Finite-size scaling and Monte Carlo simulations of first-order phase transitions

Jooyoung Lee and J. M. Kosterlitz
Physics Department, Brown University, Providence, Rhode Island 02912
(Received 11 June 1990)

We develop a detailed finite-size-scaling theory at a general, asymmetric, temperature-driven, strongly first-order phase transition in a system with periodic boundary conditions. We compute scaling functions for various cumulants of energy in the form $U(L,t)=U_0(tL^d)+L^{-d}U_1(tL^d)$ with $t=1-T_c/T$. In particular, we consider the specific heat and Binder's fourth cumulant and show this has a minimum value of $\frac{2}{3}-(e_1/e_2-e_2/e_1)^2/12+O(L^{-d})$ at a temperature $T_c^{(2)}(L)-T_c=O(L^{-d})$. Various other pseudocritical temperatures corresponding to extrema of other cumulants are evaluated. We compare these theoretical predictions with extensive Monte Carlo simulations of the nominally strong first-order transitions in the eight- and ten-state Potts models in two dimensions for system sizes $L\leq 50$. The ten-state simulations agree with theory in all details in contrast to the eight-state data, and we give estimates for the bulk specific heats at T_c using all exactly known analytic results. A criterion is developed to estimate numerically whether or not system sizes used in a simulation of a first-order transition are in the finite-size-scaling regime.

I. INTRODUCTION

This paper is concerned with the analysis of Monte Carlo (MC) data on well-equilibrated finite systems which, in the thermodynamic limit, undergo a first-order phase transition. We consider only systems with periodic boundary conditions with linear size L and volume L^d. In a simulation one is always faced with the problem that computer time is limited and, consequently, system sizes are also limited. There are two issues to be faced in such simulations. The first is to identify the nature of the transition by purely numerical methods in a finite-sized system and the second is to determine quantitatively various thermodynamic quantities at the transition. These are, for example, the equilibrium magnetizations and susceptibilities just above and below the transition line in a field driven case and, in a temperature driven transition, the energies, specific heats, and latent heat at T_c. To evaluate these requires extrapolation from finite size to the thermodynamic limit. In principle, this problem has been solved by using finite-size scaling[1-10] but one must still know if the system sizes used in the simulations are sufficiently large so that leading-order finite-size corrections are enough to perform the extrapolation. For example, it is difficult to even see the first-order nature in the five-state Potts model.[11,12] In particular, we demonstrate that, in reasonably sized systems with $L\leq 50$, the ten-state Potts model[13] in two dimensions obeys finite-size scaling very well but the eight-state model, although apparently scaling as predicted by theory, is misleading and leading-order finite-size corrections are quite inadequate.

In Sec. II we summarize a recent method[14] to identify the nature of the transition by purely numerical means which can also indicate if the system sizes used are sufficiently large to enable comparison with finite-scaling predictions. This involves the simulation of the probability distribution of energy whose logarithm will have a double-peaked structure at a first-order transition. For finite-size scaling to work, the difference between the maxima and minima must both scale as L^{d-1} and also must be sufficiently large so that contributions to estimates of thermodynamic quantities are completely dominated by the minima. For example, for accessible system sizes, simulations on the three-state Potts model in three dimensions[15] show that this scales as L^2 but is still too small numerically for agreement with analytic finite-size-scaling theory. For a five-state model in two dimensions, this never scales as L for accessible system sizes so that comparison with finite-size scaling is pointless.

In Sec. III, we discuss theoretical predictions for several quantities which have characteristic behavior in finite systems which have a first-order transition in the thermodynamic limit. The mimic partition function is defined for a q-state Potts model and, using this, the specific heat of a finite system is calculated up to $O(L^{-d})$ as is Binder's fourth cumulant of energy.[16] In particular, we derive a corrected value of the minimum which differs significantly from that quoted earlier.[5,12,17] Various characteristic temperatures $T_c(L)$ are defined, all of which tend, as $L\rightarrow\infty$, to the bulk critical temperature and are calculated to $O(L^{-2d})$. Although the bulk of the theoretical discussion and all the simulations are on cumulants of energy for the Potts model, we also briefly discuss order-parameter cumulants[4,5] for both symmetric and asymmetric first-order transitions.

Section IV discusses extensive MC simulations for the $q=8$ and 10 state Potts models in two dimensions for system sizes up to $L=50$. Detailed comparison with finite-size-scaling theory is made and it is found that there is excellent agreement for $q=10$ but there is serious disagreement for $q=8$, although the latter has a strong first-order transition. This comparison is made using all available exact results to reduce the number of fitting parameters to a minimum. It turns out that some quantities

such as the specific-heat maximum can be analytically determined to leading order in terms of exactly known quantities and the next order correction in terms of one unknown bulk specific heat. This enables quite accurate estimates of this to be made. In Sec. V we summarize our theoretical and numerical results and discuss their implications, and in the Appendix we summarize some of the tedious calculations.

We note that many of the theoretical results obtained in this paper were independently obtained recently by Borgs[10] *et al.*, especially the value of the minimum of Binder's fourth-order cumulant[16] and the full scaling form of the specific heat to leading order in L^{-d}.

II. FIRST-ORDER TRANSITIONS AND SIMULATIONS

Many methods have been proposed to distinguish numerically between continuous and first-order transitions.[6] They depend either on the hysteresis at first-order transitions or on being able to detect unambiguously first-order finite-size scaling of the form L^{-d}. However, hysteresis may be present at a continuous transition for small systems at short times but if one runs for a sufficiently long time one will explore all of phase space and there will be no hysteresis. Hysteresis is then difficult to detect in an unambiguous way because its presence or absence will depend on the rate of cooling or heating. Finite-size scaling of the form L^{-d} will be present even when the finite system is well equilibrated and does explore all of phase space but will set in only for system size $L \gg \xi$, where ξ is the correlation length at the transition. Since ξ may be many lattice spacings as, for example, in the q-state Potts models where $\xi(q=5)\sim 2000$ (Ref. 11) and $\xi(q=8)\sim 20$, very large systems and prohibitive amounts of computer time may be required.

Recently[14] the problem of detecting a temperature-driven first-order transition by MC in a finite system of volume L^d with periodic boundary conditions has been solved by computing the histogram of energy distribution[17,18]

$$N(E;\beta,L)=\mathcal{N}Z^{-1}(\beta,L)\Omega(E,L)\exp-\beta E \ , \tag{2.1}$$

where $\mathcal{N}$ is the number of MC sweeps, Z is the partition function, and $\Omega(E,L)$ is the number of states of energy E. For a field-driven transition, one computes the corresponding order-parameter distribution. For q-state Potts models with q equivalent ordered and one disordered state, $N(E;\beta,L)$ has a characteristic double-peak structure in the vicinity of T_c. The two peaks at $E_1(L)$ and $E_2(L)$ corresponding, respectively, to the ordered and disordered states are separated by a minimum at $E_m(L)$. Defining

$$A(E;\beta,L,\mathcal{N})\equiv -\ln N(E;\beta,L) \ ,$$

it is easy to see that, at $\beta=\beta_c(L)$ defined by $A(E_1;\beta,L)=A(E_2;\beta,L)$,

$$A(E_m;\beta,L,\mathcal{N})-A(E_1;\beta,L,\mathcal{N})=\Delta F(L) \ , \tag{2.2}$$

where $\Delta F(L)$ is a bulk free-energy barrier between the states and is independent of $\mathcal{N}$. At a continuous transition, $\Delta F(L)$ is independent of L and at a first-order transition is an increasing function of L which, for $L \gg \xi$, $\Delta F(L)\sim L^{d-1}$. This follows because, at a first-order transition, the free energy as a function of E has the expansion

$$F(E,L)=L^d f_0(E)+L^{d-1}f_1(E)+\cdots \ , \tag{2.3}$$

where $f_0(E)$ is the bulk free energy which is a minimum and constant for $E_1 \le E \le E_2$ and $f_1(E)$ is a surface free energy which has a maximum at E_m. Then $F(E,L)$ has minima at

$$E_1(L)=E_1-O(L^{-1})$$

and at

$$E_2(L)=E_2+O(L^{-1})$$

with a maximum at $E_m(L)$, so that

$$\Delta F(L)=F(E_m,L)-F(E_1,L)\sim L^{d-1} \ . \tag{2.4}$$

In the vicinity of a critical point, $\Delta F(L)$ grows more slowly than this at a rate controlled by the critical point for $L<\xi$.[14] Under the assumption that L is sufficiently large so that irrelevant fields have scaled away so that the only fields left are those relevant at the critical point, this constitutes an unambiguous method of numerically detecting a first-order transition. Then, at $\beta_c(L)$, $A(E;\beta,L)$ develops deeper minima and higher barriers as L is increased so, if a set of minima is detected for some L and the barriers grow with L, then the transition is first order. This has been checked numerically for the $q=5$ and 6 Potts models in $d=2$ for system sizes up to $L=60$ (Ref. 14) and for $q=3$ in $d=3$ up to $L=14$ (Ref. 15). In all cases, $\Delta F(L)$ increases monotonically with L as expected, although $L<\xi$.

For the $q=8$ and 10 state models in $d=2$ analyzed in detail in this paper, we find that for $q=10$, $\Delta F(L)\sim L^{d-1}$ but, for $q=8$, ΔF is increasing more slowly for $L \lesssim 50$. We therefore expect that, although $\Delta F/kT_c$ is quite large in both cases, the ten-state model will obey first-order finite scaling well but the eight-state model will show significant departures from theory. As we shall see in Sec. IV, these expectations are borne out when detailed comparisons with theory are made. This is not too surprising in view of the fact that one may estimate the correlation lengths as $\xi(10)\simeq 9$ and $\xi(8)\simeq 22$ by defining the correlation length by $\Delta F(L=\xi)=1$ in $d=2$. This estimate follows from $N(E_m)/N(E_1)=1/e$, where $N(E)$ is defined in Eq. (2.1). These numbers are almost the same as those of Peczak and Landau[11] who define their $\xi(10)\simeq 6$ by fitting the cluster size to an assumed Ornstein-Zernicke form.

