Simulation of Statistical-Mechanics Models

Alan D. Sokal

Department of Physics
New York University
4 Washington Place
New York, NY 10003
USA

E-mail: SOKAL@NYU.EDU

Lectures at the Les Houches Summer School on
"Mathematical Statistical Mechanics"
July 2005

Abstract

My course at the 2005 Les Houches summer school was based on the course "Monte Carlo Methods in Statistical Mechanics: Foundations and New Algorithms" that I gave at the 1996 Cargèse summer school and which was published in that school's proceedings [1]. The basic ideas have not changed radically in the past ten years, though many new applications have been made and some new algorithms have been invented. In these notes, I have added some updated information concerning the Swendsen–Wang algorithm.

1 Overview

The goal of these lectures is to give an introduction to the use of Monte Carlo methods in statistical mechanics and quantum field theory, with an emphasis on:

1) the conceptual foundations of the method, including the possible dangers and misuses, and the correct use of statistical error analysis; and

2) new (collective-mode) Monte Carlo algorithms for problems in critical phenomena and quantum field theory, aimed at reducing or eliminating the "critical slowing-down" found in conventional (local) algorithms.

These lectures are aimed at a mixed audience of theoretical, computational and mathematical physicists — some of whom are currently doing or want to do Monte Carlo studies themselves, others of whom want to be able to evaluate the reliability of published Monte Carlo work, and still others of whom might want to contribute to the mathematical analysis of Markov-chain Monte Carlo algorithms (currently a very lively branch of probability theory).

The principal topics to be covered are the following:

1. Introduction: Dangers of Monte Carlo methods
2. Dynamic (Markov-chain) Monte Carlo methods: General theory
3. Statistical analysis of dynamic Monte Carlo data
4. Conventional (local) Monte Carlo algorithms for spin models
5. Cluster algorithms (Swendsen–Wang and related algorithms)
6. Rigorous analysis of Markov-chain Monte Carlo methods

Details on all these topics can be found in [1]. Here I would simply like to give some updated information concerning the Swendsen–Wang algorithm.

2 The Swendsen–Wang algorithm: Some recent progress

The Swendsen–Wang [2] algorithm for simulating the ferromagnetic q-state Potts model [3–5] is the prototype for many other collective-mode Monte Carlo algorithms in statistical mechanics that are generically termed *cluster algorithms* or *auxiliary-variable algorithms*. The basic idea behind all algorithms of Swendsen-Wang type is to augment the given model by means of auxiliary variables, and then to simulate this augmented model.

Here I shall begin by reviewing the theory underlying the Swendsen–Wang algorithm. Next I shall present the most recent numerical data on its dynamic critical behavior, and discuss some of the things that remain to be understood.

2.1 Fortuin–Kasteleyn representation and Swendsen–Wang algorithm

The q-state Potts model [3–5] is a generalization of the Ising model in which each spin σ_i can take q distinct values rather than just two; here q is an integer ≥ 2. At each site i we therefore have a "spin" σ_i taking values in the set $\{1, 2, \ldots, q\}$. Neighboring spins prefer to be in the same state (here I am considering the *ferromagnetic* case), and pay an energy price if they are not. The Hamiltonian is therefore

$$H(\sigma) \;=\; \sum_{\langle ij \rangle} J_{ij} \, (1 - \delta_{\sigma_i,\sigma_j}) \tag{2.1}$$

with $J_{ij} \geq 0$ for all i, j, and the partition function is

$$\begin{aligned} Z &= \sum_{\{\sigma\}} \exp[-H(\sigma)] \\ &= \sum_{\{\sigma\}} \exp\left[\sum_{\langle ij \rangle} J_{ij} \, (\delta_{\sigma_i,\sigma_j} - 1)\right] \\ &= \sum_{\{\sigma\}} \prod_{\langle ij \rangle} [(1 - p_{ij}) \;+\; p_{ij} \delta_{\sigma_i,\sigma_j}] \end{aligned} \tag{2.2}$$

where we have defined $p_{ij} = 1 - \exp(-J_{ij})$. The Gibbs measure $\mu_{\mathrm{Potts}}(\sigma)$ is, of course,

$$\begin{aligned} \mu_{\mathrm{Potts}}(\sigma) &= Z^{-1} \exp\left[\sum_{\langle ij \rangle} J_{ij} \, (\delta_{\sigma_i,\sigma_j} - 1)\right] \\ &= Z^{-1} \prod_{\langle ij \rangle} [(1 - p_{ij}) \;+\; p_{ij} \delta_{\sigma_i,\sigma_j}] \, . \end{aligned} \tag{2.3}$$

