Extrapolating Monte Carlo Simulations to Infinite Volume: Finite-Size Scaling at $\xi/L \gg 1$

Sergio Caracciolo*
Dipartimento di Fisica, Università di Lecce and INFN–Sezione di Lecce, I-73100 Lecce, Italy

Robert G. Edwards†
Supercomputer Computations Research Institute, Florida State University, Tallahassee, Florida 32306

Sabino José Ferreira‡
Departamento de Física, Instituto de Ciências Exatas, Universidade Federal de Minas Gerais, Caixa Postal 702, Belo Horizonte, Minas Gerais 30161, Brasil

Andrea Pelissetto§ and Alan D. Sokal
Department of Physics, New York University, 4 Washington Place, New York, New York 10003
(Received 9 September 1994)

We present a simple and powerful method for extrapolating finite-volume Monte Carlo data to infinite volume, based on finite-size-scaling theory. We discuss carefully its systematic and statistical errors, and we illustrate it using three examples: the two-dimensional three-state Potts antiferromagnet on the square lattice, and the two-dimensional $O(3)$ and $O(\infty)$ σ models. In favorable cases it is possible to obtain reliable extrapolations (errors of a few percent) even when the correlation length is 1000 times larger than the lattice.

PACS numbers: 02.70.Lq, 05.70.Jk, 11.15.Ha, 64.60.Fr

No matter how powerful computers become, physicists will always want to study problems that are too difficult for the computers at hand. For example, in statistical mechanics and quantum field theory, physicists want to push to ever larger correlation lengths ξ. But Monte Carlo simulations must perforce be carried out on lattices of finite linear size L (limited by computer memory and speed); the data are then extrapolated to the infinite-volume limit $L = \infty$. Obviously this extrapolation—which is based on the theory of finite-size scaling (FSS) [1]—is feasible in practice only if ξ/L is not too large. But how large?

In this Letter we present a simple and powerful method for performing the extrapolation to $L = \infty$, and discuss carefully its systematic and statistical errors. We illustrate the method using three examples: the two-dimensional three-state Potts antiferromagnet on the square lattice [2], and the two-dimensional $O(3)$ and $O(\infty)$ σ models [3,4]. We found—much to our surprise—that in favorable cases it is possible to obtain reliable extrapolations (errors of a few percent) at ξ/L as large as 10–1000.

Consider, for a start, a model controlled by a renormalization-group (RG) fixed point having *one* relevant operator. Let us work on a periodic lattice of linear size L. Let $\xi(\beta, L)$ be a suitably defined finite-volume correlation length [5], and let $\mathcal{O}$ be any long-distance observable (e.g., the correlation length or the susceptibility). Then finite-size-scaling theory [1] predicts that [7]

$$\frac{\mathcal{O}(\beta, L)}{\mathcal{O}(\beta, \infty)} = f_{\mathcal{O}}(\xi(\beta, \infty)/L) + O(\xi^{-\omega}, L^{-\omega}), \quad (1)$$

where $f_{\mathcal{O}}$ is a universal function and ω is a correction-to-scaling exponent. It follows that if s is any fixed scale factor (usually we take $s = 2$) then

$$\frac{\mathcal{O}(\beta, sL)}{\mathcal{O}(\beta, L)} = F_{\mathcal{O}}(\xi(\beta, L)/L) + O(\xi^{-\omega}, L^{-\omega}), \quad (2)$$

where $F_{\mathcal{O}}$ can be expressed in terms of $f_{\mathcal{O}}$ and f_ξ.

Our method proceeds as follows [8]. Make Monte Carlo runs at numerous pairs (β, L) and (β, sL). Plot $\mathcal{O}(\beta, sL)/\mathcal{O}(\beta, L)$ versus $\xi(\beta, L)/L$, using those points satisfying both $\xi(\beta, L) \geq$ some value $\xi_{\min}$ and $L \geq$ some value $L_{\min}$. If all these points fall with good accuracy on a single curve—thus verifying the ansatz (2) for $\xi \geq \xi_{\min}$, $L \geq L_{\min}$—choose a smooth fitting function $F_{\mathcal{O}}$. Then, using the functions F_ξ and $F_{\mathcal{O}}$, extrapolate the pair $(\xi, \mathcal{O})$ successively from $L \to sL \to s^2L \to \cdots \to \infty$.