III. THEORY

Given that the system of interest undergoes a strong first-order transition in the thermodynamic limit, the δ-function singularities in, for example, the specific heat, will be smeared out[1-6] over a temperature range $T-T_c\sim L^{-d}$. There is no critical temperature but one may define a size-dependent temperature $T_c(L)$ which

tends to the bulk T_c as $L \to \infty$. The definition of $T_c(L)$ is somewhat arbitrary and differs according to the quantity being considered. In this paper, we consider three different cumulants of energy: the specific heat per site

$$C(L)L^{-d} \equiv (\langle E^2 \rangle_L - \langle E \rangle_L^2)/T^2 ,$$

$$U_2(L) \equiv \langle E^2 \rangle_L / \langle E \rangle_L^2 ,$$

and

$$U_4(L) \equiv \langle E^4 \rangle_L / \langle E^2 \rangle_L^2$$

which have maxima at $T_c^{(0)}(L)$, $T_c^{(1)}(L)$, and $T_c^{(2)}(L)$, respectively. All these maxima tend to finite constants at T_c as L^{-d} but with different slopes as do $T_c^{(i)}(L)$. In this section, we calculate the $C(L)L^{-d}$, etc., up to $O(L^{-d})$ and $T_c^{(i)} - T_c$ up to $O(L^{-2d})$ and in the next section we compare with simulation results.

It has been rigorously established that, at a general first-order transition where k bulk phases coexist, the partition function of a finite system of linear size L with periodic boundary conditions may be written as[8-10]

$$Z = \sum_{i=1}^{k} \exp -\beta_c f_i(\beta_c) L^d , \tag{3.1}$$

where $\beta_c = 1/kT_c$ is the bulk critical temperature and $f_i(\beta)$ is the free energy per site of the ith bulk phase in the thermodynamic limit. At the transition, all $f_i(\beta_c)$ are equal, so each phase contributes equally to Z. We make the very plausible assumption, as done by other workers,[2,4,19] that, in the immediate vicinity of T_c, all thermodynamic averages may be calculated from Eq. (3.1) with β_c replaced by β.[20] Since the q ordered states of the Potts model are all equivalent

$$Z(\beta, L) = q \exp -\beta f_1(\beta) L^d + \exp -\beta f_2(\beta) L^d , \tag{3.2}$$

where f_1 is the free energy of an ordered state and f_2 that of the disordered state. When $f_1 \neq f_2$, the partition function of Eq. (3.2) reduces to that of a single phase when $L \to \infty$. We shall use the subscript 1 (2) to denote an ordered (disordered) phase.

Near T_c, when $t \equiv 1 - T_c/T \ll 1$, the free energies $f_i(\beta)$ may be Taylor expanded in t about β_c where $f_1(\beta_c) = f_2(\beta_c)$,

$$\beta f_i(\beta) = \beta_c f_i(\beta_c) - \beta_c e_i t - \tfrac{1}{2} C_i t^2 + O(t^3) , \tag{3.3}$$

where $e_i = \partial(\beta f_i)/\partial\beta|_{\beta_c}$ and

$$C_i = -\beta_c^2 \partial^2(\beta f_i)/\partial\beta^2|_{\beta_c}$$

are the energy and specific heat of the ith bulk phase. Defining new variables $x = \beta_c e_1 t L^d$, $a = C_1/\beta_c^2 e_1^2 L^d$, $r = e_2/e_1 < 1$, and $y = C_2/C_1$, using Eq. (3.3) we can write the partition function of Eq. (3.2) as

$$Z = q \exp(x + \tfrac{1}{2} a x^2) + \exp(rx + \tfrac{1}{2} a y x^2) , \tag{3.4}$$

where we have ignored an overall multiplicative constant of $\exp[-\beta_c f(\beta_c) L^d]$. In the temperature range of interest, $t = O(L^{-d})$, since e_i and C_i are finite at the bulk transition, the quantities x, r, y are $O(1)$ and $a = O(L^{-d})$ so we can expand in powers of a to evaluate quantities calculated from Eq. (3.4) to the desired order in L^{-d}.

Before performing such calculations, we note that the form of Z of Eq. (3.4) implies that the probability distribution of energy $P(E)$ is a normalized sum of individually normalized Gaussians. This is easily seen by regarding $Z(\beta)$ as the Laplace transform of $\Omega(E)$ and inverting. After performing the straightforward algebra we obtain

$$P(E) = \frac{1}{(2\pi)^{1/2} 2 \cosh\Delta} \left[e^{\Delta} \left[\frac{\beta_c^2 L^d}{C_1} \right]^{1/2} \exp -\frac{\beta_c^2 L^d}{2C_1} \left[E - e_1 - \frac{C_1 t}{\beta_c} \right]^2 + e^{-\Delta} \left[\frac{\beta_c^2 L^d}{C_2} \right]^{1/2} \exp -\frac{\beta_c^2 L^d}{2C_2} \left[E - e_2 - \frac{C_2 t}{\beta_c} \right]^2 \right] , \tag{3.5}$$

where

$$\Delta = \tfrac{1}{2}(e_1 - e_2)\beta_c t L^d + \tfrac{1}{4}(C_1 - C_2) t^2 L^d + \tfrac{1}{2} \ln q .$$

This generalization of the form used by Challa *et al.*[5] is, of course, equivalent to Eq. (3.4) but differs from that of Ref. 5 where the Gaussians were not individually normalized. From Eq. (3.5), it is obvious that we are approximating the true $P(E)$ by a superposition of Gaussians which ignores in the statistical sum the effects of domain walls with energy $\sim L^{d-1}$. Thus, for theoretical estimates of cumulants derived from Eq. (3.4) or (3.5) to be accurate approximations of their actual values, the minima in $P(E)$ corresponding to stable phases, must be separated by sufficiently high barriers. A necessary condition in a simulation is that $\Delta F(L) \gg 1$ [see Eq. (2.2)]. This is not sufficient as $\Delta F(L) \sim L^{d-1}$ must also hold for a strong first-order transition.

We choose to use Eq. (3.4) to calculate the energy cumulants as these are obtained by differentiating Z with respect to x which is a straightforward (but tedious) procedure. We first study the ratio of cumulants

$$U_4(L) \equiv \langle E^4 \rangle_L / \langle E^2 \rangle_L^2 ,$$

which is related to Binder's fourth cumulant[16]

$$V_4(L) = 1 - U_4(L)/3 .$$

This was introduced as a quantity which could distinguish between first- and second-order transitions since it will have a nontrivial value at a first-order transition. In either a disordered or ordered phase, $T \neq T_c$, $V_4(L) = \frac{2}{3}$ in the thermodynamic limit but, at T_c, all phases contribute

to $V_4(L)$ so it will have a minimum value. However, the value quoted in earlier works,[5,6,12,17,19]

$$V_4(\infty)=\tfrac{2}{3}-(e_1^2-e_2^2)^2/3(e_1^2+e_2^2)^2 , \tag{3.6}$$

is derived on the assumption that, at the transition, $P(E)$ is a sum of equally weighted Gaussians or δ functions. There is no *a priori* reason to assume equal weighting since all that is required is that, as $L\to\infty$, only the disordered state contributes for $T>T_c$ and only the ordered state for $T<T_c$. Any superposition of Gaussians will do this provided only that the weighted of the unstable phase vanishes as $L\to\infty$. In a finite system in the temperature range $tL^d=O(1)$, all states will contribute significantly and, to obtain the $O(L^{-d})$ corrections, it is vital that the relative weights are chosen correctly.

To obtain the correct value of $V_4(\infty)$, all that is necessary is to take a form of the probability distribution which, when $L\to\infty$, becomes a combination of δ functions of the form

$$P(E)=[K(tL^d)\delta(E-e_1)+\delta(E-e_2)]/[1+K(tL^d)] , \tag{3.7}$$

where $K(x)\to\infty$ as $x\to-\infty$ and $K(x)\to 0$ as $x\to\infty$. Both the probability distribution of Eq. (3.5) and that of Eq. (1) of Ref. 5 are of this form. From Eq. (3.7) we immediately obtain

$$V_4(L)=1-(1+K)(Ke_1^4+e_2^4)/(Ke_1^2+e_2^2)^2 . \tag{3.8}$$

Minimizing this with respect to t or, equivalently, K, one obtains

$$V_4(L)|_{\min}=\tfrac{2}{3}-(e_1/e_2-e_2/e_1)^2/12 , \tag{3.9}$$

which differs significantly from the previously quoted value of Eq. (3.6). The next correction of $O(L^{-d})$ depends on the detailed form of $P(E)$ and, since the form chosen by Challa *et al.*[5] differs from Eq. (3.5) or (3.4), we obtain different results. The value given by Eq. (3.9) agrees very well with our simulations, while Eq. (3.6) is inconsistent.

