilde{C}}_{c,c}(t)$ is biased by terms of order $1/n$. The covariance matrices of $\hat{C}_{c,c}$ and $\hat{\tilde{C}}_{c,c}$ are the same to leading order in the large- n limit (i.e., $n \gg \tau_{\text{int},c}$), and we have^(30,31)

$$\begin{aligned} & \text{cov}(\hat{C}_{c,c}(t), \hat{C}_{c,c}(u)) \\ &= \frac{1}{n} \sum_{m=-\infty}^{\infty} [C_{c,c}(m) C_{c,c}(m+u-t) + C_{c,c}(m+u) C_{c,c}(m-t) \\ & \quad + \kappa(t, m, m+u)] + o\left(\frac{1}{n}\right) \end{aligned} \tag{4.6}$$

where $t, u \geq 0$ and κ is the connected four-point autocorrelation function

$$\begin{aligned} \kappa(r, s, t) &\equiv \langle (\mathcal{O}_i - \mu_c)(\mathcal{O}_{i+r} - \mu_c)(\mathcal{O}_{i+s} - \mu_c)(\mathcal{O}_{i+t} - \mu_c) \rangle \\ & \quad - C_{c,c}(r) C_{c,c}(t-s) - C_{c,c}(s) C_{c,c}(t-r) \\ & \quad - C_{c,c}(t) C_{c,c}(s-r) \end{aligned} \tag{4.7}$$

The natural estimator for the normalized autocorrelation function $\rho_{c,c}(t)$ is

$$\hat{\rho}_{c,c}(t) \equiv \hat{C}_{c,c}(t) / \hat{C}_{c,c}(0) \tag{4.8}$$

if the mean μ_c is known, and

$$\hat{\hat{\rho}}_{c,c}(t) \equiv \hat{\tilde{C}}_{c,c}(t) / \hat{\tilde{C}}_{c,c}(0) \tag{4.9}$$

if the mean μ_c is unknown. The estimators $\hat{\rho}_{c,c}(t)$ and $\hat{\hat{\rho}}_{c,c}(t)$ are biased by terms of order $1/n$ as a result of the ratios of random variables in (4.8)/(4.9). The covariance matrices of $\hat{\rho}_{c,c}$ and $\hat{\hat{\rho}}_{c,c}$ are the same to leading order in $1/n$. If the process is Gaussian, this covariance matrix is given in the large- n limit by⁽³¹⁾

$$\begin{aligned} & \text{cov}(\hat{\rho}_{c,c}(t), \hat{\rho}_{c,c}(u)) \\ &= \frac{1}{n} \sum_{m=-\infty}^{\infty} [\rho_{c,c}(m) \rho_{c,c}(m+t-u) + \rho_{c,c}(m+u) \rho_{c,c}(m-t) \\ & \quad + 2\rho_{c,c}(t) \rho_{c,c}(u) \rho_{c,c}^2(m) - 2\rho_{c,c}(t) \rho_{c,c}(m) \rho_{c,c}(m-u) \\ & \quad - 2\rho_{c,c}(u) \rho_{c,c}(m) \rho_{c,c}(m-t)] + o\left(\frac{1}{n}\right) \end{aligned} \tag{4.10}$$

If the process is *not* Gaussian, then there are additional terms proportional to the fourth cumulant $\kappa(m, t, t-u)$; these terms are, like those in (4.10), of order $1/n$. The simplest assumption is to consider the stochastic process to be “not too far from Gaussian” and drop all the terms involving κ . If this assumption is *not* justified, then we are introducing a bias in the estimate of this covariance.

Finally, we shall take the estimator for the integrated autocorrelation time to be⁽³²⁾

$$\hat{\tau}_{\text{int},c} \equiv \frac{1}{2} \sum_{t=-M}^M \hat{\rho}_{cc}(t) \quad (4.11)$$

[or the same thing with $\hat{\rho}_{cc}(t)$], where M is a suitably chosen number. The reason behind the cutoff M is the following: if we were to make the “obvious” choice $M=n+1$, then the resulting estimator would have a variance of order 1 even in the limit $n \rightarrow \infty$; this is because the terms $\hat{\rho}_{cc}(t)$ with large t have errors (of order $1/n$) that *do not* vanish as t grows [cf. (4.10)], and their number is also large ($\sim n$). Taking $M \ll n$ restores the good behavior of the estimator as $n \rightarrow \infty$. The bias introduced by this rectangular cutoff⁸ is given by

$$\text{bias}(\hat{\tau}_{\text{int},c}) = -\frac{1}{2} \sum_{|t|>M} \rho_{cc}(t) + o\left(\frac{1}{n}\right) \quad (4.12)$$

The variance of the estimator $\hat{\tau}_{\text{int},c}$ can be computed from the covariance (4.10); the final result is⁽³²⁾

$$\text{var}(\hat{\tau}_{\text{int},c}) \approx \frac{2(2M+1)}{n} \tau_{\text{int},c}^2 \quad (4.13)$$

where the approximation $\tau_{\text{int},c} \ll M \ll n$ has been made. A good (self-consistent) choice of M is the following⁽³²⁾: let M be the smallest integer such that $M \geq c \hat{\tau}_{\text{int},c}(M)$, where c is a suitable constant. If the normalized autocorrelation function is roughly a pure exponential,⁹ then the choice $c \approx 6$ is reasonable. Indeed, if we take $\rho_{cc}(t) = e^{-t/|\tau|}$ and minimize the mean-square error

$$\text{MSE}(\hat{\tau}_{\text{int},c}) \equiv \text{bias}(\hat{\tau}_{\text{int},c})^2 + \text{var}(\hat{\tau}_{\text{int},c}) \quad (4.14)$$

⁸ We could use more general cutoff functions, but this rectangular cutoff is the most convenient for the present purposes.

⁹ This has been found empirically to be true in the SW algorithm for the 2D Ising⁽³³⁾ and four-state Potts models⁽¹³⁾ and is confirmed here for the three-state Potts model (see Section 5.5).

using (4.12)/(4.13), we find that the optimal window width is

$$M_{\text{opt}} = \frac{\tau}{2} \log \left(\frac{n}{2\tau} \right) - 1 \quad (4.15)$$

For $n/\tau \approx 10^6$ (resp. 10^4), we have $M_{\text{opt}}/\tau \approx 6.56$ (resp. 4.26).

As notes above, we expect the estimator $\hat{\tau}_{\text{int},\epsilon}$ to have a bias of order $\tau_{\text{int},\epsilon}/n$, due to the nonlinearities in (4.8)/(4.9).¹⁰ To make this bias negligible, we need long runs. It has been shown empirically that this procedure works fairly well when $n \gtrsim 10^4 \hat{\tau}_{\text{int},\epsilon}$.⁽⁴⁾

Remarks. 1. The estimation of the error bar for the specific heat is a little bit more complicated. One can obtain $\text{var}(C_H)$ by computing $\text{var}(\mathcal{E})$, $\text{var}(\mathcal{E}^2)$, and $\text{cov}(\mathcal{E}, \mathcal{E}^2)$. This procedure has a numerical drawback: sometimes the covariance matrix for the observables \mathcal{E} and \mathcal{E}^2 is nearly singular; then a small statistical fluctuation can cause the *estimator* of this matrix to be non-positive-definite. We are not aware of any procedure that ensures that the estimator of a covariance matrix is also positive-definite. To overcome this difficulty, we considered the observable $\mathcal{O} = (\mathcal{E} - \mu_\epsilon)^2$, which can be studied using the standard method. As we do not know exactly the value of μ_ϵ , we can use instead the sample mean $\bar{\mathcal{E}}$ (which should be computed first). To leading order in $1/n$, this procedure gives the right error bar for the specific heat.

2. We have a similar problem when computing the error bar for the second-moment correlation length ξ , which is a function of the two quantities χ and F [cf. (3.6)]. In this case we considered the random variable

$$\mathcal{O}' = \frac{\mathcal{M}^2}{\mu_{\mathcal{M}^2}} - \frac{\mathcal{F}}{\mu_{\mathcal{F}}} \quad (4.16)$$

which has automatically a zero mean. The error bar for the second-moment correlation length can be written easily as a function of the susceptibility χ , the quantity F , and the variance of the above-mentioned observable \mathcal{O}' . With this trick we take into account the cross-correlation between \mathcal{M}^2 and \mathcal{F} . This method needs the mean values $\mu_{\mathcal{M}^2}$ and $\mu_{\mathcal{F}}$; in practical situations they are substituted by the corresponding sample means $\bar{\mathcal{M}^2}$ and $\bar{\mathcal{F}}$ (which must be computed first).