We have chosen to use functions $F_{\mathcal{O}}$ of the form

$$F_{\mathcal{O}}(x) = 1 + a_1 e^{-1/x} + a_2 e^{-2/x} + \cdots + a_n e^{-n/x}. \quad (3)$$

This form is partially motivated by theory, which tells us that $F(x) \to 1$ exponentially fast as $x \to 0$ [11]. Typically a fit of order $3 \leq n \leq 12$ is sufficient; we increase n until the χ^2 of the fit becomes essentially constant. The resulting χ^2 value provides a check on the systematic errors arising from corrections to scaling and/or from the inadequacies of the form (3).

The *statistical* error on the extrapolated value of $\mathcal{O}_\infty(\beta) \equiv \mathcal{O}(\beta, \infty)$ comes from three sources: (i) error on $\mathcal{O}(\beta, L)$, which gets multiplicatively propagated to $\mathcal{O}_\infty$; (ii) error on $\xi(\beta, L)$, which affects the argument $x \equiv \xi(\beta, L)/L$ of the scaling functions F_ξ and $F_{\mathcal{O}}$; and (iii) statistical error in our estimate of the coefficients $a_1, \ldots, a_n$ in F_ξ and $F_{\mathcal{O}}$. The errors of types (i) and (ii) depend on the statistics available at the single point (β, L), while the error of type (iii) depends on the

statistics in the whole set of runs. Errors (i) + (ii) [(i) + (ii) + (iii)] can be quantified by performing a Monte Carlo experiment in which the input data at (β, L) [the whole set of input data] are varied randomly within their error bars and then extrapolated [12–14].

The discrepancies between the extrapolated values from different lattice sizes at the same β—to the extent that these exceed the estimated statistical errors—indicate the presence of systematic errors and thus the necessity of increasing $L_{\min}$ and/or $\xi_{\min}$ and/or n.

A figure of (de)merit of the method is the relative variance on the extrapolated value $\mathcal{O}_\infty(\beta)$, multiplied by the computer time needed to obtain it [15]. We expect this *relative variance-time product* (RVTP) [for errors (i) + (ii) only] to scale as

$$\mathrm{RVTP}(\beta, L) \approx \xi_\infty(\beta)^{d+z_{\mathrm{int},\mathcal{O}}} G_{\mathcal{O}}(\xi_\infty(\beta)/L), \qquad (4)$$

where d is the spatial dimension and $z_{\mathrm{int},\mathcal{O}}$ is the dynamic critical exponent of the Monte Carlo algorithm being used; here $G_{\mathcal{O}}$ is a combination of several static and dynamic finite-size-scaling functions, and depends both on the observable $\mathcal{O}$ and on the algorithm but not on the scale factor s. As ξ_∞/L tends to zero, we expect $G_{\mathcal{O}}$ to diverge as $(\xi_\infty/L)^{-d}$ (it is wasteful to use a lattice $L \gg \xi_\infty$). As ξ_∞/L tends to infinity, we expect $G_{\mathcal{O}} \sim (\xi_\infty/L)^p$ [16], but the *power p can be either positive or negative.* If $p > 0$, there is an optimum value of ξ_∞/L; this determines the best lattice size at which to perform runs for a given β. If $p < 0$, it is most efficient to use the *smallest* lattice size for which the corrections to scaling are negligible compared to the statistical errors. (This neglects errors of type (iii) [17]; the optimization becomes much more complicated if they are included.)

Our first example [2] is the two-dimensional three-state Potts antiferromagnet on the square lattice, which is believed to have a critical point at $\beta = \infty$ [18]. We used the Wang-Swendsen-Kotecký cluster algorithm [19], which appears to have *no* critical slowing down ($\tau_{\mathrm{int},\mathcal{M}^2_{\mathrm{stagg}}} < 5$ uniformly in β and L) [2]. We ran on lattices $L = 32, 64, 128, 256, 512, 1024, 1536$ at 153 different pairs (β, L) in the range $5 \lesssim \xi_\infty \lesssim 20\,000$. Each run was between 2×10^5 and 2.2×10^7 iterations, and the total CPU time was modest by our standards (about 2 yr on an IBM RS-6000/370). We took $\xi_{\min} = 10$ and $L_{\min} = 128$ and used a quintic fit in (3); the result for F_ξ is shown in Fig. 1 [$\chi^2 = 75.41$, 66 DF (No. of degrees of freedom), level = 20%]. The extrapolated values from different lattice sizes at the same β agree within the estimated statistical errors ($\chi^2 = 43.03$, 75 DF, level > 99%); see Table I for an example. The result for G_ξ is shown in Fig. 2; the errors are roughly constant for $\xi_0/L \gtrsim 0.4$ but rise sharply for smaller ξ_∞/L. The theoretical exponent p computed [16] from the fitted F_ξ is equal to $p = -0.10 \pm 0.06$; the curve suggests that $p = 0$. In practice we were able to obtain ξ_∞ to an accuracy of about 1% (2%,3%,5%) at $\xi_\infty \approx 1000$ (2000,5000,10 000).