We choose to calculate the cumulants from the partition function Z of Eq. (3.4) rather than the probability distribution function of Eq. (3.5) and the fourth-order cumulant ratio $U_4(L)$ is given by

$$U_4(L)=\langle E^4\rangle_L/\langle E^2\rangle_L=ZZ^{(4)}(Z^{(2)})^{-2} , \tag{3.10}$$

where $Z^{(n)}=\partial^n Z/\partial x^n$. To find the minimum of $V_4(L)$, we maximize $U_4(L)$ of Eq. (3.10) and, after some rather tedious but straightforward algebra (see the Appendix), we find

$$V_4(L)|_{\min}=\tfrac{2}{3}-(e_1/e_2-e_2/e_1)^2/12+A^{(2)}L^{-d} . \tag{3.11}$$

The coefficient $A^{(2)}$ is a rather complicated expression involving C_1, C_2, e_1, e_2, and β_c and is given by Eq. (A18). The temperature $T_c^{(2)}(L)$ at which $V_4(L)$ is a minimum is

$$[T_c^{(2)}(L)/T_c-1]L^d = T_c(e_2-e_1)^{-1}\ln q e_1^2/e_2^2+B^{(2)}L^{-d} , \tag{3.12}$$

where $B^{(2)}$ is given by Eq. (A19). Note that the leading terms of Eqs. (3.11) and (3.12) are exactly known for the q-state Potts models in $d=2$ since e_1, e_2, and T_c are known.[13] Since $A^{(2)}$ and $B^{(2)}$ depend on the bulk specific heats C_1 and C_2, they cannot be exactly calculated but depend on one unknown parameter which must be determined from simulations.

The specific heat per site $C(L)$ of a finite system of volume L^d is given by

$$\begin{aligned}C(L)L^{-d}&=\beta^2L^{-2d}(\langle E^2\rangle_L-\langle E\rangle_L^2)\\&=\beta^2e_1^2[Z^{(2)}/Z-(Z^{(1)}/Z)^2] ,\end{aligned} \tag{3.13}$$

where $\langle E\rangle$ is the average of the total energy and we find, for the specific-heat maximum,

$$\begin{aligned}C(L)L^{-d}|_{\max}&=(e_1-e_2)^2/4T_c^2\\&\quad+\tfrac{1}{2}L^{-d}[C_1+C_2\\&\quad+(C_2-C_1+\beta_c e_1-\beta_c e_2)\ln q] ,\end{aligned} \tag{3.14}$$

which occurs at $T_c^{(0)}(L)$ given by

$$[T_c^{(0)}(L)/T_c-1]L^d=T_c(e_2-e_1)^{-1}\ln q+B^{(0)}L^{-d} , \tag{3.15}$$

where $B^{(0)}$ is given by Eq. (A22). Note that the form of the L^{-d} correction to the specific-heat maximum differs from that of Challa *et al.*[5] Note also that the leading-order expressions for $T_c^{(0)}$ and $T_c^{(2)}$ differ from each other.

Similar results follow for the ratio of second cumulants

$$U_2(L)=\langle E^2\rangle_L/\langle E\rangle_L^2 ,$$

which have a maximum at $T_c^{(1)}(L)$,

$$[T_c^{(1)}(L)/T_c-1]L^d=T_c(e_2-e_1)^{-1}\ln q e_1/e_2+B^{(1)}L^{-d} \tag{3.16}$$

with a maximum value

$$U^{(2)}(L)|_{\max}=(e_1+e_2)^2/4e_1e_2+A^{(1)}L^{-d} . \tag{3.17}$$

$A^{(1)}$ and $B^{(1)}$ are given by Eqs. (A16) and (A17).

The form of the probability distribution $P(E)$ used by Challa *et al.*[5] will, in fact, give identical results to ours for the leading terms of the maxima of $C(L)L^{-d}$, $U_2(L)$, and $U_4(L)$ but different results for the L^{-d} corrections and for $T_c^{(i)}(L)$. We also calculate the complete scaling function for the specific heat

$$\begin{aligned}C(L)L^{-d}&=\frac{\beta_c^2(e_1-e_2)^2}{4\cosh^2[\frac{1}{2}(\beta_c-\beta)(e_1-e_2)L^d+\frac{1}{2}\ln q]}\\&\quad+O(L^{-d}) .\end{aligned} \tag{3.18}$$

This agrees with Ref. 5 at the maximum only and differs at all other temperatures. We note that, including the $O(L^{-d})$, corrections given by Eq. (A21) give excellent agreement with our simulation results for $q=10$ as do the expressions for $T_c^{(i)}(L)$.

Although all our simulations discussed in Sec. IV are for energy cumulants only, it is of some interest to calculate the order-parameter cumulants simulated in Ref. 5 such as

$$\tilde{V}_4(L)\equiv 1-\langle m^4\rangle_L/3\langle m^2\rangle_L^2 \ . \tag{3.19}$$

For the Potts models, we define the order parameter m by[5]

$$m=(qN_{\max}L^{-d}-1)/(q-1) \ , \tag{3.20}$$

where $N_{\max}=\max(N_1,N_2\cdots N_q)$ and N_i is the number of spins in the ith state. The partition function is now written as

$$Z=q\exp-\beta g_1(\beta,h)L^d+\exp-\beta g_2(\beta,h)L^d \tag{3.21}$$

with $g_i(\beta,h)$ the Gibbs potential of the ith bulk phase. Expanding $g_i(\beta,h)$ about the transition point $\beta=\beta_c$, $h=0$,

$$\beta g_i(\beta,h)=\beta_c g_i(\beta_c,0)-\beta_c e_i t-\tfrac{1}{2}C_i t^2-\beta_c m_i h + ht\beta_c m_i\left[1+\frac{\partial\ln m_i}{\partial\ln\beta}\bigg|_{\beta_c}\right]-\tfrac{1}{2}\beta_c\chi_i h^2 \ , \tag{3.22}$$

where m_i is the magnetization of the ith phase ($m_2=0$) and $\chi_i=\partial m_i/\partial h$ is the susceptibility of the ith bulk phase at the transition point. Using Eqs. (3.21) and (3.22) one may readily calculate the cumulants. We find that $\tilde{V}_4(L)$ has a minimum at $T_c(L)>T_c$ at

$$T_c(L)/T_c-1=(dL^{-d}\ln L)/\beta_c(e_2-e_1)+\cdots \tag{3.23}$$

with a value

$$\tilde{V}_4(L)=-(\beta_c m_1^2/4\chi_2)L^d \ . \tag{3.24}$$

This behavior may already have been seen in the simulations of Ref. 5 in Fig. 16 where a deep minimum is apparent. However, at $T>T_c(L)$, the simulations show that $\tilde{V}_4(L)>0$ while this simple theory predicts that it should tend to zero from below as $t(L)L^d\to\infty$. There are several effects not taken into account by the partition function of Eqs. (3.21) and (3.22) which may be responsible for this discrepancy.

The special case of a symmetric first-order transition between two phases, as in the field-driven transition in an Ising model at fixed $T<T_c$ and magnetic field $h\simeq 0$ studied extensively by Binder and Landau,[4] is a simpler special case and must be studied separately. The order parameters of the bulk phases at the transition are related by $m_1(\beta)=-m_2(\beta)=m$ and the susceptibilities $\chi_1(\beta)=\chi_2(\beta)\equiv\chi$ and the partition function of Eqs. (3.21) and (3.22) simplifies to

$$Z=2\cosh\beta mhL^d\exp\tfrac{1}{2}\beta\chi h^2L^d \ .$$

From this it is trivial to show that the finite-size susceptibility defined by

$$\chi(L)\equiv\beta L^{-d}\partial^2\ln Z/\partial(\beta h)^2$$

has a maximum value

$$\chi(L)L^{-d}|_{\max}=\beta m^2+\chi L^{-d}+O(L^{-2d})$$

at $h_c=0$. However, the fourth cumulant used in Ref. 4 defined by

$$\tilde{V}_4^{(c)}(L)\equiv-\langle S^4\rangle^c/3\langle S^2\rangle^2 \ , \tag{3.25}$$

where the connected cumulant

$$\langle S^4\rangle^{(c)}=\langle S^4\rangle-3\langle S^2\rangle^2-4\langle S\rangle\langle S^3\rangle+12\langle S^2\rangle\langle S\rangle^2-6\langle S\rangle^4$$

has nontrivial behavior at fixed $\beta>\beta_c$ as a function of h. At the bulk transition point $h=0$ it has the value $\frac{2}{3}-O\langle L^{-d}\rangle$ and vanishes as $L\to\infty$ in both bulk phases $h\neq 0$. For finite L, $\tilde{V}_4^{(c)}(L)$ has minima at

$$\pm h_c(\beta,L)=a_1L^{-d}/\beta m-a_2\chi L^{-2d}/\beta^2m^3 \ ,$$

where the numbers a_1 and a_2 can be calculated exactly by the methods of this paper to be $a_1=1.1462\ldots$ and $a_2=1.4720\ldots$. At the minima

$$\tilde{V}_4^{(c)}(L)|_{\min}=-\tfrac{2}{9}+a_3\chi L^{-d}/\beta m^2+O(L^{-2d}) \tag{3.26}$$

with $a_3=4[1+\sqrt{2/3}\ln(5+2\sqrt{6})]/9=1.2763\ldots$. It is interesting to note that the leading term $-\frac{2}{9}$ in Eq. (3.26) is exact and independent of the bulk thermodynamic parameters m and χ, in contrast to the expression of $V_4(L)$ of Eq. (3.11). Also, at low temperatures, $\beta>\beta_c$ in the Ising model $\chi(\beta)$ will also be small so that the $O(L^{-d})$ term of Eq. (3.26) will be numerically negligible. Inspection of Fig. 4 of Ref. 4 shows that the simulations agree very well with theory.