This is the standard procedure we have used to analyze each of our Monte Carlo runs. Estimates coming from multiple independent runs

¹⁰ The bias on the estimator $\hat{\tau}_{\text{int},\epsilon}$ also induces a bias on the estimated variance (4.3) of the sample mean $\bar{\mathcal{E}}$. This bias is of order $1/n^2$, i.e., a factor $1/n$ down from the variance (4.3) itself.

(for $L \geq 256$) were merged using the standard formulas for statistically independent data. The results are reported in Table I (static quantities) and Table II (dynamic quantities). The results on $\tau_{\text{int},\sigma}$ for $16 \leq L \leq 256$ can be compared with those reported in ref. 11; the agreement is excellent ($\chi^2 = 6.17$, five degrees of freedom, confidence level 56%).

4.2. Method Based on Independent Bunches

Unfortunately, this standard method does not provide an easy way to compute the error bar for complicated quantities such as ratio $\tau_{\text{int},\sigma}/C_H$, which will play a central role in our analysis of the sharpness of the Li-Sokal bound (see Section 5.3). We can, of course, give an *upper bound* on the actual error bar by using the triangle inequality, but this upper bound will be a significant overestimate of the true value. If we were then to use this overestimate in fits, we would find artificially small values of χ^2 ; as a result, the confidence levels would be artificially high, and useless for distinguishing good from bad fits. (At best we could distinguish *better* versus *worse* fits by looking at the *relative* values of χ^2 .)

This fact motivated us to look for an alternative method to compute the error bars. There is also another advantage in having an alternate method: we can independently check the assumptions and approximations made in the standard procedure.

The second method works as follows. First, we split the whole sample of n MC measurements $\{\mathcal{O}_1, \mathcal{O}_2, \dots, \mathcal{O}_n\}$ into m bunches of equal (or almost equal) length $l \equiv n/m$. For each bunch i ($1 \leq i \leq m$) we can compute the sample mean of the observable \mathcal{O} and an estimate of its integrated autocorrelation times ($\bar{\mathcal{O}}_i$ and $\hat{\tau}_{\text{int},\mathcal{O},i}$, respectively). Indeed, we can also estimate the corresponding variances, but these estimates do *not* play any role in what follows. When computing the autocorrelation functions within each bunch, we used the whole sample mean $\bar{\mathcal{O}}$ as our estimate of $\mu_{\mathcal{O}}$ (instead of the bunch sample mean $\bar{\mathcal{O}}_i$); this trick reduces the bias in the estimates of the autocorrelation functions.

In this way we obtain two sequences of single-bunch estimates $\{\bar{\mathcal{O}}_i\}$ and $\{\hat{\tau}_{\text{int},\mathcal{O},i}\}$. If the bunches are long enough (i.e., $l \gg \tau_{\text{int},\mathcal{O}}$), then the estimates from distinct bunches are almost statistically independent. Thus, we can define our estimates as follows:

$$\hat{\mathcal{O}}' \equiv \frac{1}{m} \sum_{i=1}^m \bar{\mathcal{O}}_i \quad (4.17a)$$

$$\widehat{\text{var}}(\hat{\mathcal{O}}') \equiv \frac{1}{m(m-1)} \sum_{i=1}^m (\bar{\mathcal{O}}_i - \hat{\mathcal{O}}')^2 \quad (4.17b)$$

Table I. Static Data from the MC Simulations at the Critical Point of the Three-State Potts Model^a

L	MCS	χ	C_H	ξ	E
4	Exact	12.204711	1.496719	3.862380	1.710062
4	4.9	12.2093 ± 0.0054	1.4952 ± 0.0021	3.8635 ± 0.0043	1.71051 ± 0.00039
8	6.9	41.3527 ± 0.0204	2.4712 ± 0.0031	7.5419 ± 0.0072	1.65391 ± 0.00026
16	9.9	138.4491 ± 0.0723	3.7162 ± 0.0046	14.9267 ± 0.0134	1.62070 ± 0.00016
32	14.9	462.7014 ± 0.2408	5.2941 ± 0.0064	29.8509 ± 0.0250	1.60223 ± 0.00010
64	19.9	1542.6921 ± 0.8428	7.3839 ± 0.0094	59.6964 ± 0.0501	1.59163 ± 0.00006
128	29.9	5135.9512 ± 2.7481	10.1007 ± 0.0127	119.3401 ± 0.0952	1.58552 ± 0.00003
256	40.8	17082.8221 ± 9.2802	13.7116 ± 0.0176	238.4565 ± 0.1888	1.58202 ± 0.00002
512	12.9	56760.2838 ± 65.0770	18.4752 ± 0.0507	475.9494 ± 0.7779	1.58001 ± 0.00003
1024	5.5	189676.5530 ± 387.3500	24.5281 ± 0.1247	958.8929 ± 2.7875	1.57893 ± 0.00003
∞	Exact				1.577350

^a For each lattice size L , we include the number of performed measurements MCS in units of 10^6 , the susceptibility χ , the specific heat C_H , the second-moment correlation length ξ , and the energy E . The quoted errors correspond to one standard deviation (i.e., confidence level $\approx 68\%$). The first row (“exact”) gives the *exact* results for the 4×4 lattice, and the last row (“ $L = \infty$ ”) gives the *exact* energy in the limit $L \rightarrow \infty$.

Table II. Autocorrelation Times for the Runs Performed at the Critical Point of the Three-State Potts Model^a

L	τ_{int, μ^2}	$\tau_{\text{int}, \uparrow}$	$\tau_{\text{int}, \sigma}$	$\tau_{\text{int}, \sigma'}$
4	4.013 ± 0.018	3.205 ± 0.013	4.023 ± 0.018	4.034 ± 0.018
8	5.983 ± 0.028	5.194 ± 0.023	6.033 ± 0.028	6.080 ± 0.028
16	8.816 ± 0.041	8.084 ± 0.036	9.025 ± 0.043	9.117 ± 0.043
32	12.648 ± 0.057	12.200 ± 0.055	13.280 ± 0.062	13.438 ± 0.063
64	18.101 ± 0.085	18.313 ± 0.086	19.549 ± 0.095	19.769 ± 0.097
128	25.665 ± 0.117	27.094 ± 0.127	28.525 ± 0.137	28.820 ± 0.139
256	35.691 ± 0.164	39.204 ± 0.189	40.824 ± 0.200	41.212 ± 0.203
512	49.775 ± 0.480	56.667 ± 0.583	58.511 ± 0.612	58.974 ± 0.619
1024	68.387 ± 1.183	80.489 ± 1.511	82.488 ± 1.567	83.030 ± 1.583

^a For each lattice size L , we include the integrated autocorrelation times for the squared magnetization (τ_{int, μ^2}), the bond occupation ($\tau_{\text{int}, \uparrow}$), the energy ($\tau_{\text{int}, \sigma}$), and the nearest neighbor connectivity ($\tau_{\text{int}, \sigma'}$).

$$\hat{t}'_{\text{int},\mathcal{E}} \equiv \frac{1}{m} \sum_{i=1}^m \hat{t}_{\text{int},\mathcal{E},i} \quad (4.17c)$$

$$\widehat{\text{var}}(\hat{t}'_{\text{int},\mathcal{E}}) \equiv \frac{1}{m(m-1)} \sum_{i=1}^m (\hat{t}_{\text{int},\mathcal{E},i} - \hat{t}'_{\text{int},\mathcal{E}})^2 \quad (4.17d)$$

The quality of the results depends on the total number of measurements n and the number m of bunches we use. The merging of data coming from different runs is trivial in this case.