We now use the deep identity

$$a + b \;=\; \sum_{n=0}^{1} [a \delta_{n,0} \;+\; b \delta_{n,1}] \tag{2.4}$$

on each bond $\langle ij \rangle$; that is, we introduce on each bond $\langle ij \rangle$ an auxiliary variable n_{ij} taking the values 0 and 1, and obtain

$$Z \;=\; \sum_{\{\sigma\}} \sum_{\{n\}} \prod_{\langle ij \rangle} [(1 - p_{ij}) \, \delta_{n_{ij},0} \;+\; p_{ij} \delta_{n_{ij},1} \delta_{\sigma_i,\sigma_j}] \, . \tag{2.5}$$

Let us now take seriously the $\{n\}$ as dynamical variables: we can think of n_{ij} as an *occupation variable* for the bond $\langle ij \rangle$ (1 = occupied, 0 = empty). We therefore define the *Fortuin–Kasteleyn–Swendsen–Wang (FKSW) model* to be a joint model having q-state Potts spins σ_i at the sites and occupation variables n_{ij} on the bonds, with joint probability distribution

$$\mu_{\mathrm{FKSW}}(\sigma, n) \;=\; Z^{-1} \prod_{\langle ij \rangle} \left[(1 - p_{ij})\, \delta_{n_{ij},0} \;+\; p_{ij} \delta_{n_{ij},1} \delta_{\sigma_i,\sigma_j} \right] . \tag{2.6}$$

Finally, let us see what happens if we sum over the $\{\sigma\}$ at fixed $\{n\}$. Each occupied bond $\langle ij \rangle$ imposes a constraint that the spins σ_i and σ_j must be in the same state, but otherwise the spins are unconstrained. We therefore group the sites into connected clusters (two sites are in the same cluster if they can be joined by a path of occupied bonds); then all the spins within a cluster must be in the same state (all q values are equally probable), and distinct clusters are independent. It follows that

$$Z \;=\; \sum_{\{n\}} \left(\prod_{\langle ij \rangle:\, n_{ij}=1} p_{ij} \right) \left(\prod_{\langle ij \rangle:\, n_{ij}=0} (1 - p_{ij}) \right) q^{\mathcal{C}(n)} , \tag{2.7}$$

where $\mathcal{C}(n)$ is the number of connected components (including one-site components) in the graph whose edges are the bonds having $n_{ij} = 1$. The corresponding probability distribution,

$$\mu_{\mathrm{RC}}(n) \;=\; Z^{-1} \left(\prod_{\langle ij \rangle:\, n_{ij}=1} p_{ij} \right) \left(\prod_{\langle ij \rangle:\, n_{ij}=0} (1 - p_{ij}) \right) q^{\mathcal{C}(n)} , \tag{2.8}$$

is called the *random-cluster model with parameter q*. This is a generalized bond-percolation model, with non-local correlations coming from the factor $q^{\mathcal{C}(n)}$; for $q = 1$ it reduces to ordinary bond percolation. Note, by the way, that in the random-cluster model (unlike the Potts and FKSW models), q is merely a parameter; it can take any positive real value, not just $2, 3, \ldots$. So the random-cluster model defines, in some sense, an analytic continuation of the Potts model to non-integer q; ordinary bond percolation corresponds to the "one-state Potts model".

We have already verified the following facts about the FKSW model:

a) $Z_{\mathrm{Potts}} = Z_{\mathrm{FKSW}} = Z_{\mathrm{RC}}$.

b) The marginal distribution of μ_{FKSW} on the Potts variables $\{\sigma\}$ (integrating out the $\{n\}$) is precisely the Potts model $\mu_{\mathrm{Potts}}(\sigma)$.

c) The marginal distribution of μ_{FKSW} on the bond occupation variables $\{n\}$ (integrating out the $\{\sigma\}$) is precisely the random-cluster model $\mu_{\mathrm{RC}}(n)$.

The conditional distributions of $\mu_{\rm FKSW}$ are also simple:

d) The conditional distribution of the $\{n\}$ given the $\{\sigma\}$ is as follows: independently for each bond $\langle ij \rangle$, one sets $n_{ij} = 0$ in case $\sigma_i \neq \sigma_j$, and sets $n_{ij} = 0, 1$ with probability $1 - p_{ij}, p_{ij}$, respectively, in case $\sigma_i = \sigma_j$.

e) 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 the cluster to the same value, chosen equiprobably from $\{1, 2, \ldots, q\}$.