IV. SIMULATION RESULTS

A sufficiently large system undergoing a first-order transition in the $L\to\infty$ limit should be described in detail by the partition function of Eq. (3.4) and, in principle, the five thermodynamic parameters such as T_c, e_i, and C_i should be obtainable from simulations using the leading-order finite-size-scaling corrections of $O(L^{-d})$. We tested this by performing extensive MC simulations on the $q=8$ and 10 state Potts models in $d=2$ for which four of the five parameters are exactly known.[13,21,22] We restricted ourselves to system sizes $L\leq 50$ since, to obtain good statistics, the amount of computer time required becomes prohibitive for larger L. However, since both these models undergo a rather strong first-order transition, one might expect these system sizes to agree well with our simple-minded finite-size-scaling theory. To our surprise, we find that, for $q=10$, the agreement with theory is excellent but for $q=8$ there is essentially no agreement. The implications of this are that, for $q=8$, one requires much larger system sizes but, more seriously, to evaluate all the thermodynamic quantities from MC data seems to require a transition at least as strong as in the $q=10$ model.

We simulated the probability distribution of energy by

performing about 5×10^6 Monte Carlo steps (MCS) per site on the local IMB 3090. From the histograms of energy at carefully chosen temperatures near the bulk T_c at which the double-peak structure was very apparent, extrapolations in T were performed using the method due to Ferrenberg and Swendsen[17] to obtain, from the same data, probability distributions at nearby T. The validity of the extrapolation from T to $T+\Delta T$ depends on the variation in the Boltzmann factor $\sim\exp\Delta T L^d$ being sufficiently small compared to $\Omega(E)$, the number of states of energy E. Since $\Delta T\sim L^{-d}$ in the range of interest, the Boltzmann factor varied by about 10 while $\Omega(E)$ was typically $\sim10^4$. However, as L is increased, the system will stay longer in one of the bulk states so that, during a fixed number of MCS, it flips between the ordered and disordered states less frequently. The statistics become rather poor for $L=50$ for the $q=10$ model. This problem was alleviated to some extent by using the algorithm due to Swendsen and Wang[23] to reduce the time spent in one bulk state, but even this requires more computer time than was available to us to make a major improvement in the statistics for the largest systems. For example, for the $q=10$ model with $L=50$ there were less than ten flips between ordered and disordered states in 5×10^6 MCS. This causes a systematic error in $t(L)$ of about 3×10^{-4}. However, one of the major advantages of the Ferrenberg-Swendsen extrapolation method is that the systematic error at every temperature is about the same so that a curve for, say, the specific heat, is bodily translated in temperature. Curves for different L values may be compared by superposing them, allowing for a bodily shift of temperature.

We assessed the strength of the transition by the methods described in Sec. II demanding that, for all L, $\Delta F(L)>1$ and $\Delta F(L)L^{-d+1}$ is constant. In Fig. 1, we show one plot of $-\ln P(E)$ for $L=32\times32$ and $q=8$, and in Fig. 2 is shown $\Delta F(L)/L$ against L^{-1}. From these we see that the $q=10$ data, although rather noisy due to the small number of events at the minimum of $P(E)$, gives a constant $\Delta F/L$ to within errors for $L>18$. On the other hand, $\Delta F/L$ for $q=8$ has not settled down to a constant even for the largest L. Note that the errors for $q=8$ are

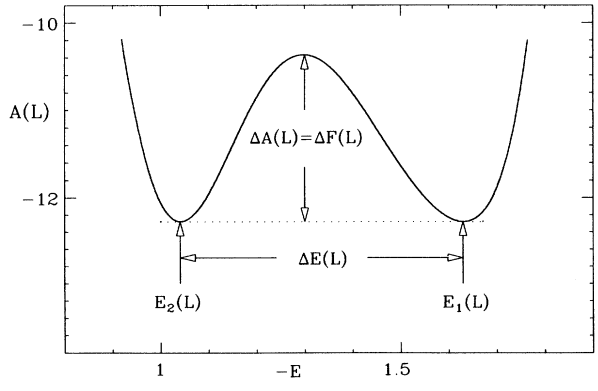

FIG. 1. A typical plot of $A(L)\equiv-\ln P_L(E)$ for the $q=8$ Potts model with $L=32$. Units of the energy per site E are such that $E=-2$ for a completely ordered state.

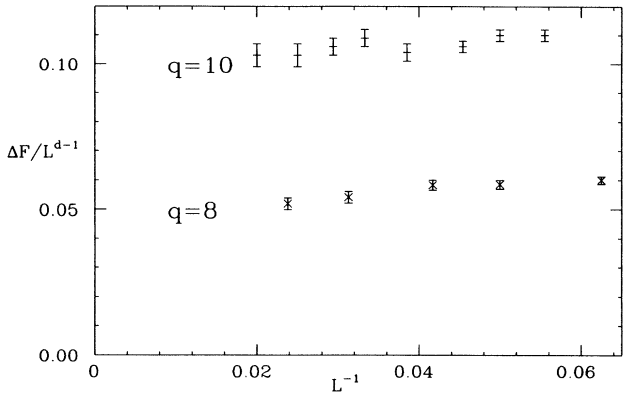

FIG. 2. The plots of barrier heights $\Delta F(L)L^{-d+1}$ for the $q=8$ and 10 Potts models.

smaller than for $q=10$ because of the greater number of flips and the consequent better statistics. As a final check, we plot in Fig. 3 the positions of the minima $E_i(L)$ of $-\ln P_L(E)$ for $q=10$. From Eq. (2.3) it is clear that

$$E_i(L)-e_i=O(L^{-1})$$

so that the positions of the minima $E_i(L)$ should tend to the bulk energies e_i linearly in L^{-1} if the systems are all in the strongly first-order region. From the linear extrapolation to the known values of e_i shown in Fig. 3, we conclude that this system should obey the finite-size-scaling formulas of the previous section. We note that the corresponding plots of the $q=8$ data (not shown) do not lie on straight lines for the L values simulated and we do not expect this system to agree with the theoretical formulas.

The ferromagnetic q-state Potts model is defined by the Hamiltonian

$$H/k_BT=-\beta\sum_{\langle i,j\rangle}\delta(\sigma_i,\sigma_j)\,, \tag{4.1}$$

where $\delta(\sigma_1,\sigma_2)$ is a Kronecker δ function, $\beta=1/k_BT$, and $\langle ij\rangle$ denotes nearest-neighbor pairs. Baxter[21] has

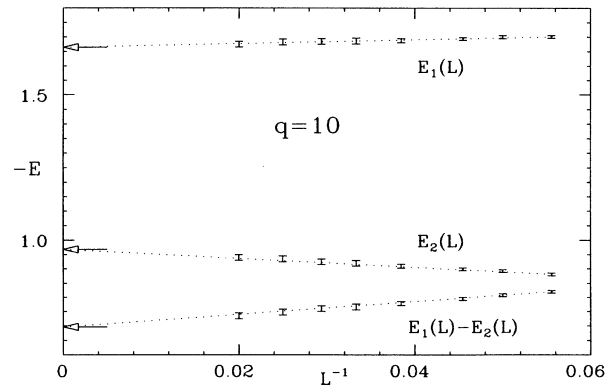

FIG. 3. The energy minima $E_i(L)$ of $-\ln P_L(E)$. Arrows denote exact values at $L=\infty$.

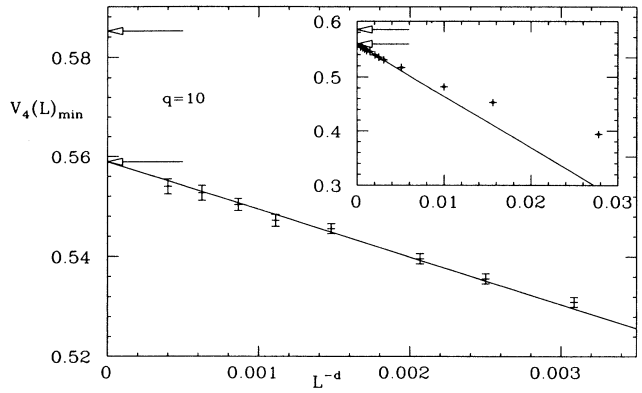

FIG. 4. The fourth-order cumulant ratio minima $V_4(L)|_{\min}$ for $q=10$. The lower arrow is the exact value from Eq. (3.9) and the upper arrow is from Eq. (3.6). The solid line is the best straight-line fit according to Eq. (3.11). The inset shows data for $L \geq 6$ and the lack of scaling for $L < 18$.

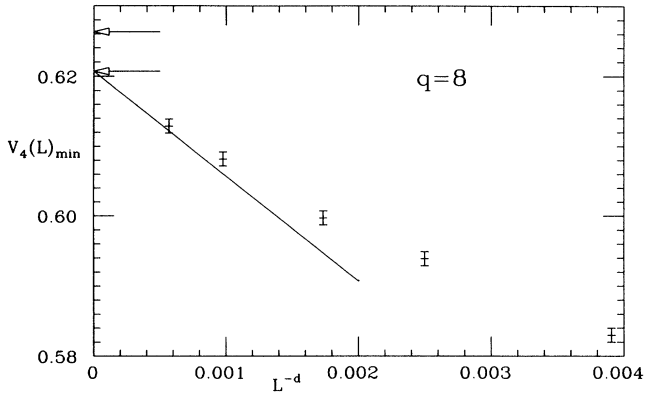

FIG. 5. $V_4(L)|_{\min}$ for $q=8$. The solid straight line is drawn according to procedure described in the text.

shown that the model has a continuous transition for $q \leq 4$ and a first order for $q > 4$ in two dimensions. Several physical quantities are known exactly[13,21,22] at T_c in units of $k_B = 1$ and for completeness we list them below:

$$\beta_c = \ln(1+\sqrt{q}) , \tag{4.2a}$$

$$e_2 - e_1 = 2(1+q^{-1/2})\tanh\tfrac{1}{2}\theta \prod_{n=1}^{\infty} \tanh^2 n\theta , \tag{4.2b}$$

$$2\cosh\theta = q^{1/2} , \tag{4.2b}$$

$$e_1 + e_2 = -2(1+q^{-1/2}) , \tag{4.2c}$$

$$C_2 - C_1 = (e_2 - e_1)\beta_c^2 q^{-1/2} . \tag{4.2d}$$

Numerical values for the $d=2$, $q=8$ and 10 models are listed in Table I. Note that $C_2 - C_1 < 0.5$ in these units, whereas our best estimates give the individual C_i an order of magnitude larger which greatly simplifies the analysis of the data. The only undetermined parameter is (say) C_1 so these models are an excellent testing ground to compare simulations with theory.