For $4 \leq L \leq 256$ we have extremely good statistics ($n \sim 10^6 \tau_{\text{int},\mathcal{E}}$). This allows us to vary m over at least one order of magnitude and thereby to provide a cross-check on the standard time-series analysis. In this discussion it is useful to divide the observables into three categories: linear static, nonlinear static, and dynamic. The first category includes \mathcal{E} , \mathcal{E}' , \mathcal{N} , \mathcal{M}^2 , and \mathcal{F} , whose sample mean values are linear in the raw MC data. The second category includes the specific heat and the second-moment correlation length, whose mean values are nonlinear functions of the raw MC data. Finally, the third category contains all the autocorrelation times, which are also nonlinear functions of the raw data.

For $4 \leq L \leq 256$ we first divided the whole sample into $m = 100$ bunches, each of them with a length $l \sim 10^4 \tau_{\text{int},\mathcal{E}}$. For $4 \leq L \leq 64$ we repeated the analysis using $m = 1000$ bunches, each of them with a length $l \sim 10^3 \tau_{\text{int},\mathcal{E}}$.

For the linear static observables, the “standard” and “bunch” methods always give identical mean values; this is a trivial identity, provided that the bunch lengths are exactly equal. As for the error bars on these observables, we find *unsystematic* discrepancies between the estimates given by the two methods: for $m = 100$ the discrepancies are of order 10%, and for $m = 1000$ they are of order 2%. In order words, the size of these discrepancies is roughly of order $\sim 1/\sqrt{m}$ and the sign is random; this is exactly what one expects on theoretical grounds for the statistical fluctuations in the estimators (4.17b) and (4.17d).

For the nonlinear static observables, the agreement between the mean values coming from the standard and the bunch methods is exact only in the case of the specific heat; this is because we used the same estimator $\bar{\mathcal{E}}$ in both methods. For the correlation length, the mean values show small *systematic* discrepancies between the two methods, of order 0.05–0.1 standard deviations when $m = 100$; the bunch method always gives a larger estimate than the standard method. In absolute value, these discrepancies range from 6×10^{-4} (for $L = 4$) to 2×10^{-2} (for $L = 256$). The same qualitative behavior is found when we repeat the analysis with $m = 1000$ bunches. Now the differences in the mean value of the correlation length

are 0.6–1.6 standard deviations. In absolute value, they range from 8×10^{-3} (for $L=4$) to 4×10^{-2} (for $L=64$), i.e., they are roughly one order of magnitude larger than in the $m=100$ case. Thus, the discrepancies in the mean value of the correlation length are systematic with a size of order $\sim m$; this is exactly what one expects for the mean value of a nonlinear observable, which is afflicted by a bias of order $1/n$ in the standard method versus $1/l = m/n$ in the bunch method. Regarding the error bars, we find *unsystematic* discrepancies of order 10% when $m=100$, and of order 2% when $m=1000$. They behave like the error bars of the static linear observables.

For the autocorrelation-time estimators, we find *systematic* differences between the two methods: the estimates coming from the standard method are consistently smaller than those coming from the bunch method. For $m=100$, these differences are rather small compared to the statistical error bars ($\lesssim 0.5$ standard deviations). When $m=1000$, these discrepancies are much more relevant: their size is roughly two standard deviations. In absolute value, the discrepancies when $m=1000$ are one order of magnitude larger than when $m=100$. Thus, we find a systematic bias of size $\sim m$, again as expected for a quantity which exhibits a bias of order $1/n$ versus $1/l$. The error bars for the autocorrelation times are consistently larger for the standard method than for the bunch method, except for the specific heat, where the behavior is consistently the opposite one. For $m=100$ bunches, these discrepancies are of order 15%; and for $m=1000$ the size remains at the same level. Thus, the discrepancies for the error bars of dynamic observables do not depend much on m ; rather, they are of order ~ 1 .

From the above discussion, we conclude that for the linear static observables the two methods show excellent agreement both for the mean values (trivial equality) and for the error bars (random discrepancies of order $1/\sqrt{m}$). The same holds for the error bars of the nonlinear static observables. This confirms that the standard method is giving accurate estimates of the error bars, at least when the run length n is large enough to provide a good determination of the autocorrelation time (which largely determines the static-quantity error bar). This $1/\sqrt{m}$ dependence also confirms our theoretical prediction that the standard method gives a *more accurate* estimate of the error bars than the bunch method. Indeed, the bunch method can be considered roughly equivalent to employing the standard method *with the window width M taken of order the bunch length l* ; but if $l \gg \tau_{\text{int},c}$ (as it *must* be in the bunch method), this is an *unnecessarily large* window width, and thus leads to unnecessarily large statistical fluctuations.

For the mean values of the nonlinear observables (both static and dynamic), we likewise confirm the validity of the standard method. Once

again the standard method is more reliable than the bunch method, but for a different reason: the bias of order $1/n$ is much smaller than the bias of order $1/l = m/n$. The latter bias becomes particularly serious when the number m of bunches is large (as it *must* be in order to get good estimates of the *error bars!*).

Finally, the least understood piece is the determination of the error bar of the autocorrelation time: our data from the bunch method suggest that the standard method may be making a systematic error of order 15%. Perhaps this systematic error (if indeed it is real) arises from our neglect of the contributions of the fourth-order cumulant κ to the covariance (4.10). This point definitely merits further investigation.

In summary, we think that the bunch method provides a good confirmation of the estimates given by the standard method. We shall hereafter consider the values given by the standard method to be the definitive ones, except for the ratio $\tau_{\text{int},\sigma}/C_H$, where the standard method does not yield any correct error bar. In this latter case we shall use the central value coming from the standard method (which in fact agrees with the bunch-method value to within 0.1–1 %; the discrepancies are of order 0.2 standard deviations), and the error bars coming from the bunch method with 100 bunches (for $4 \leq L \leq 256$), 26 bunches (for $L = 512$), or 55 bunches (for $L = 1024$). We shall also compute the upper bound on the error bar coming from the standard method combined with the triangle inequality. These results for $\tau_{\text{int},\sigma}/C_H$ are shown in Table III.

Table III. Ratio $\tau_{\text{int},\sigma}/C_H$ for the Runs Performed at the Critical Point of the Three-State Potts Model^a

L	$\tau_{\text{int},\sigma}/C_H$
4	$2.6906 \pm (0.0111, \leq 0.0161)$
8	$2.4413 \pm (0.0092, \leq 0.0145)$
16	$2.4285 \pm (0.0090, \leq 0.0145)$
32	$2.5084 \pm (0.0084, \leq 0.0147)$
64	$2.6476 \pm (0.0111, \leq 0.0163)$
128	$2.8241 \pm (0.0111, \leq 0.0171)$
256	$2.9773 \pm (0.0113, \leq 0.0184)$
512	$3.1670 \pm (0.0281, \leq 0.0418)$
1024	$3.3630 \pm (0.0442, \leq 0.0810)$

^a The parentheses give our error-bar estimates. The first number shows the error bar coming from performing the “bunch” method with 100 bunches. The second number is obtained by using the triangle inequality with the numerical results coming from the standard method (see Section 4).

Table IV. Numerical Estimates for the Static and Dynamic Critical Exponents of the Three-State Potts Model and the Exact Results

Exponent	Numerical	Exact
γ/ν	1.73444 ± 0.00043	$26/15$
α/ν	0.4240 ± 0.0030	$2/5$
$\bar{z}_{\text{int.}, \epsilon}$	0.515 ± 0.006	$\geq 2/5$
$\bar{z}_{\text{int.}, \epsilon'}$	0.514 ± 0.006	$\geq 2/5$
$\bar{z}_{\text{int.}, t}$	0.529 ± 0.006	$\geq 2/5$
$\bar{z}_{\text{int.}, \omega^2}$	0.475 ± 0.006	

5. DATA ANALYSIS

For each quantity \mathcal{O} , we carry out a fit to the power-law Ansatz $\mathcal{O} = AL^p$ using the standard weighted least-squares method. As a precaution against corrections to scaling, we impose a lower cutoff $L \geq L_{\min}$ on the data points admitted in the fit, and we study systematically the effects of varying L_{\min} on the estimates A and p and on the χ^2 value. In general, our preferred fit corresponds to the smallest L_{\min} for which the goodness of fit is reasonable (e.g., the confidence level¹¹ is ≥ 10 – 20%), and for which subsequent increases in L_{\min} do not cause the χ^2 to drop vastly more than one unit per degree of freedom.