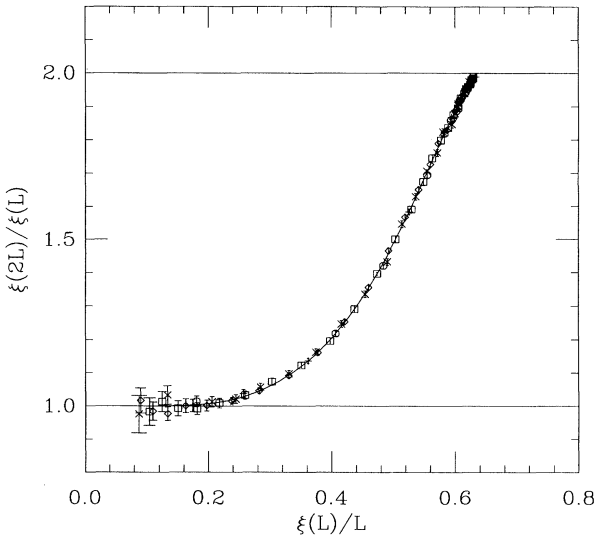

FIG. 1. $\xi(\beta, 2L)/\xi(\beta, L)$ vs $\xi(\beta, L)/L$ for the two-dimensional three-state Potts antiferromagnet. Symbols indicate $L = 32$ (+), 64 (×), 128 (□), 256 (◇), and 512 (○). Error bars are 1 standard deviation. Curve is a quintic fit in (3), with $\xi_{\min} = 10$ and $L_{\min} = 128$.

Next let us consider [3,4] the two-dimensional $O(3)$ σ model. We used the Wolff embedding algorithm with standard Swendsen-Wang updates [6,20,21]; again critical slowing down appears to be completely eliminated. We ran on lattices $L = 32, 48, 64, 96, 128, 192, 256, 384, 512$ at 180 different pairs (β, L) in the range $20 \lesssim \xi_\infty \lesssim 10^5$. Each run was between 10^5 and 5×10^6 iterations, and the total CPU time was 7 yr on an IBM RS-6000/370. We took $\xi_{\min} = 20$ and used a tenth-order fit. There appear to be weak corrections to scaling (of order ≲1.5%) in the region $0.3 \lesssim \xi_L/L \lesssim 0.7$ for lattices with $L \lesssim 64$–96. We therefore chose $L_{\min} = 128$ for $\xi_L/L \leq 0.7$, and $L_{\min} = 64$ for $\xi_L/L > 0.7$. The result for F_ξ is shown in Fig. 3 ($\chi^2 = 72.91$, 73 DF, level = 48%). The result for

TABLE I. Raw and extrapolated correlation lengths for the two-dimensional three-state Potts antiferromagnet at $\beta = 3.5$. Extrapolation based on $\xi_{\min} = 10$ and $L_{\min} = 128$ and a quintic fit. For each extrapolated value we reported the standard deviation of the estimate, including errors of all three types. The mean value and the χ^2 have been computed, taking into account the full covariance matrix [14].

L	Iterations	Raw Data $\xi(\beta, L)$	Extrapolated $\xi(\beta, \infty)$
32	2.2×10^7	19.02 (0.01)	90.68 (1.03)
64	10^6	35.52 (0.04)	92.47 (0.81)
128	10^6	60.61 (0.09)	93.01 (0.44)
256	10^6	84.69 (0.15)	93.01 (0.31)
512	5×10^5	92.48 (0.33)	92.85 (0.39)
1024	2×10^5	93.78 (1.17)	93.78 (1.17)
mean			93.01 (0.29)
($L \geq 128$)		$\chi^2 = 0.68$ (3 DF, level = 88%)	