We first investigate the scaling of the minimum of Binder's fourth cumulant[16] and in Figs. 4 and 5 are shown plots of $V_4(L)|_{\min}$ against L^{-d}. For $q=10$, the data for $18 \leq L \leq 50$ fits very well to a straight line and extrapolates well to the theoretical $L=\infty$ value of Eq. (3.9) denoted by the lower arrow in Fig. 4 and is clearly inconsistent with the value of Eq. (3.6) shown by the upper arrow. In the inset of Fig. 4 is shown the simulation results including system sizes $L < 18 \simeq 2\xi(10)$, for which $\Delta F(L) < 1$. As expected, the small systems do not obey the finite-size-scaling predictions. The solid line is the best fit to the data with a slope of -9.5 ± 1 obtained by demanding it passes through the known $L=\infty$ value. Using Eqs. (4.2d) and (A18), this value of the slope yields values for the bulk specific heats

$$C_1 = 10.7 \pm 1.0, \quad C_2 = C_1 + 0.45 . \tag{4.3}$$

In contrast to the ten-state data, the $q=8$ data shown in Fig. 5 do not scale as L^{-d} for the values of L used in the simulations and it is not possible to distinguish between the theoretical values for $V_4(\infty)|_{\min}$ of Eqs. (3.6)

TABLE I. Previously known exact results in units in which $J=1=k_B$ and some thermodynamic quantities evaluated from the theory and simulations of this paper.

	$q=10$	$q=8$
e_1	$-1.664\,25 \cdots$ [a]	$-1.596\,73 \cdots$ [a]
e_2	$-0.968\,20 \cdots$ [a]	$-1.110\,37 \cdots$ [a]
$r=e_2/e_1$	$0.581\,76 \cdots$	$0.695\,40 \cdots$
T_c	$0.701\,23 \cdots$ [a]	$0.744\,90 \cdots$ [a]
$V_4(\infty)\|_{\min}=\frac{2}{3}-(r-r^{-1})^2/12$	$0.5589 \cdots$ [b]	$0.6207 \cdots$ [b]
$U_2(\infty)\|_{\max}=(1+r)^2/4r$	$1.0751 \cdots$ [b]	$1.0333 \cdots$ [b]
$C_2-C_1=(e_2-e_1)/T_c^2\sqrt{q}$	$0.4476 \cdots$ [a]	$0.3098 \cdots$ [a]
C_1 [from $V_4(L)$]	10.6 ± 1 [b]	22.8 ± 3 [b]
C_1 [from $U_2(L)$]	10.7 ± 1 [b]	23.2 ± 3 [b]
$\lim[C(L)-AL^d]$	10.3 ± 1 [b]	22.8 ± 3 [b]
$(T_c^{(2)}-T_c^{(1)})L^d$	$0.3826+7.8L^{-d} \cdots$ [b]	

[a]From Refs. 21 and 22.
[b]This work.

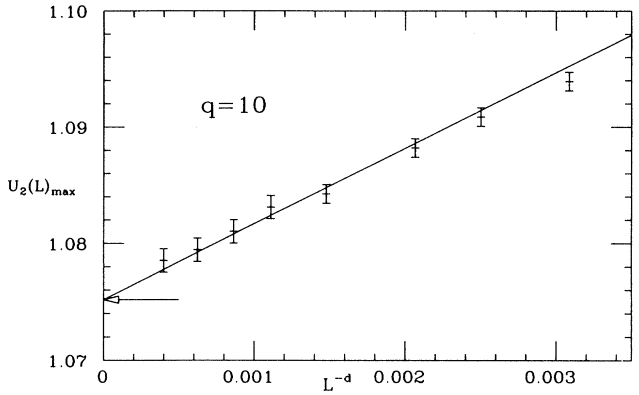

FIG. 6. The maxima of $U_2(L)\equiv\langle E^2\rangle_L/\langle E\rangle_L^2$ for $q=10$. The arrow at $L=\infty$ is the exact value $(e_1+e_2)^2/4e_1e_2$ and the solid line is the best fit according to Eq. (3.17).

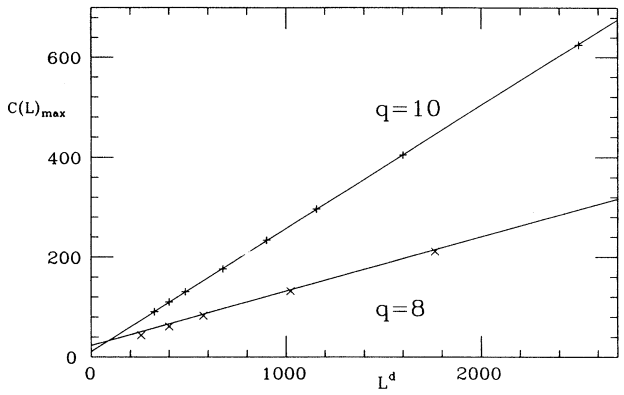

FIG. 8. $C(L)|_{max}$ as function of L^d for $q=8$ and 10. The solid lines are drawn with theoretical slope $(e_2-e_1)^2/4T_c^2$ and the intercepts according to Eq. (3.14) using our best estimates of C_1 and C_2.

and (3.9) from the data. In view of our estimates of the correlation length $\xi(8)\simeq 22$, this is not surprising since the largest system $L=42$ for $q=8$ corresponds roughly to $L=18$ for $q=10$ and the data behaves similarly to the small-L data of the latter as shown in the inset of Fig. 4. We can obtain an estimate for $C_i(q=8)$ by drawing the solid line in Fig. 5 through its known value at $L=\infty$ and the $L=42$ point yielding a slope of -15 ± 2. This yields

$$C_1(8)=23\pm 3,\quad C_2(8)=C_1(8)+0.31\ . \tag{4.4}$$

The maximum of the second cumulant

$$U_2(L)=\langle E^2\rangle_L/\langle E\rangle_L^2$$

shows similar behavior and is shown in Fig. 6 for $q=10$ and Fig. 7 for $q=8$. As expected, the $q=10$ data scales well with L^{-d} and extrapolates to the theoretical $L=\infty$ value of $(e_1+e_2)^2/4e_1e_2$ [Eq. (3.17)]. Using the same procedure as above, the best straight-line fit has a slope of 6.5 ± 0.5 which yields a value of $C_1=10.7\pm 0.9$ with the aid of Eq. (A16). The $q=8$ data of Fig. 7 shows distinct curvature but, using the same reasoning as before, the same estimate of C_1 as Eq. (4.4) is obtained.

In principle, very accurate values of C_1 and C_2 could be obtained by this procedure but very accurate estimates of the L^{-d} corrections are required. This would need much larger system sizes for the $q=8$ model, up to $L=120$, and better statistics for the $q=10$ model, both of which imply a much larger investment in computer time that we could afford.

We now turn to a discussion of the specific-heat simulations which we analyzed and compared to theoretical predictions in considerably more detail than for the cumulant ratios V_4 and U_2. The specific-heat maximum is predicted to scale as L^d with a known coefficient given by Eq. (3.14) and our data is shown in Fig. 8. The solid lines are theoretical fits using the known values of the slopes $(e_1-e_2)^2\beta_c^2/4$ and the intercepts at $L=0$ from Eq. (3.14) and our previous estimates of $C_{1,2}$. Again, the ten-state data are very good fits but the eight-state data show definite deviations from theory at small L as expected. A more detailed fit of the data is shown in Fig. 9 in which the exactly known L^d contribution to $C(L)|_{max}$ is subtracted and the constant (according to theory) residue

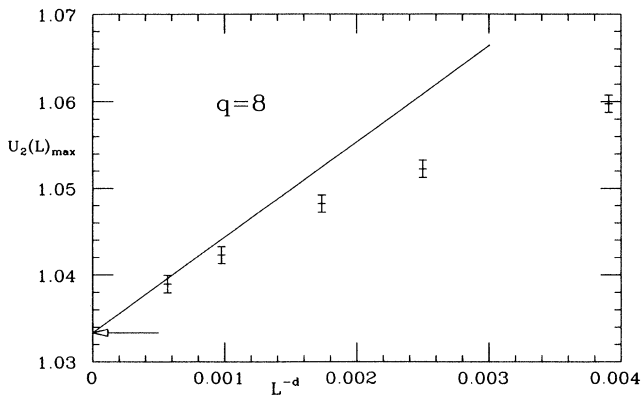

FIG. 7. $U_2(L)|_{max}$ for $q=8$. The solid line is drawn according to procedure described in the text.

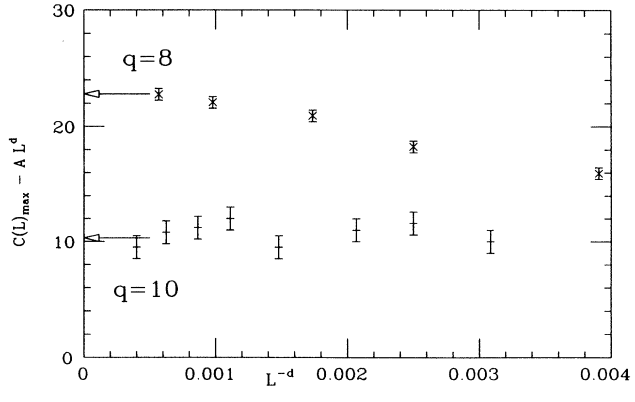

FIG. 9. The plot of $C(L)|_{max}-(e_2-e_1)^2L^d/4T_c^2$. The arrows at $L=\infty$ are according to Eq. (3.14) using our best estimates of $C_1(8)=23$, $C_1(10)=10.7$.