Our final estimates for static and dynamic critical exponents are collected in Table IV.

5.1. Static Quantities

There are a few exactly known results concerning the 2D three-state Potts model. We know all the critical exponents^(19,21) and in particular the ratios

$$\frac{\gamma}{\nu} = \frac{26}{15} \approx 1.73333 \quad (5.1)$$

$$\frac{\alpha}{\nu} = \frac{2}{5} = 0.4 \quad (5.2)$$

¹¹ “Confidence level” is the probability that χ^2 would exceed the observed value, assuming that the underlying statistical model is correct. An unusually low confidence level (e.g., less than 5%) thus suggests that the underlying statistical model is *incorrect*—the most likely cause of which would be corrections to scaling.

which are the relevant quantities we can directly estimate from our Monte Carlo data. The leading correction-to-scaling exponent is also known^(34,35):

$$\theta = 2/3 \quad (5.3)$$

using the notation in which the correction term is $|\beta - \beta_c|^{\theta}$ or $L^{-\theta/\nu}$.

We can write the singular part of the free energy as a function of the thermal field $t = \beta - \beta_c$, the ordering field h , the leading irrelevant field u , and the linear size of the system L as⁽³⁶⁾

$$f_s(t, h, L) = L^{-d} F(tL^{1/\nu}, hL^{d-\beta/\nu}, uL^{-\theta/\nu}) \quad (5.4)$$

where d is the dimensionality of the lattice. If we differentiate (5.4) twice with respect to the thermal field t and then take the limit $t = h = 0$, we get the specific heat at criticality on a finite lattice:

$$C_H(0, 0, L) \sim \frac{\partial^2 f_s}{\partial t^2}(0, 0, L) = L^{\alpha/\nu} G(0, 0, uL^{-\theta/\nu}) \quad (5.5a)$$

$$= L^{\alpha/\nu} [A + BL^{-\theta/\nu} + \dots] \quad (5.5b)$$

where $\alpha/\nu = 2/\nu - d$, G is the second derivative of F with respect to its first argument, and the dots indicate subdominant corrections. Thus, the corrections to the specific heat at criticality are given by $L^{-\Delta}$ with $\Delta = \theta/\nu = 4/5 = 0.8$. A similar analysis can be carried out for the magnetization and the susceptibility, giving again corrections proportional to $L^{-\Delta}$.

The energy E is obtained by differentiating the full free energy $f = f_s + f_{ns}$ with respect to the thermal field t . The contribution of the nonsingular piece is believed to be trivial: there is numerical evidence that $f_{ns}(t, L) = f_{ns}(t, \infty)$.⁽³⁶⁾ In other words, this contribution has no L -dependence and gives merely the infinite-volume value of the energy at the given temperature, $E(\beta, \infty)$. The contribution of the singular piece can be obtained by differentiating (5.4) once with respect to t ; it goes to zero like $L^{-d+1/\nu}$, which thus gives the leading correction to scaling for the energy. This correction is of order $L^{-4/5}$, which (by pure coincidence as far as we can tell) is exactly the same order as the correction $L^{-\Delta}$ for the divergent static observables. Finally, we note that the energy of the 2D $q = 3$ Potts model at criticality is also exactly known⁽¹⁹⁾ to be $E(\beta_c, \infty) = 1 + 1/\sqrt{3} \approx 1.577350$.

We can check these predictions by performing the fit $E - E(\beta_c, \infty) = AL^{-\nu}$. The quality of the fit is very good already for $L_{\min} = 16$:

$$w = 0.803 \pm 0.002 \quad (5.6)$$

with $\chi^2 = 3.23$ (5 DF, level = 66%). The agreement with the predicted exponent $2 - 1/\nu = 4/5 = 0.8$ is truly spectacular. (Indeed, it is probably a coincidence that the agreement is *so* good. In general, one cannot expect to obtain anywhere near this accuracy for correction-to-scaling exponents.)

The fits of the susceptibility to a power law $AL^{\gamma/\nu}$ are quite stable. Our preferred fit corresponds to $L_{\min} = 64$:

$$\gamma/\nu = 1.73444 \pm 0.00043 \quad (5.7)$$

with $\chi^2 = 4.09$ (3 DF, confidence level = 25%). This result is 2.6 standard deviations away from the exact value $\gamma/\nu = 1.73333$. This discrepancy could be due to corrections to scaling. We can try to fit our data to $AL^{26/15}(1 + BL^{-\Delta})$ with various choices for $\Delta = 1.1, 1.0, \dots, 0.1$ as well as $0 \times \log$ (i.e., a correction $1/\log L$). We find that the best fits correspond to $\Delta \approx 0.8$, in agreement with the theoretical prediction; for $\Delta = 0.8$ and $L_{\min} = 8$ we obtain $\chi^2 = 6.55$ (6 DF, level = 36%). Surprisingly, in all these fits we find that the χ^2 remains almost constant when L_{\min} is increased beyond our “preferred” value (and the confidence level consequently deteriorates): for example, with $L_{\min} = 256$ we get $\chi^2 = 3.37$ (1 DF, level = 7%) for the pure power-law behavior and $\chi^2 = 3.91$ (1 DF, level = 5%) for the fit with $\gamma/\nu = 26/15$ and $\Delta = 0.8$. This suggests that the point with $L = 1024$ is off by about two standard deviations (possibly because the error bar is underestimated¹²). As a matter of fact, if we drop this point, we obtain a good power-law fit for $L_{\min} = 128$:

$$\gamma/\nu = 1.73337 \pm 0.00080 \quad (5.8)$$

with $\chi^2 = 0.39$ (1 DF, level = 53%). This result agrees excellently with the exact value. If we impose the right power $\gamma/\nu = 26/15$ and try to fit the first correction-to-scaling exponent Δ , we again find that the best fits correspond to $\Delta \approx 0.8$. For $L_{\min} = 8$ we get $\chi^2 = 3.84$ (5 DF, level = 57%).

For the specific heat we find that the fits to the power law $AL^{\alpha/\nu}$ are not stable at all: the confidence levels are horrible, and there is a clear trend toward smaller values of α/ν as L_{\min} is increased. The least-bad fit is obtained for $L_{\min} = 256$:

$$\alpha/\nu = 0.4240 \pm 0.0030 \quad (5.9)$$

with $\chi^2 = 3.80$ (1 DF, level = 5%). This value is eight standard deviations away from the exact result $\alpha/\nu = 2/5 = 0.4$. Unlike the four-state Potts

¹² This is quite possible: though the total run length at $L = 1024$ is $7 \times 10^4 \tau_{\text{int},\sigma}$, the *individual* runs (on which the time-series analysis was performed) ranged in length from only $10^4 \tau_{\text{int},\sigma}$ to $2.5 \times 10^4 \tau_{\text{int},\sigma}$: and with such short runs the time-series analysis may not be completely reliable.

model,^(23,4) we do not expect multiplicative corrections to the leading term of the specific heat. We do, however, expect additive corrections to scaling of the form $AL^{2.5}(1+BL^{-A})$. If we try the same exponents A as in the susceptibility, we find a decent fit for $A \approx 0.6$: for $L_{\min} = 128$ we get $\chi^2 = 1.21$ (2 DF, level = 55%). This value of the exponent A is not far from the expected value 0.8, but the cause of the discrepancy is unknown; perhaps there is a large next-to-leading correction to scaling.