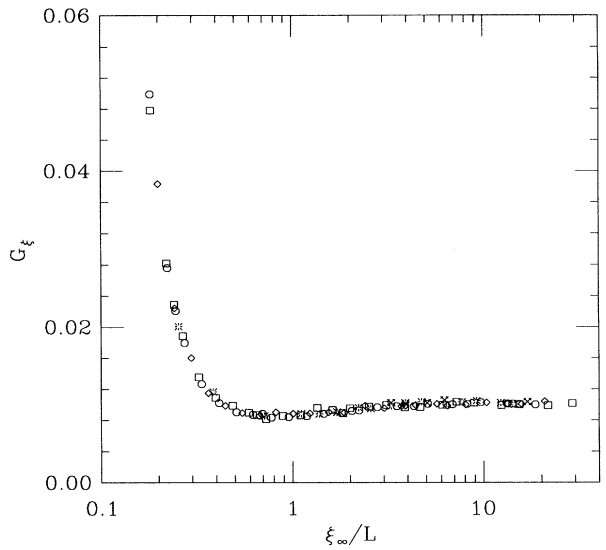

FIG. 2. Relative variance-time product [for errors (i) + (ii) only] divided by $\xi_\infty(\beta)^2$, plotted vs $\xi_\infty(\beta)/L$, for the two-dimensional three-state Potts antiferromagnet. Symbols indicate $L = 128$ (□), 256 (◇), 512 (○), 1024 (*), and 1536 (+).

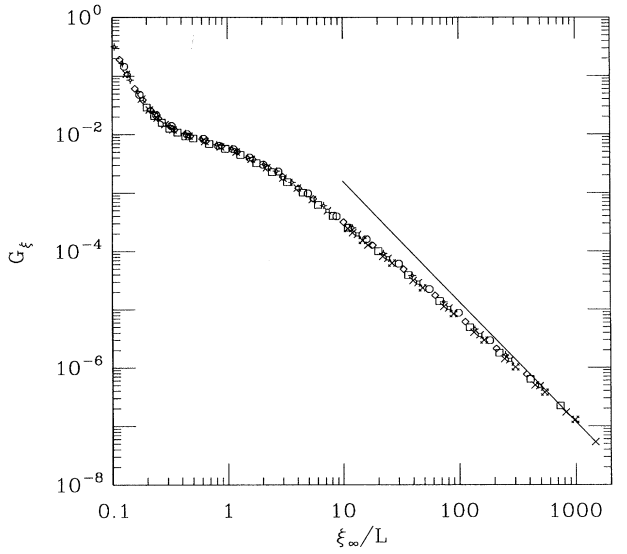

FIG. 4. Relative variance-time product [for errors (i) + (ii) only] divided by $\xi_\infty(\beta)^2$, plotted vs $\xi_\infty(\beta)/L$, for the two-dimensional $O(3)$ σ model. Symbols indicate $L = 64$ (×), 96 (+), 128 (□), 192 (⌑), 256 (◇), 384 (✧), and 512 (○). For comparison, the line shows the theoretical limiting slope –2.

G_ξ is shown in Fig. 4; at large ξ_∞/L it *decreases* sharply, with a power $p \approx -2$, in agreement with theory [16]. In practice we obtained ξ_∞ to an accuracy of about 0.2% (0.7%,1.1%,1.6%) at $\xi_\infty \approx 10^2$ (10^3,10^4,10^5).

We also carried out a "simulated Monte Carlo" experiment for the $O(N)$ σ model at $N = \infty$, by generating data