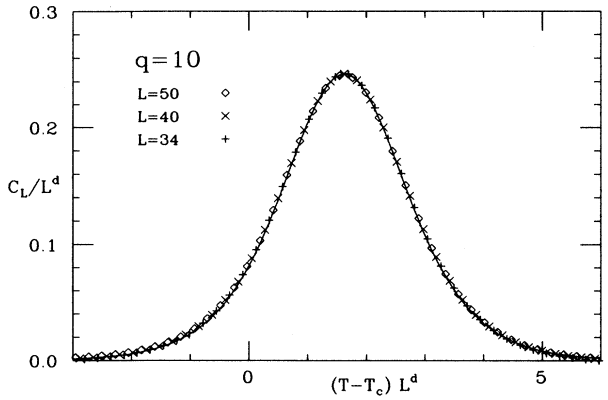

FIG. 10. The plot of the specific heat $C(T,L)L^{-d}$ against the scaling variable $(T-T_c)L^d$ for $q=10$ with system sizes $L=50$, 40, and 34. The solid curve is the theoretical curve of Eq. (3.18) at $L=\infty$. The data points are from simulations with the L^{-d} correction according to Eq. (A21) removed using the best estimates of C_1 (see the text).

plotted against L^{-d}. Because of the lack of statistics for the $q=10$ model, there is some numerical uncertainty here, but, to within numerical error, this is constant and agrees with Eq. (A23) using the estimates of C_i of Eq. (4.3). The $q=8$ data have better statistics but are not in the scaling regime. The arrows indicate our estimates of

$$C(L)|_{\max}-\beta_c^2(e_1-e_2)^2L^d/4$$

at $L=\infty$.

A more detailed comparison of the simulation data for the full temperature-dependent specific heat is shown in Fig. 10 for $q=10$ and Fig. 11 for $q=8$ in which $C(L)L^{-d}$ is plotted against the scaling variable $(T-T_c)L^d$. The theoretical prediction for $C(L)L^{-d}$is given by Eq. (A21) which may be written as

$$C(L)L^{-d}=\tfrac{1}{4}\beta_c^2(e_1-e_2)^2\mathrm{sech}^2[\tfrac{1}{2}(e_1-e_2)(\beta_c-\beta)L^d + \tfrac{1}{2}\ln q] + L^{-d}A[(\beta_c-\beta)L^d], \quad (4.5)$$

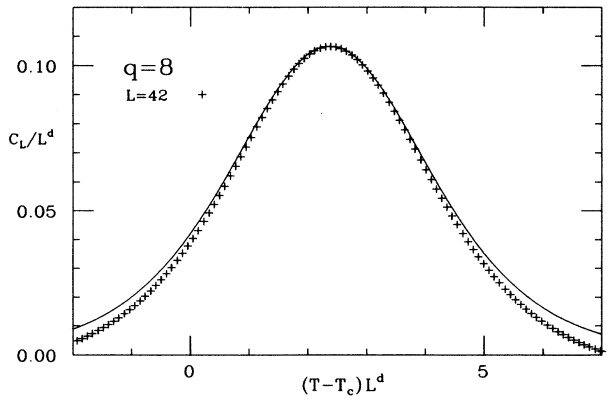

FIG. 11. The specific heat $C(T,L)L^{-d}$ for $q=8$. The solid curve is the theoretical $L=\infty$ curve and the crosses are the corrected data points for $L=42$.

where the $O(L^{-d})$ term is a complicated expression given by Eq. (A21). The solid line of Fig. 10 is the $L=\infty$ limit of Eq. (4.5) given by the first term only and the data points for $L=50$, 40, and 34 are obtained by subtracting from each raw data point the $O(L^{-d})$ correction as computed from Eq. (A21). Each curve for the different L values is obtained by extrapolation in temperature from a single simulation for each L at one temperature.[17] As discussed earlier, this procedure leads to a systematic error in temperature of $O(10^{-4})$ and, to eliminate this, each curve was shifted bodily in temperature to attempt to superimpose it on the theoretical $L=\infty$ curve. The result of this procedure causes a very good data collapse in excellent agreement with theory. The identical procedure for $q=8$ was carried out for only one system size, $L=42$, and, as expected, there are significant deviations from the theoretical $L=\infty$ curve.

The final test of the theory of Sec. III is a study of the size dependence of the pseudocritical temperatures $T_c^{(i)}(L)$ where the cumulants $C(L)L^{-d}$, $U_2(L)$, and $U_4(L)$ have maxima. These all tend to the bulk critical temperature T_c as $L\to\infty$ with different slopes given by Eqs. (3.12), (3.15), and (3.16) which may be summarized by

$$T_c^{(i)}(L)/T_c-1=L^{-d}T_c(e_2-e_1)^{-1}\ln q(e_1/e_2)^i + B^{(i)}L^{-2d}, \quad (4.6)$$

where $i=0,1,2$ and $B^{(i)}$ are given by Eqs. (A22), (A19), and (A17). The comparison between the theoretical prediction of Eq. (4.6) and the simulations for $q=10$ is shown in Fig. 12 and for $q=8$ in Fig. 13. To within numerical uncertainty, the data points lie on very good straight lines when $T_c^{(i)}(L)$ is plotted against L^{-d} and extrapolate to the bulk $T_c=1/\ln(1+\sqrt{q}\,)$ for both $q=10$ and 8. However, the leading-order corrections in L^{-d}

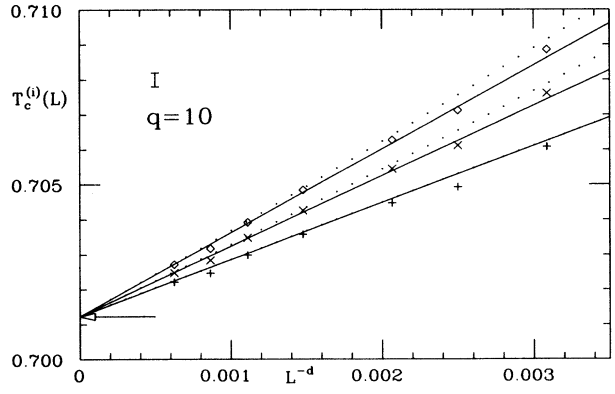

FIG. 12. The pseudocritical temperatures $T_c^{(i)}(L)$ for the $q=10$ model. The arrow is exact $T_c(\infty)=1/\ln(1+\sqrt{q}\,)$. The plus signs (+) denote $T_c^{(0)}(L)$, maxima of $C(L)L^{-d}$; the crosses (×) denote $T_c^{(1)}(L)$, maxima of $\langle E^2\rangle_L/\langle E\rangle_L^2$; the open diamonds (◇) denote $T_c^{(2)}(L)$, maxima of $\langle E^4\rangle_L/\langle E^2\rangle_L^2$. The solid lines are the exact theoretical leading-order L^{-d} corrections of slope $T_c^2(e_2-e_1)^{-1}\ln q(e_1/e_2)^i$ and the dotted lines include the theoretical L^{-2d} corrections of the Appendix. These are ordered bottom ($T_c^{(0)}$) to top ($T_c^{(2)}$).

are known exactly in Eq. (4.6) and are shown as solid lines in Figs. 12 and 13. The next order corrections using our best estimates for the specific heats are shown as dotted lines and are about 6% for the smallest L for $q=10$ but as large as 25% for $q=8$. The agreement between theory and simulations is very good for $q=10$ but nonexistent for $q=8$ despite the linear extrapolation to T_c in L^{-d} because our L values are too small for comparison with strong first-order finite-size scaling in this case. An inspection of the $-\ln P(E)$ curves of Fig. 1 shows that they are highly asymmetric. Since the bulk specific heats C_1 and C_2 are controlled by the curvature at the minima and $C_2 > C_1$ from known exact results,[22] as $L \to \infty$ the ordered phase minimum should be the narrower. However, for $L < 42$, the relative width of the minima is reversed which has the effect of lowering $T_c^{(i)}(L)$ below its theoretical value. However, the contribution to the statistical sums from the region between the minima is quite large, the effect of which is not taken into account by our theoretical analysis.

The absolute values of the data points for $T_c^{(i)}(L)$ are subject to some systematic errors due to the limited statistics of the histograms. However, because $T_c^{(0)}(L)$, $T_c^{(1)}(L)$, and $T_c^{(2)}(L)$ for the same L are generated from the same data extrapolated to the appropriate values of T, much of these systematic errors will be eliminated by considering the differences

$$[T_c^{(i)}(L) - T_c^{(j)}(L)]L^d .$$

In Fig. 14 we plot $(T_c^{(2)} - T_c^{(1)})L^d$ which, as $L \to \infty$, tends to the known value

$$T_c^2(e_2 - e_1)^{-1} \ln e_1/e_2 \simeq 0.3827$$

and the slope from Eqs. (A17) and (A19), drawn as the straight line. Although the errors in Fig. 14 appear to be very large, there is a factor of L^d multiplying the temperature differences and the error in $\Delta T L^d$ is only about 3%. The agreement between the simulation data and theory for this L^{-2d} correction may be taken as very satisfactory.