Finally, the second-moment correlation length ξ is expected to behave linearly in L as $L \rightarrow \infty$. In particular, the ratio $x \equiv \xi/L$ should approach a constant x^* . We have tested this behavior. Already for $L_{\min} = 16$ our data are consistent with a constant value

$$x^* = \lim_{L \rightarrow \infty} \frac{\xi(L)}{L} = 0.93235 \pm 0.00033 \quad (5.10)$$

with $\chi^2 = 8.08$ (6 DF, level = 23%). The fact that a good fit can be obtained with such a small L_{\min} implies that the corrections to scaling are very small for this observable. [If we fit to $\xi(L)/L = x^* + BL^{-A}$ with $0.1 \lesssim A \lesssim 1.5$, we find a very slight improvement in the goodness of fit, and the estimated x^* decreases somewhat (by less than 0.0008 if $A \gtrsim 0.6$). But the estimated coefficient B is consistent with zero at the 1.5σ level.] As in the case of the susceptibility, we also find that the goodness of fit deteriorates as L_{\min} is increased (e.g., for $L_{\min} = 512$ we have $\chi^2 = 4.80$, 1 DF, level = 3%). This might be due to the fact that the value for $L = 1024$ is a little bit off (or its error bar is underestimated). If we drop this point, we obtain a good fit again for the same $L_{\min} = 16$:

$$x^* = 0.93229 \pm 0.00034 \quad (5.11)$$

with $\chi^2 = 5.82$ (5 DF, level = 32%). But now the fit with $L_{\min} = 256$ is reasonable ($\chi^2 = 1.24$, 1 DF, level = 27%). We remark that the value x^* should in principle be calculable by conformal field theory; we hope that someone will perform this calculation and test our results (5.10)/(5.11).

5.2. Dynamic Quantities

In this section we are going to fit the autocorrelation times for the observables $\mathcal{O} = \mathcal{E}, \mathcal{E}', \mathcal{N}$, and \mathcal{H}^2 to a simple power law $\tau_{\text{int}, \mathcal{O}} = AL^{-z_{\text{int}, \mathcal{O}}}$.

Let us start with the energy \mathcal{E} . The fit $\tau_{\text{int}, \mathcal{E}} = AL^{-z_{\text{int}, \mathcal{E}}}$ is not very stable: the estimate of the power decreases systematically as L_{\min} is increased, and the χ^2 is poor until $L_{\min} = 128$, where the estimate stabilizes within errors (this behavior suggests that there are strong corrections to scaling). Our preferred fit corresponds to $L_{\min} = 128$:

$$z_{\text{int}, \mathcal{E}} = 0.515 \pm 0.006 \quad (5.12)$$

with $\chi^2 = 0.44$ (2 DF, level = 80%). This value is greater than our estimate for α/ν ; thus, the Li-Sokal bound (1.2) holds, though apparently not as a strict equality. To compare our result with the one reported in ref. 11, we redid the fit using only the data with $L \leq 256$. The fit again shows a systematic decrease of the exponent as L_{\min} is increased, as well as a poor χ^2 , until $L_{\min} = 64$, where we get $z_{\text{int.},\sigma} = 0.531 \pm 0.005$ with $\chi^2 = 2.67$ (1 DF, level = 10%). This is consistent with the result $z_{\text{int.},\sigma} = 0.55 \pm 0.02$ reported in ref. 11, but our error bar is one-fourth of theirs. The slightly higher estimate of $z_{\text{int.},\sigma}$ in ref. 11 seems to arise from corrections to scaling induced by their choice $L_{\min} = 16$. Actually, the best fit to the data in ref. 11 corresponds to $L_{\min} = 32$ and gives $z_{\text{int.},\sigma} = 0.52 \pm 0.02$ with $\chi^2 = 3.93$ (2 DF, level = 14%).

The fit of the autocorrelation time for the bond occupation \mathcal{N} follows the same pattern: the estimate of the power decreases strongly as L_{\min} increases, and the χ^2 is initially horrible; eventually the power stabilizes within errors, and the χ^2 becomes reasonable. Our preferred fit is $L_{\min} = 128$:

$$z_{\text{int.},\tau} = 0.529 \pm 0.006 \quad (5.13)$$

with $\chi^2 = 0.71$ (2 DF, level = 70%). The difference with respect to $z_{\text{int.},\sigma}$ is only 2.3 standard deviations, consistent with the theoretical prediction (2.29) that $z_{\text{int.},\tau} = z_{\text{int.},\sigma}$. To check this result, we studied the ratio $\tau_{\text{int.},\tau}/\tau_{\text{int.},\sigma}$. Since the standard time-series-analysis method would give only an upper bound on the right error bar for this quantity, we used instead the error bar provided by the ‘‘bunch’’ method. Among all the Ansätze we tried, only one gave a good fit for $L_{\min} = 128$: asymptotically constant with corrections $\sim L^{-1/4}$ ($\chi^2 = 0.75$, 2 DF, level = 69%). Thus, we conclude that in the SW algorithm the dynamic critical behavior of the energy and that of the bond density are the same: $z_{\text{int.},\tau} = z_{\text{int.},\sigma}$.

In the same way, we considered the nearest-neighbor connectivity \mathcal{E}' . The pattern of the fit is the same as above and our preferred fit again corresponds to $L_{\min} = 128$:

$$z_{\text{int.},\mathcal{E}'} = 0.514 \pm 0.006 \quad (5.14)$$

with $\chi^2 = 0.47$ (2 DF, level = 79%). This time, the agreement with $z_{\text{int.},\sigma}$ is extremely good, confirming the theoretical prediction (2.41) that $z_{\text{int.},\mathcal{E}'} = z_{\text{int.},\sigma}$.

Finally, a similar behavior is observed in the fit for the autocorrelation time of the squared magnetization \mathcal{M}^2 : the estimate of the power shows a clear trend toward smaller values as L_{\min} is increased, and the χ^2 is initially poor. Our preferred fit is $L_{\min} = 128$, for which we get

$$z_{\text{int.},\mathcal{M}^2} = 0.475 \pm 0.006 \quad (5.15)$$

with $\chi^2 = 0.36$ (2 DF, level = 84%). This power is slightly smaller than $z_{\text{int},\mathcal{E}}$ (the difference is seven standard deviations). It would be interesting to know whether this difference is significant or not. Let us consider the ratio $\tau_{\text{int},\mathcal{M}^2}/\tau_{\text{int},\mathcal{E}}$, again using the error bar provided by the ‘‘bunch’’ method. We tried to fit this ratio to various Ansätze, but only two gave good results for $L_{\text{min}} = 32$: a pure power-law behavior AL^p with $p = -0.0409 \pm 0.0007$ ($\chi^2 = 0.57$, 5 DF, level = 97%), and a constant plus corrections of the type $\sim L^{-1/16}$ ($\chi^2 = 0.41$, 5 DF, level = 98%). We are therefore unable to resolve whether $z_{\text{int},\mathcal{M}^2}$ is exactly equal to $z_{\text{int},\mathcal{E}}$ or not; but if it is not equal, then it is only very slightly smaller ($z_{\text{int},\mathcal{M}^2} - z_{\text{int},\mathcal{E}} \approx -0.04$).

In conclusion, we have shown numerically that the energy \mathcal{E} , the bond occupation \mathcal{N} , and the nearest-neighbor connectivity \mathcal{E}' all have the same dynamic critical exponent z_{int} , as was ‘‘almost proved’’ in Section 2. On the other hand, the observable \mathcal{M}^2 has a *similar but perhaps not identical* dynamic critical behavior: $z_{\text{int},\mathcal{M}^2}$ may coincide with $z_{\text{int},\mathcal{E}}$ or it may be slightly smaller.

5.3. Analysis of the Li–Sokal Bound

In the previous subsection we saw that all the observables considered have the same (or, in the case of \mathcal{M}^2 , almost the same) dynamic critical exponent, and the common z_{int} is strictly larger than α/ν . This implies that the Li–Sokal bound (1.2) is *not sharp*. However, the difference $z_{\text{int}} - \alpha/\nu \approx 0.115$ is actually not very large. There are a few arguments in favor of a more detailed analysis:

(i) The power-law fit to the specific heat was not very good: the estimated value of α/ν decreased as L_{min} increased, and we were unable to reach (within statistical errors) the exact value $\alpha/\nu = 2/5$. If we compare the *observed* values of z_{int} and α/ν , we find that $z_{\text{int}} - \alpha/\nu$ is only ≈ 0.09 .

(ii) The power-law fits to the autocorrelation times also exhibited this trend to smaller values as L_{min} increased. Even though the fits seemed stable for $L_{\text{min}} = 128$, this might well be an artifact due to the large error bars associated with the largest lattices ($L \geq 512$) compared to the smallest ones ($L \leq 256$).

(iii) In refs. 12 and 13 it was shown that differences $z_{\text{int}} - \alpha/\nu$ of order 0.1 could actually be due to multiplicative logarithmic corrections. Indeed, such a multiplicative logarithmic correction was found to be a likely scenario even for models *not* having such a logarithmic correction in the specific heat.