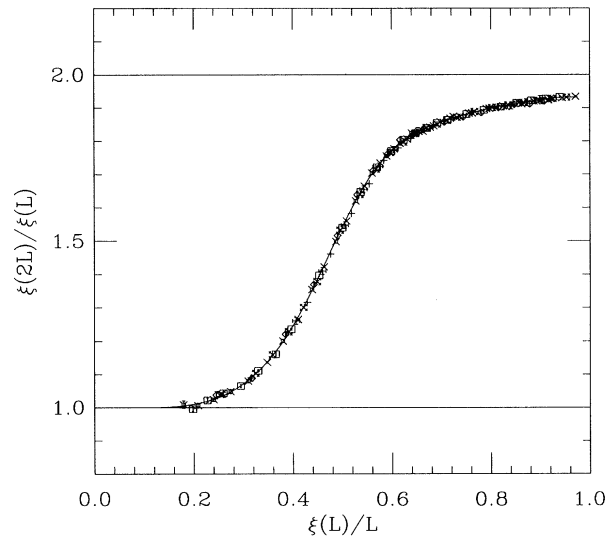

FIG. 3. $\xi(\beta,2L)/\xi(\beta,L)$ vs $\xi(\beta,L)/L$ for the two-dimensional $O(3)$ σ model. Symbols indicate $L = 32$ (+), 48 (⌑), 64 (×), 96 (⊠), 128 (□), 192 (⌑), and 256 (◇). Error bars are 1 standard deviation. Curve is a tenth-order fit in (3), with $\xi_{\min} = 20$ and $L_{\min} = 128$ (64) for $\xi(L)/L \leq 0.7$ (> 0.7).

from the exact finite-volume solution plus random noise of 0.1% for $L = 64,96,128$, 0.2% for $L = 192,256$, and 0.5% for $L = 384,512$ [which is the order of magnitude we attain in practice for $O(3)$]. We considered 35 values of β in the range $20 \lesssim \xi_\infty \lesssim 10^6$. We used $\xi_{\min} = 20$ and $L_{\min} = 64$ (in fact much smaller values could have been used, as corrections to scaling are very small here) and a tenth-order fit; for two different data sets we obtain $\chi^2 = 159.18$ and 179.39 with 165 DF. In practice we obtain ξ_∞ with an accuracy of 0.6% (1.2%,2%,3%) at $\xi_\infty \approx 10^3$ (10^4,10^5,10^6). Here we can also compare the extrapolated values $\xi_\infty^{\rm extr}(\beta)$ with the exact values $\xi_\infty^{\rm exact}(\beta)$. Defining $\mathcal{R} = \sum_\beta [\xi_\infty^{\rm extr}(\beta) - \xi_\infty^{\rm exact}(\beta)]^2/\sigma^2(\beta)$, we find for the two data sets $\mathcal{R} = 28.05$ (32.53) with 35 DF. Only 13 (10) points differ from the exact value by more than 1 standard deviation, and only 1 (1) by more than 2. Details on all of these models will be reported separately [2,4].

The method is easily generalized to a model controlled by a RG fixed point having k relevant operators. It suffices to choose $k - 1$ dimensionless ratios of long-distance observables, call them $R = (R_1, \ldots, R_{k-1})$, then the function $F_{\mathcal{O}}$ will depend parametrically on $R(\beta, L)$. In practice one can divide R space into "slices" within which $F_{\mathcal{O}}$ is empirically constant within error bars, and perform the fit (3) within each slice. We used this approach to study the mixed isovector-isotensor σ model, taking R to be the ratio of isovector to isotensor correlation length [3,4].

The method can also be applied to extrapolate the exponential correlation length (inverse mass gap). For

this purpose one must work in a system of size $L^{d-1} \times T$ with $T \gg \xi_{\rm exp}(\beta, L)$ (compare [9]).

We wish to thank Martin Hasenbusch and especially Jae-Kwon Kim for sharing their data with us, and for challenging us to push ever larger values of ξ/L. The authors' research was supported by CNR, INFN, CNPq, FAPEMIG, and DOE Contracts No. DE-FG05-85ER250000 and No. DE-FG05-92ER40742, NSF Grant No. DMS-9200719, and NATO CRG 910251.

*Electronic address: CARACCIO@LE.INFN.IT
†Electronic address: EDWARDS@SCRI.FSU.EDU
‡Electronic address: SABINO@FISICA.UFMG.BR
§Permanent address: Dipartimento di Fisica and INFN–Sezione di Pisa, Università degli Studi di Pisa, I-56100 Pisa, Italy. Electronic address: PELISSET@SUNTHPI1.DIFI.UNIPI.IT SOKAL@ACF4.NYU.EDU