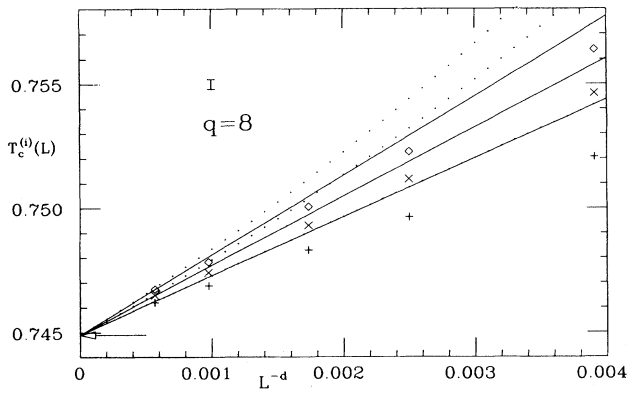

FIG. 13. The pseudocritical temperatures for the $q=8$ model. The data points and theoretical curves are as for Fig. 12. Note that the data points lie on good straight lines extrapolating to known T_c but the slopes disagree badly with theory.

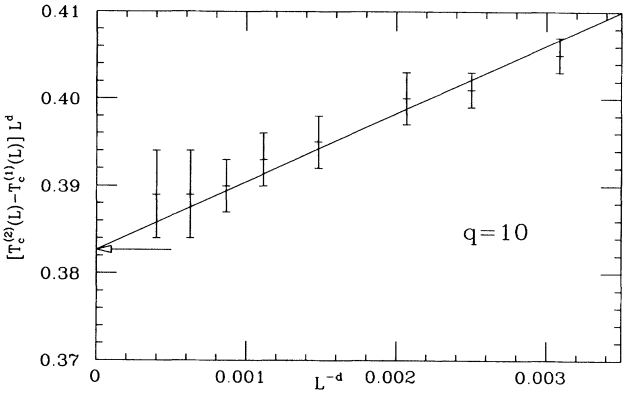

FIG. 14. The plot of $[T_c^{(2)}(L) - T_c^{(1)}(L)]L^d$ to estimate the L^{-2d} corrections to $T_c(L)$ for the $q=10$ model. The arrow is the exactly known value $T_c^2(e_2 - e_1)^{-1} \ln e_1/e_2 = 0.382\,67$ and the solid line is the theoretical line according to Eqs. (A17) and (A19). Error bars seem large because of the expanded scale and correspond to the uncertainty in $\Delta T L^2$ of 3% (see the text).

V. CONCLUSIONS

We have developed a detailed theory of finite-size scaling for systems with periodic BC which undergo a strong first-order transition and calculated the scaling form for various thermodynamic quantities U, such as the specific heat, to $O(L^{-d})$ in the form

$$U(T,L) = U_0(tL^d) + L^{-d} U_1(tL^d) ,$$

where $t \equiv 1 - T_c/T$. For the Potts models in two dimensions, the leading term $U_0(tL^d)$ is known in terms of the exactly known quantities e_1, e_2, and T_c while the next correction $U_1(tL^d)$ also depends on the bulk specific heats C_1 and C_2 of the ordered and disordered phases. The ratio of fourth-order cumulants of energy introduced by Binder[16] as a quantity to distinguish between a first-order and continuous transition is reconsidered and a corrected value for the minimum calculated. Explicit expressions for various pseudocritical temperatures $T_c^{(i)}(L)$ at which the quantities $U(T,L)$ have maxima are given and it is found that they all differ, depending on the definition of $U(L)$. Deviations from the bulk T_c are computed to $O(L^{-2d})$. The corresponding cumulants of the order parameter in the Potts models and in the symmetric field-driven transition in an Ising model are also briefly discussed and some features of earlier simulations explained.

Extensive MC simulations on the two-dimensional $q=8$ and 10 Potts models for system sizes up to 50×50 and a detailed comparison with finite-size-scaling theory are reported. Although both models supposedly have very strong first-order transitions, for these sizes, the agreement with theory is good only for $q=10$ but not for $q=8$. The lack of agreement for the latter model is explained in terms of our simulations of the energy distribution function $P(E)$ from which all quantities are calculated. These latter simulations allow for an immediate as-

sessment of the validity of finite-size scaling by studying the height of the free-energy barrier between the bulk states. We find that this barrier $\Delta F(L)$ must be both numerically much larger than unity and also must scale as L^{d-1}. The ten-state model obeys both criteria and agrees well with theory while the eight-state model violates both and disagrees with theory for our system sizes. We note that the scaling of several quantities should be studied because, although the eight-state model seems to obey $T_c(L)-T_c \sim L^{-d}$ very well, the coefficient is incorrect and this apparent agreement with finite-size scaling is accidental. Estimates for the bulk specific heats are also obtained from the simulations by comparing with theoretical expressions. We also note that, to evaluate all bulk quantities such as T_c, e_i, and C_i from simulation data alone, it is essential to verify that one is in the strong first-order scaling regime and that one has very accurate data with excellent statistics.

Several of our theoretical results were independently obtained by Borgs *et al.*[10] from a more rigorous point of view. They restricted themselves to a calculation of the leading finite-size-scaling expressions $U_0(tL^d)$ and our results are identical to theirs. To compare with simulation data, the next order correction $U_1(tL^d)$ improves the agreement considerably and we have checked in several cases that the numerical data and the theory for U_1 agree with each other.

There is one discrepancy between theory and simulation that is still a puzzle despite the excellent agreement for the thermodynamic averages. In Sec. III, we derived the probability distribution of energy $P(E)$ of Eq. (3.5) which is exactly equivalent to the partition function of Eq. (3.4) from which all theoretical expressions are derived. The minima of $-\ln P(E)$ are at $e_1+O(\exp -L^d)$ and $e_2-O(\exp -L^d)$. However, from the discussion of Sec. II one concludes that the minima are at $E_1(L)=e_1-A_1L^{-1}$ and $E_2(L)=e_2+A_2L^{-1}$ with $A_{1,2}>0$ which agrees rather well with our simulation data of Fig. 3 in which it is clear that the $E_i(L)$ certainly do not scale as $\exp -L^d$ but rather as L^{-1}, at least for the L values of the simulations. This discrepancy between the direction of the finite-size shifts and the L dependence of these shifts is unexplained. Presumably this is due to surface or domain-wall corrections in the simulations which are not included in the analytic theory of Sec. III. However, the thermodynamic averages $U_i(L)$ as calculated from the simulated distribution of energy $P_S(E)$ agree well with those calculated from the analytic distribution $P_A(E)$ of Eq. (3.5) despite the rather significant differences in detailed shapes of $P_S(E)$ and $P_A(E)$. The main ones are in the positions of the minima and in the curvatures at the minima which seem to behave oppositely in the two cases. We have no understanding of how the overall behaviors of $P_S(E)$ and $P_A(E)$ conspire to give such good agreement for the cumulants $U_i(L)$ and leave this point as an interesting puzzle.

ACKNOWLEDGMENTS

The authors thank T. Ala-Nissila, E. Granato, W. K. Han, and S. C. Ying for useful and informative discussions and C. Borgs for making his work available before publication. They would also like to thank R. H. Swendsen for explaining some of the history of the histogram method. This research was supported by the National Science Foundation (NSF) under Grant No. DMR-8918358.

APPENDIX

This appendix is devoted to the simple but tedious calculations involved in computing the finite-size corrections to the various quantities discussed in the body of the paper:

$$C(L)\equiv(\langle E^2\rangle_L-\langle E\rangle_L^2)/T^2L^d\ ,$$
$$U_2(L)\equiv\langle E^2\rangle_L/\langle E\rangle_L^2\ , \tag{A1}$$
$$U_4(L)\equiv\langle E^4\rangle_L/\langle E^2\rangle_L^2\ .$$

We assume that the system is in the strong first-order regime $L\gg\xi$, where the partition function is a superposition of partition functions of the $q+1$ bulk phases in the vicinity of the bulk transition temperature T_c. In this region, we may write

$$Z=\sum_k \exp -\beta f_k(\beta)L^d\ , \tag{A2}$$

where $f_k(\beta)$ is the free energy of the kth bulk phase, L is the system size, and d is the spatial dimension. We take the free energies to be analytic expansions about β_c,

$$\beta f_k(\beta)=\beta_c f_k(\beta_c)-\beta_c e_k t-\tfrac{1}{2}C_k t^2+O(t^3)\ , \tag{A3}$$

where

$$t=1-\beta/\beta_c=1-T_c/T\ .$$

e_k and C_k are the energy and specific heat of the kth bulk phase

$$e_k=\partial(\beta f_k)/\partial\beta|_{\beta_c},\quad C_k=-\beta_c^2\partial e_k/\partial\beta|_{\beta_c}\ . \tag{A4}$$

In the above we have used units of $k_B=1$. Since the rounding of the transition takes place over a range $t\sim L^{-d}$, we rewrite the partition function in terms of

$$x=\beta_c e_1 tL^d,\quad a=L^{-d}C_1/e_1^2\beta_c^2\ ,$$
$$r=e_2/e_1<1,\quad y=C_2/C_1>1\ , \tag{A5}$$

where e_1 and C_1 are the energy and specific heat of the q equivalent ordered states and e_2 and C_2 are for the disordered state. In terms of these variables, the partition function becomes

$$Z=qe^{x+(1/2)ax^2}+e^{rx+(1/2)ayx^2}$$
$$=qe^x+e^{rx}+\tfrac{1}{2}ax^2(qe^x+ye^{rx})+O(a^2)\ . \tag{A6}$$

Since $x=O(1)$ and $a=O(L^{-d})$, we may calculate the $O(L^{-d})$ corrections to $C(L)$, etc., by expanding to $O(a)$.