Here, we will follow the approach of refs. 12 and 13, and consider the ratio $\tau_{\text{int},\mathcal{E}}/C_H$; we shall use the error bars obtained from the ‘‘bunch’’ method (see Table III).¹³ We tried to fit this ratio to different Ansätze:

1. A pure power-law behavior AL^p .
2. A logarithmic growth, either as $A + B \log L$ or as $A \log^p L$.
3. Asymptotically constant with additive corrections to scaling $A + BL^{-A}$. We considered the cases $A = 2, 1, 1/2, 1/4, 1/8$, and $0 \times \log$ (i.e., $\tau_{\text{int},\mathcal{E}}/C_H = A + B/\log L$).

Among all these Ansätze, only two were reasonably good. The best one is the simple power-law behavior AL^p . For $L_{\min} = 32$ we obtain

$$p = 0.084 \pm 0.002 \quad (5.16)$$

with $\chi^2 = 1.72$ (4 DF, level = 79 %). When $L_{\min} > 32$, the value of the power p stays stable and χ^2 decreases slowly and consistently.

The second-best fit corresponds to the logarithmic growth $A + B \log L$. For $L_{\min} = 64$, we get $\chi^2 = 1.93$ (3 DF, level = 59 %). Again, the estimates are stable for $L_{\min} > 64$ and the χ^2 is reasonable. However, this Ansatz seems to be slightly inferior to the power-law fit: both L_{\min} and the χ^2 are greater than in the power-law case. To test the logarithmic Ansatz, we can impose the known critical behavior of the specific heat and perform the fit $\tau_{\text{int},\mathcal{E}} = L^{2/5}(A + B \log L)$. A reasonably good result is obtained for $L_{\min} = 32$, giving $\chi^2 = 1.54$ (4 DF, level = 82 %).

The asymptotically constant fits are always horrible unless one takes $L_{\min} = 256$ or larger. The only semi-exception is $A = 1/8$, which gives a tolerable fit ($\chi^2 = 5.54$, 3 DF, level = 14 %) already for $L_{\min} = 64$.

Of course for $L_{\min} = 256$ we obtain reasonably good fits for all Ansätze. But this is because the error bars for $L > 256$ are so large that we are unable to distinguish among the various scenarios.

In summary, we have shown that there are only two likely scenarios for the ratio $\tau_{\text{int},\mathcal{E}}/C_H$:

$$\tau_{\text{int},\mathcal{E}}/C_H \sim \begin{cases} AL^p & \text{with } p = 0.084 \pm 0.002 \\ A + B \log L \end{cases} \quad (5.17)$$

or equivalently

$$\tau_{\text{int},\mathcal{E}} \sim \begin{cases} AL^{z_{\text{int},\mathcal{E}}} & \text{with } z_{\text{int},\mathcal{E}} = 0.515 \pm 0.006 \\ L^{2/5}(A + B \log L) \end{cases} \quad (5.18)$$

¹³ In this section we only consider the energy \mathcal{E} , as the other observables have the same (or very slightly smaller) dynamic critical exponent.

The first scenario implies that the Li–Sokal bound (1.2) is a strict inequality, with $p \equiv z_{\text{int},\sigma} - \alpha/\nu \approx 0.08\text{--}0.12$; while the second scenario means that this bound is violated only by a logarithm (i.e., the bound is sharp modulo logarithms). Thus, the same “dynamic continuity” we found interpolating between the two-state and the four-state Potts models along the AT self-dual line is also found along the q -state Potts model critical line. Moreover, the power p found is midway between the power found for the Ising model ($p \approx 0.05$) and the one found for the four-state Potts model ($p \approx 0.12$).⁽¹³⁾

5.4. Further Discussion of the Li–Sokal Bound

The proof of the Li–Sokal bound (Section 2.2) is based on computing the autocorrelation function for the observable \mathcal{N} at time lags 0 and 1 in terms of static observables and then exploiting the inequality $\rho_{\cdot,1,1}(t) \geq \rho_{\cdot,1,1}(1)^{|t|}$. The apparent nonsharpness of this bound indicates that the large- t behavior of $\rho_{\cdot,1,1}(t)$ is *not* fully predicted by its behavior at $t=1$. Otherwise put, if we define the *initial* autocorrelation time

$$\tau_{\text{init},e} \equiv \frac{1}{2} \times \frac{1 + \rho_{ee}(1)}{1 - \rho_{ee}(1)} \quad (5.19)$$

then the Li–Sokal bound

$$\tau_{\text{int},1} \geq \tau_{\text{init},1} = \frac{p}{1-p} \frac{C_H}{E} + \frac{1}{2} \quad (5.20)$$

is apparently *not sharp*: that is, $\tau_{\text{int},1}$ and $\tau_{\text{init},1}$ diverge with *different* critical exponents $z_{\text{int},1}$ and $z_{\text{init},1} = \alpha/\nu$.

One might ask whether the situation could be improved by computing $C_{\cdot,1,1}(t)$ exactly at (for example) time lag $t=2$ or $t=3$. Because of the identities (2.25)/(2.36), this would be equivalent to carrying out the Li–Sokal proof using \mathcal{E} or \mathcal{E}' as the test observable in place of \mathcal{N} . At present we do not know how carry out this computation—the trouble is that we do not know how to express $\langle \gamma_{ij} \gamma_{kl} \rangle$ in terms of spin observables—but we can nevertheless test numerically whether we *would* obtain a better bound on $z_{\text{int},\sigma}$ and/or $z_{\text{int},\sigma'}$ if we *could* carry out this computation.

We thus computed $\tau_{\text{init},e}$ for $\mathcal{O} = \mathcal{N}, \mathcal{E}, \mathcal{E}'$, with error bars given by (4.10)¹⁴; we then tried various fits, among others the fit $\tau_{\text{init},e}/C_H = AL^r$.¹⁵

¹⁴ Since the autocorrelation function for these observables is almost exactly a pure exponential, it suffices to perform the sums in (4.10) *analytically* and then insert the appropriate value of τ .⁽¹³⁾

¹⁵ For the error bars on $\tau_{\text{init},e}/C_H$, we used for simplicity the triangle inequality rather than the (more correct) bunch method.

For $\mathcal{O} = \mathcal{N}$ we have the identity

$$\frac{\tau_{\text{init},\mathcal{N}}}{C_H} = \frac{p}{1-p} \frac{1}{E} + \frac{1}{2C_H} \tag{5.21a}$$

$$= \left[\frac{p}{1-p} \frac{1}{E(\beta, \infty)} + aL^{-4/5} + \dots \right] + [bL^{-2/5} + cL^{-6/5} + \dots] \tag{5.21b}$$

where $E(\beta, \infty)$ is the infinite-volume energy at inverse temperature β ; and we have taken into account the leading terms and the corrections to scaling discussed in Section 5.1. At criticality $\beta = \beta_c = \log(1 + \sqrt{3})$ [hence $p = p_c = \sqrt{3}/(1 + \sqrt{3})$] we know⁽¹⁹⁾ the exact constant term:

$$\frac{p_c}{1-p_c} \frac{1}{E(\beta_c, \infty)} = \frac{3}{1 + \sqrt{3}} \approx 1.098076 \tag{5.22}$$

Indeed, if we fit our data to the Ansatz $\tau_{\text{init},\mathcal{N}}/C_H \sim 3(1 + \sqrt{3}) = bL^{-s}$ we obtain a good fit for $L_{\min} = 16$, and the estimate is

$$s = 0.415 \pm 0.017 \tag{5.23}$$

($\chi^2 = 1.92$, 3 DF, level = 86%), in excellent agreement with the theoretical prediction. Of course, we could try also the more primitive fit $\tau_{\text{init},\mathcal{N}}/C_H = AL^r$, and we get a good fit for $L_{\min} = 128$:

$$r = -0.010 \pm 0.002 \tag{5.24}$$

($\chi^2 = 0.90$, 2 DF, level = 64%), which is very close to the exact result $r = 0$. All this is trivially to be expected, since our raw data for $\rho_{\mathcal{N},\mathcal{N}}(1)$, C_H , and E are in good agreement with the theoretical identity (2.16).