[1] *Finite Size Scaling and Numerical Simulation of Statistical Systems* edited by V. Privman (World Scientific, Singapore, 1990).
[2] S. J. Ferreira and A. D. Sokal, Phys. Rev. B (to be published); (to be published).
[3] S. Caracciolo, R. G. Edwards, A. Pelissetto, and A. D. Sokal, Phys. Rev. Lett. **71**, 3906 (1993).
[4] S. Caracciolo, R. G. Edwards, A. Pelissetto, and A. D. Sokal, Reports No. hep-lat/9411009 and No. 9411064; (to be published).
[5] It is necessary to choose a definition of ξ that makes sense in a fully finite system. We use the second-moment correlation length defined by Eqs. (4.11)–(4.13) of Ref. [6].
[6] S. Caracciolo, R. G. Edwards, A. Pelissetto, and A. D. Sokal, Nucl. Phys. **B403**, 475 (1993).
[7] This form of finite-size scaling assumes hyperscaling, and thus is expected to hold only below the upper critical dimension of the model. See, e.g., Ref. [1], Chap. I, Sec. 2.7.
[8] Our method has many features in common with those of Lüscher, Weisz, and Wolff [9] and Kim [10]. In particular, all these methods share the property of working only with observable quantities (ξ, $\mathcal{O}$, and L) and not with bare quantities (β). Therefore, they rely on "scaling" and not on "asymptotic scaling;" they differ from other FSS-based methods such as phenomenological renormalization.
[9] M. Lüscher, P. Weisz, and U. Wolff, Nucl. Phys. **B359**, 221 (1991).
[10] J.-K. Kim, Phys. Rev. Lett. **70**, 1735 (1993); Nucl. Phys. B (Proc. Suppl.) **34**, 702 (1994); Phys. Rev. D **50**, 4663 (1994); Europhys. Lett. **28**, 211 (1994).
[11] H. Neuberger, Phys. Lett. B **233**, 183 (1989); S. Caracciolo and A. Pelissetto (to be published).
[12] In principle, ξ and $\mathcal{O}$ should be generated from a *joint* Gaussian with the correct covariance. We ignored this subtlety and simply generated *independent* fluctuations on ξ and $\mathcal{O}$.
[13] Errors of types (i) and (ii) can be exactly computed in terms of f_ξ and $f_\mathcal{O}$. For ξ we have $\Delta\xi_\infty/\xi_\infty = K_\xi(\xi_\infty/L)\Delta\xi_L/\xi_L$, where $\Delta\xi_L$ ($\Delta\xi_\infty$) is the standard deviation on the raw (extrapolated) value, and $K_\xi(z) = f_\xi(z)/[f_\xi(z) + z f'_\xi(z)]$. For a deviation, see Ref. [4].
[14] The extrapolated estimates from different (β, L) are correlated. In setting the error bars we have kept account of the full covariance matrix between extrapolations at the same β but different L, but we have ignored correlations between extrapolations at different values of β.
[15] This variance-time product tends to a constant as the CPU time tends to infinity. However, if the CPU time used is too small, then the variance-time product can be significantly larger than its asymptotic value, due to nonlinear cross terms between error sources (i) and (ii).
[16] The exponent p, which does not depend on s, can be computed in terms of F_ξ. Define an exponent h by assuming that $\Delta\xi_L/\xi_L \approx \xi_\infty^z H(\xi_\infty/L)/N_{\rm iter} \sim \xi_\infty^{z+h}/N_{\rm iter}$, where $\Delta\xi_L$ is the error on ξ_L for a run of length $N_{\rm iter}$. If $\xi_L/L \to x_* < \infty$, when $\xi_\infty/L \to \infty$ (this is the case for models with critical exponent $\eta > 0$) and $R \equiv 1 + x_* F'_\xi(x_*)/s > 1$, then $p = -d + 2\log R/\log s + 2h$. If $R = 1$, logarithmic terms appear and $G_\xi(z) \sim z^{-d+2h}(\log z)^q$, with $q > 0$. If instead $\xi_L/L \to \infty$, when $\xi_\infty/L \to \infty$ (this is the case for asymptotically free theories), then $G_\xi(z) \sim z^{-d+2h}(\log z)^2$.
[17] In our three examples, the total error (i) + (ii) + (iii) is never more than twice the error (i) + (ii), and is usually much less.
[18] R. J. Baxter, Proc. R. Soc. London A **383**, 43 (1982).
[19] J.-S. Wang, R. H. Swendsen, and R. Kotecký, Phys. Rev. Lett. **63**, 109 (1989); Phys. Rev. B **42**, 2465 (1990); M. Lubin and A. D. Sokal, Phys. Rev. Lett. **71**, 1778 (1993).
[20] U. Wolff, Phys. Rev. Lett. **62**, 361 (1989).
[21] R. H. Swendsen and J.-S. Wang, Phys. Rev. Lett. **58**, 86 (1987).