We first consider the second energy cumulant

$$U_2(L)=\langle E^2\rangle_L/\langle E\rangle_L^2=ZZ^{(2)}/(Z^{(1)})^2\ , \tag{A7}$$

where $Z^{(n)}=\partial^n Z/\partial x^n$. Expanding to $O(L^{-d})$, we may

write

$$U_2(x)=g(x)[1+ah(x)]\ , \tag{A8}$$

where

$$g(x)=(qe^x+e^{rx})(qe^x+r^2e^{rx})(qe^x+e^{rx})^{-2}\ , \tag{A9}$$

$$h(x)=\tfrac{1}{2}x^2\left[\frac{qe^x+ye^{rx}}{qe^x+e^{rx}}+\frac{qe^x+yr^2e^{rx}}{qe^x+r^2e^{rx}}-\frac{2(qe^x+yre^{rx})}{qe^x+re^{rx}}\right]+2x\left[\frac{qe^x+yre^{rx}}{qe^x+e^{rx}}-\frac{qe^x+ye^{rx}}{qe^x+re^{rx}}\right]+\frac{qe^x+ye^{rx}}{qe^x+r^2e^{rx}}\ . \tag{A10}$$

We wish to find the maximum of $U_2(x)$ which occurs at $x=x_c$. To find these we maximize Eq. (A8) with respect to x:

$$U_2'(x)=g'(x)+a[g'(x)h(x)+g(x)h'(x)]=0\ . \tag{A11}$$

In the thermodynamic limit $L=\infty$, $a=0$ at which point

$$g'(x_0)=0 \text{ and } g'(x)=g''(x_0)(x-x_0)\ .$$

Then, from Eq. (A11) we obtain

$$x_c-x_0=-ag(x_0)h'(x_0)/g''(x_0)+O(a^2) \tag{A12}$$

and

$$U_2(x)|_{\max}=g(x_0)[1+ah(x_0)]+O(a^2)\ . \tag{A13}$$

From Eq. (A9) we find that, in the thermodynamic limit $L\to\infty$,

$$x_0^{(1)}=-(\ln q/r)/(1-r)<0 \tag{A14}$$

and

$$U_2(x_0)=(1+r)^2/4r\ . \tag{A15}$$

The $O(L^{-d})$ corrections are very tedious but straightforward to evaluate and the results are

$$U_2(x,L)|_{\max}=\frac{(1+r)^2}{4r}+\frac{a(1+r)}{4r^2}[(y-r)\ln q/r+r+y]\ , \tag{A16}$$

$$x_c^{(1)}-x_0^{(1)}=-\frac{a}{(1-r)^3}\left[\tfrac{1}{2}(1-y)\left[\ln\frac{q}{r}\right]^2+\frac{(y-r)}{r}\left[2(1+r)+(1-r)\ln\frac{q}{r}\right]+\frac{2(y-r^2)}{r}\right]\ . \tag{A17}$$

Using the identical procedure to evaluate the minimum of Binder's fourth cumulant, we find, for $V_4=1-\langle E^4\rangle/3\langle E^2\rangle^2$,

$$V_4(x,L)|_{\max}=\tfrac{2}{3}-\frac{1}{12}\left[r-\frac{1}{r}\right]^2-\frac{a}{6}\left[1+\frac{1}{r^2}\right]\left[\left[1+\frac{1}{r}\right](y-r)\ln\frac{q}{r^2}+3(1+y)-\tfrac{1}{2}(y+r^2)\left[1+\frac{1}{r^2}\right]\right]\ , \tag{A18}$$

which occurs at $x=x_c^{(2)}$, where

$$x_c^{(2)}-x_0^{(2)}=-\frac{a}{(1-r)^3}\left[\frac{1}{2}(1-y)\left[\ln\frac{q}{r^2}\right]^2+\frac{2(y-r)}{r(1+r)}\left[(1-r^2)\ln\frac{q}{r^2}+2+2r^2\right]+\frac{y-r^2}{(1+r)^2}\left[10-\frac{1}{r^2}-r^2\right]\right] \tag{A19}$$

and

$$x_0^{(2)}=-(\ln q/r^2)/(1-r)\ . \tag{A20}$$

The last quantity we evaluate is the specific heat:

$$C(x,L)L^{-d}=\frac{\beta_c^2(e_1-e_2)^2qe^{(r+1)x}}{(qe^x+e^{rx})^2}+\frac{C_1}{L^d}\frac{qe^{(r+1)x}}{(qe^x+e^{rx})^2}\times\left[\frac{1}{2}x^2\frac{(y-1)(1-r)^2(qe^x-e^{rx})}{(qe^x+e^{rx})}+2x[(y-1)(r-1)-(1-r)^2e_1\beta_c/C_1]+\frac{(qe^x+e^{rx})(qe^x+ye^{rx})}{qe^{(r+1)x}}\right]\ . \tag{A21}$$

This gives the full scaling form of the specific heat which we use to obtain the plots shown in Figs. 10 and 11. The maximum of $C(x,L)L^{-d}$ occurs at

$$x_c^{(0)} = -\frac{\ln q}{1-r} - 2a\left[\frac{(y-1)}{(1-r)^3}[3-\tfrac{1}{4}(\ln q)^2] + \frac{2\beta_c e_1}{C_1(1-r)^2}\right] \tag{A22}$$

with a value

$$C(x,L)L^{-d}|_{\max} = \tfrac{1}{4}\beta_c^2(e_1-e_2)^2 + \tfrac{1}{2}L^{-d}[(C_2-C_1+\beta_c e_1-\beta_c e_2)\ln q + C_2+C_1]\,. \tag{A23}$$

Note that, in the thermodynamic limit, this is related to the latent heat e_2-e_1 but the correction term depends not only on the combination C_1+C_2 but also on the difference C_1-C_2. Note also that Eqs. (A17), (A19), and (A22) give corrections to $\beta_c(L)-\beta_c(\infty)$ up to $O(L^{-2d})$ which may be obtained from an expansion of the partition function up to $O(a)$ because the leading $O(L^{-d})$ correction to $\beta_c(L)$ is obtained from the leading $O(1)$ expressions for $x_0(L)$.

[1]Y. Imry, Phys. Rev. B **21**, 2042 (1980).

[2]M. E. Fisher and A. N. Berker, Phys. Rev. B **26**, 2507 (1982): V. Privman and M. E. Fisher, J. Stat. Phys. **33**, 385 (1983); Phys. Rev. B **32**, 447 (1985).

[3]M. N. Barber, in *Phase Transitions and Critical Phenomena,* edited by C. Domb and J. L. Lebowitz (Academic, New York, 1983), Vol. 8, p. 145.

[4]K. Binder and D. P. Landau, Phys. Rev. B **30**, 1477 (1984).

[5]M. S. S. Challa, D. P. Landau, and K. Binder, Phys. Rev. B **34**, 1841 (1986).

[6]K. Binder, Rep. Prog. Phys. **50**, 783 (1987).

[7]For recent reviews, see, *Finite Size Scaling and Numerical Simulation of Statistical Systems,* edited by V. Privman (World Scientific, Singapore, 1990).

[8]C. Borgs and J. Imbrie, Commun. Math. Phys. **123**, 305 (1989).

[9]C. Borgs and R. Kotecki (unpublished).

[10]C. Borgs, R. Kotecki, and S. Miracle-Sole (unpublished).

[11]P. Peczak and D. P. Landau, Phys. Rev. B **39**, 11 932 (1989).

[12]J. F. McCarthy, Phys. Rev. B **41**, 9530 (1990).

[13]R. B. Potts, Proc. Cambridge Philos. Soc. **48**, 106 (1952); F. Y. Wu, Rev. Mod. Phys. **54**, 235 (1982).

[14]J. Lee and J. M. Kosterlitz, Phys. Rev. Lett. **65**, 137 (1990).

[15]J. Lee and J. M. Kosterlitz, Phys. Rev. B (to be published).

[16]K. Binder, Z. Phys. B **43**, 119 (1981).

[17]A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **61**, 2635 (1988).

[18]The histogram method of analyzing simulation data has a long history. Some early work was done by Z. W. Salsburg, J. D. Jackson, W. Fickett, and W. W. Wood, J. Chem. Phys. **30**, 65 (1959); D. A. Chestnut and Z. W. Salsburg, *ibid.* **38**, 2861 (1963). To our knowledge, the first use of extrapolation methods was by I. R. McDonald and K. Singer, Discuss. Faraday Soc. **43**, 40 (1967); J. Chem. Phys. **47**, 4766 (1967). The method was rediscovered, refined, and extended in various ways: J. Valleau and D. N. Card, J. Chem. Phys. **57**, 5457 (1972); G. Torrie and J. P. Valleau, Chem. Phys. Lett. **28**, 578 (1974); M. Falconi, E. Marinari, M. L. Paciello, G. Parisi, and B. Taglienti, Phys. Lett. **108B**, 331 (1982); E. Marinari, Nucl. Phys. **B235**, 123 (1984); G. Bhanot, S. Black, P. Carter, and R. Salvador, Phys. Lett. B **183**, 331 (1987); G. Bhanot, K. M. Bitar, and R. Salvador, *ibid.* **188**, 246 (1987); G. Bhanot, R. Salvador, S. Black, P. Carter, and R. Toral, Phys. Rev. Lett. **59**, 803 (1987); A. M. Ferrenberg and R. H. Swendsen, *ibid.* **63**, 1195 (1989); A. N. Alves, B. A. Berg, and R. Villanova, Phys. Rev. B **41**, 383 (1990).

[19]V. L. Privman, in *Finite Size Scaling and Numerical Simulation of Statistical Systems,* Ref. 7.

[20]Very recently (since this work was completed) this assumption has been placed on a rigorous basis in Refs. 9 and 10, although it was foreshadowed by Ref. 8.

[21]R. J. Baxter, J. Phys. C **6**, L445 (1973).

[22]T. Kihara, Y. Midzuno, and T. Shizume, J. Phys. Soc. Jpn. **9**, 681 (1954).

[23]R. H. Swendsen and J.-S. Wang, Phys. Rev. Lett. **58**, 86 (1987).