The analysis becomes nontrivial when we look at $\mathcal{O} = \mathcal{E}$ and \mathcal{E}' . For $\mathcal{O} = \mathcal{E}$ we fit $\tau_{\text{init},\mathcal{E}}/C_H = AL^r$ and get a good fit for $L_{\min} = 64$:

$$r = 0.003 \pm 0.003 \tag{5.25}$$

($\chi^2 = 0.91$, 3 DF, level = 82%). For $\mathcal{O} = \mathcal{E}'$ the analogous fit is good for $L_{\min} = 32$:

$$r = 0.016 \pm 0.001 \tag{5.26}$$

($\chi^2 = 0.51$, 4 DF, level = 97%). These exponents r are a factor of ≈ 5 smaller than the exponent estimates for $p \equiv z - \alpha/\nu$ found in Section 5.3. This suggests that $\tau_{\text{init},\mathcal{E}}/C_H$ and $\tau_{\text{init},\mathcal{E}'}/C_H$ in fact tend to *finite constants* as $L \rightarrow \infty$ (as we know rigorously to be the case for $\tau_{\text{init},\mathcal{N}}/C_H$) and that

the apparent nonsharpness of the Li–Sokal bound cannot be remedied by using \mathcal{E} or \mathcal{E}' as a test observable in place of \mathcal{N} . Rather, the nonsharpness of the bound arises from the fact that the long-time behavior of $\rho_{\ell\ell}(t)$ is not sufficiently well predicted by its behavior at *any* small time.

5.5. Exponential Autocorrelation Time

The exponential autocorrelation time for an observable \mathcal{O} is defined as¹⁶

$$\tau_{\text{exp},\ell} = \lim_{t \rightarrow \infty} \frac{-|t|}{\log \rho_{\ell\ell}(t)} \quad (5.27)$$

This autocorrelation time measures the decay rate of the “slowest mode” of the system, provided that this mode is not orthogonal to \mathcal{O} .

The critical behavior of $\tau_{\text{exp},\ell}$ is, in general, different from the behavior of $\tau_{\text{int},\ell}$. This fact can be seen from the standard dynamic finite-size-scaling Ansatz for the autocorrelation function $\rho_{\ell\ell}(t)$:

$$\rho_{\ell\ell}(t; L) \approx |t|^{-p_\ell} h_\ell \left(\frac{t}{\tau_{\text{exp},\ell}}; \frac{\xi(L)}{L} \right) \quad (5.28)$$

(Here the dependence on the coupling constants has been suppressed for notational simplicity.) Summing (5.28) over t , it follows that

$$\tau_{\text{int},\ell} \sim \tau_{\text{exp},\ell}^{1-p_\ell} \quad (5.29)$$

or equivalently,

$$z_{\text{int},\ell} = (1 - p_\ell) z_{\text{exp},\ell} \quad (5.30)$$

Thus, only when $p_\ell = 0$ do we have $z_{\text{int},\ell} = z_{\text{exp},\ell}$.⁽⁴⁾ In this latter case the Ansatz (5.28) can be rewritten in the equivalent form

$$\rho_{\ell\ell}(t; L) \approx \hat{h}_\ell \left(\frac{t}{\tau_{\text{int},\ell}}; \frac{\xi(L)}{L} \right) \quad (5.31)$$

¹⁶ For a general Markov chain, the “lim” should strictly speaking be replaced by “lim sup” and $\rho_{\ell\ell}(t)$ should be replaced by its absolute value. But here it can be *proven* that the limit really exists and that $\rho_{\ell\ell}(t) \geq 0$ for all t ; this follows from the spectral representation (2.19) [or rather its analogue for ℓ].

To test this latter Ansatz, we have plotted $\log \rho_{\mathcal{O}}(t)$ versus $t/\tau_{\text{int},\mathcal{O}}$ for the observables $\mathcal{O} = \mathcal{N}$ (Fig. 1), \mathcal{E} (Fig. 2), and \mathcal{E}' (Fig. 3). On each figure we have plotted the data coming from different lattice sizes ($4 \leq L \leq 128$) with different symbols; the error bars are computed from (4.10), using for simplicity the approximation⁽¹³⁾ that the decay is a pure exponential (which is here almost exact). On each graph, we have also depicted for reference a line corresponding to the pure exponential $\rho_{\mathcal{O}}(t) = e^{-t/\tau_{\text{int},\mathcal{O}}}$. In these plots we have omitted the data for $L \geq 256$ for the sake of visual clarity; these data agree well with the curve found for smaller L , but with huge statistical error bars.

For $\mathcal{O} = \mathcal{N}$ we see that the data coming from $16 \leq L \leq 128$ collapse well onto a single curve. The lattices $L = 4, 8$ show slight systematic deviations from this limiting curve: these deviations are negative for $t/\tau_{\text{int},\mathcal{N}} \lesssim 1.5$ and positive for $t/\tau_{\text{int},\mathcal{N}} \gtrsim 1.5$. This trend continues for $16 \leq L \leq 128$, but the deviations are in most cases smaller than the error bars (especially for the larger lattices).

A similar behavior is found for $\mathcal{O} = \mathcal{E}$ (Fig. 2) and $\mathcal{O} = \mathcal{E}'$ (Fig. 3), but the corrections to scaling are much weaker and their sign is opposite from those seen for \mathcal{N} .

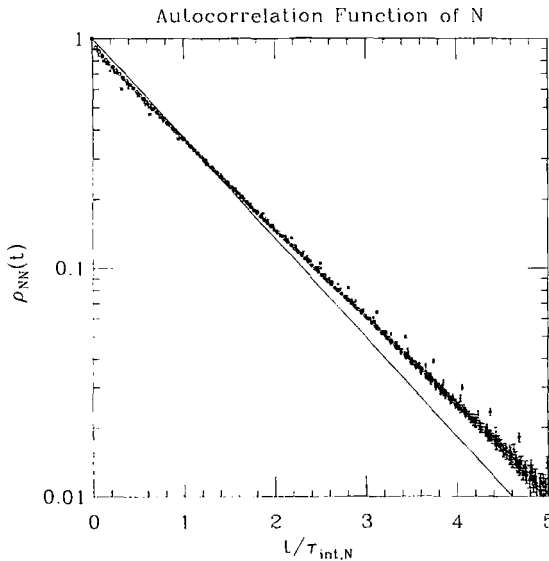


Fig. 1. Plot of $\rho_{\mathcal{N}}(t)$ versus $t/\tau_{\text{int},\mathcal{N}}$ for $4 \leq L \leq 128$. The different symbols denote the different lattice sizes: $L = 4$ (*), $L = 8$ (+), $L = 16$ (x), $L = 32$ (□), $L = 64$ (◇), and $L = 128$ (○). We have also depicted the line corresponding to the pure exponential $\rho_{\mathcal{N}}(t) = \exp(-t/\tau_{\text{int},\mathcal{N}})$.

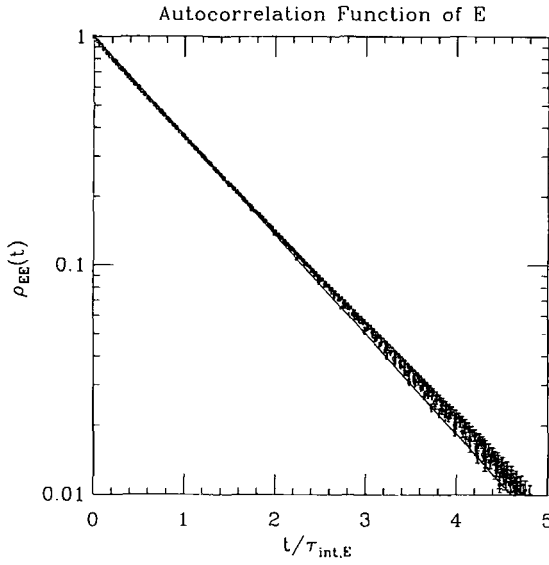


Fig. 2. Plot of $\rho_{\sigma\sigma}(t)$ versus $t/\tau_{\text{int},\sigma}$ for $4 \leq L \leq 128$. The symbols are as in Fig. 1. We have also depicted the line corresponding to the pure exponential $\rho_{\sigma\sigma}(t) = \exp(-t/\tau_{\text{int},\sigma})$.

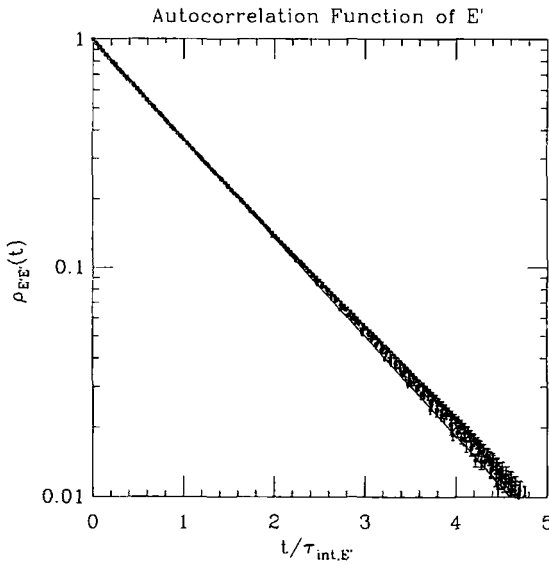


Fig. 3. Plot of $\rho_{\sigma'\sigma'}(t)$ versus $t/\tau_{\text{int},\sigma'}$ for $4 \leq L \leq 128$. The symbols are as in Fig. 1. We have also depicted the line corresponding to the pure exponential $\rho_{\sigma'\sigma'}(t) = \exp(-t/\tau_{\text{int},\sigma'})$.

Thus, within statistical errors, we have found that the Ansatz (5.31) is satisfied. This implies that the integrated and exponential autocorrelation times for these three observables have exactly the same dynamic critical exponent, i.e., $z_{\text{int},\mathcal{O}} = z_{\text{exp},\mathcal{O}}$ for $\mathcal{O} = \mathcal{N}, \mathcal{E}, \mathcal{E}'$. This equality does *not* hold as a general rule in the theory of dynamic critical phenomena,^(4,6) but it does appear to hold for algorithms of Swendsen–Wang type.

Finally, we find that $\rho_{\nu_{1,1}'}(t)$ differs slightly but noticeably from the pure exponential $\exp(-t/\tau_{\text{int},\nu_{1,1}'})$. The discrepancy from a pure exponential becomes smaller when we consider the other two observables \mathcal{E} and \mathcal{E}' . This is to be expected theoretically from the relations $\mathcal{E} \sim P_{\text{bond}}\mathcal{N}$ and $\mathcal{E}' \sim P_{\text{spin}}\mathcal{E}$ [cf. (2.14)/(2.35)]: each action of P_{bond} or P_{spin} helps to “purify” the slowest mode, so that the autocorrelation function becomes closer to a pure exponential. On the other hand, the identities (2.26)/(2.37) imply¹⁷ that the limiting (scaling) functions $\hat{h}_{\nu_{1,1}'}$, $\hat{h}_{\mathcal{E}}$, and $\hat{h}_{\mathcal{E}'}$ are *identical* (assuming they exist at all); this means that at least some of the curves in Figs. 1–3 have not yet reached their scaling limit.

Finally, a crude fit suggests that $\tau_{\text{int},\mathcal{E}}/\tau_{\text{exp},\mathcal{E}} \approx 0.96$, in agreement with the idea that $\rho_{\mathcal{E}\mathcal{E}}(t)$ is almost but not quite a pure exponential.

ACKNOWLEDGMENTS

We wish to give our *big* thanks to Wolfhard Janke for pointing out an error in an earlier version of this paper. We also thank the Cornell Theory Center for the computer time needed to complete this project. The authors’ research was supported in part by U.S. National Science Foundation grants DMS-9200719 and PHY-9520978 and by Metacenter grant MCA94P032P.

REFERENCES

1. K. Binder, ed., *Monte Carlo Methods in Statistical Physics* (Springer-Verlag, Berlin, 1978).
2. K. Binder, ed., *Applications of the Monte Carlo Method in Statistical Physics* (Springer-Verlag, Berlin, 1987).
3. K. Binder, ed., *The Monte Carlo Method in Condensed Matter Physics* (Springer-Verlag, Berlin, 1992).
4. A. D. Sokal, *Monte Carlo Methods in Statistical Mechanics: Foundations and New Algorithms*, Cours de Troisième Cycle de la Physique en Suisse Romande, Lausanne (June 1989).
5. U. Wolff, *Nucl. Phys. B (Proc. Suppl.)* **17**:93 (1990).
6. A. D. Sokal, *Nucl. Phys. B (Proc. Suppl.)* **20**:55 (1991).

¹⁷ Using the fact that $\rho_{\nu_{1,1}'}(1) \rightarrow 1$ and $\tau_{\text{int},\nu_{1,1}'} \rightarrow \infty$ as $L \rightarrow \infty$, as follows from the Li–Sokal identity (2.16) and the fact that the specific heat C_H is divergent.

7. R. H. Swendsen and J.-S. Wang, *Phys. Rev. Lett.* **58**:86 (1987).
8. D. W. Heermann and A. N. Burkitt, *Physica A* **162**:210 (1990).
9. C. F. Baillie and P. D. Coddington, *Phys. Rev. B* **43**:10617 (1991).
10. C. F. Baillie and P. D. Coddington, *Phys. Rev. Lett.* **68**:962 (1992); and private communication.
11. X.-J. Li and A. D. Sokal, *Phys. Rev. Lett.* **63**:827 (1989).
12. J. Salas and A. D. Sokal, *Nucl. Phys. B (Proc. Suppl.)* **47**:792 (1996).
13. J. Salas and A. D. Sokal, *J. Stat. Phys.* **85**:297 (1996).
14. J. Salas and A. D. Sokal, Logarithmic corrections and finite-size scaling in the two-dimensional 4-state Potts model, *J. Stat. Phys.* (to appear).
15. W. Klein, T. Ray, and P. Tamayo, *Phys. Rev. Lett.* **63**:827 (1989).
16. T. Ray, P. Tamayo, and W. Klein, *Phys. Rev. A* **39**:5949 (1989).
17. S. Wiseman and E. Domany, *Phys. Rev. E* **48**:4080 (1993).
18. J. Ashkin and J. Teller, *Phys. Rev.* **64**:178 (1943).
19. R. J. Baxter, *Exactly Solved Models in Statistical Mechanics* (Academic Press, New York, 1982).
20. A. E. Ferdinand and M. E. Fisher, *Phys. Rev.* **185**:832 (1969).
21. S. Alexander, *Phys. Lett. A* **54**:353 (1975).
22. E. Domany and E. K. Riedel, *J. Appl. Phys.* **49**:1315 (1978).
23. M. Nauenberg and D. J. Scalapino, *Phys. Rev. Lett.* **44**:837 (1980).
24. J. L. Cardy, M. Nauenberg, and D. J. Scalapino, *Phys. Rev. B* **22**:2560 (1980).
25. J. L. Black and V. J. Emery, *Phys. Rev. B* **23**:429 (1981).
26. R. G. Edwards and A. D. Sokal, *Phys. Rev. D* **38**:2009 (1988).
27. P. W. Kasteleyn and C. M. Fortuin, *J. Phys. Soc. Japan* **26**(Suppl.):11 (1969).
28. C. M. Fortuin and P. W. Kasteleyn, *Physica* **57**:536 (1972).
29. C. M. Fortuin, *Physica* **58**:393 (1972); **59**:545 (1972).
30. T. W. Anderson, *The Statistical Analysis of Time Series* (Wiley, New York, 1971).
31. M. B. Priestley, *Spectral Analysis and Time Series*, 2 Vols. (Academic Press, London, 1981).
32. N. Madras and A. D. Sokal, *J. Stat. Phys.* **50**:109 (1988).
33. A. D. Sokal, Unpublished.
34. B. Nienhuis, *J. Phys. A: Math. Gen.* **15**:199 (1982).
35. V. S. Dotsenko, *Nucl. Phys. B* **235**[FS11]:54 (1984).
36. V. Privman, In *Finite Size Scaling and Numerical Simulation of Statistical Systems*, V. Privman, ed. (World Scientific, Singapore, 1990).