These facts can be used for both analytic and numerical purposes. For example, by using facts (b), (c) and (e) we can prove an identity relating expectations in the Potts model to connection probabilities in the random-cluster model:

$$\begin{aligned}
\langle \delta_{\sigma_i,\sigma_j} \rangle_{Potts,q} &= \langle \delta_{\sigma_i,\sigma_j} \rangle_{FKSW,q} && \text{[by (b)]} \\
&= \langle E(\delta_{\sigma_i,\sigma_j} \mid \{n\}) \rangle_{FKSW,q} \\
&= \left\langle \frac{(q-1)\gamma_{ij} + 1}{q} \right\rangle_{FKSW,q} && \text{[by (e)]} \\
&= \left\langle \frac{(q-1)\gamma_{ij} + 1}{q} \right\rangle_{RC,q} && \text{[by (c)]} \qquad (2.9)
\end{aligned}$$

Here

$$\gamma_{ij} \equiv \gamma_{ij}(n) \equiv \begin{cases} 1 & \text{if } i \text{ is connected to } j \\ 0 & \text{if } i \text{ is not connected to } j \end{cases} \qquad (2.10)$$

and $E(\cdot \mid \{n\})$ denotes conditional expectation given $\{n\}$.[1] For the Ising model with the usual convention $\sigma = \pm 1$, (2.9) can be written more simply as

$$\langle \sigma_i \sigma_j \rangle_{Ising} = \langle \gamma_{ij} \rangle_{RC,q=2} \,. \qquad (2.11)$$

Similar identities can be proven for higher-order correlation functions, and can be employed to prove Griffiths-type correlation inequalities for the Potts model [7, 8].

On the other hand, Swendsen and Wang [2] exploited facts (b)–(e) to devise a radically new type of Monte Carlo algorithm. The Swendsen-Wang algorithm (SW) simulates the joint model (2.6) by alternately applying the conditional distributions (d) and (e) — that is, by alternately generating new bond occupation variables (independent of the old ones) given the spins, and new spin variables (independent of the old ones) given the bonds. Each of these operations can be carried out in a computer time of order volume: for generating the bond variables this is trivial, and for

[1] For an excellent introduction to conditional expectations, see [6].

generating the spin variable it relies on efficient (linear-time) algorithms for computing the connected clusters.[2] It is trivial that the SW algorithm leaves invariant the Gibbs measure (2.6), since any product of conditional probability operators has this property. It is also easy to see that the algorithm is ergodic, in the sense that every configuration $\{\sigma, n\}$ having nonzero $\mu_{\rm FKSW}$-measure is accessible from every other. So the SW algorithm is at least a *correct* algorithm for simulating the FKSW model. It is also an algorithm for simulating the Potts and random-cluster models, since expectations in these two models are equal to the corresponding expectations in the FKSW model.

Historical remark. The random-cluster model was introduced in 1969 by Fortuin and Kasteleyn [12]; they derived the identity $Z_{\rm Potts} = Z_{\rm RC}$, along with the correlation-function identity (2.9) and some generalizations. These relations were rediscovered several times during the subsequent two decades [13]. Surprisingly, however, no one seems to have noticed the *joint* probability distribution $\mu_{\rm FKSW}$ that underlay all these identities; this was discovered implicitly by Swendsen and Wang [2], and was made explicit by Edwards and Sokal [14].

2.2 Dynamic critical behavior: Numerical results

It is certainly plausible that the SW algorithm might have less critical slowing-down than the conventional (single-spin-update) algorithms: the reason is that a local move in one set of variables can have highly nonlocal effects in the other. In particular, setting $n_{ij} = 0$ on a single bond may disconnect a cluster, dividing it into two big pieces; and then, in the next half-step of the algorithm, the spins in these two pieces may be flipped simultaneously but independently. In some sense, therefore, the SW algorithm is a collective-mode algorithm in which the collective modes are *chosen by the system* rather than imposed from the outside as in multi-grid [15–21] or Fourier acceleration [22]. The miracle is that this is done in a way that preserves the correct Gibbs measure.

How well does the SW algorithm perform? In at least some cases, the performance is nothing short of extraordinary. Table 1 shows the latest data [23] on the two-dimensional Ising model at the bulk critical temperature, together with comparison data for a single-site Metropolis algorithm [24]. These data are consistent with a behavior $\tau_{\rm SW} \sim L^{\approx 0.22}$; but they are also consistent with $\tau_{\rm SW} \sim \log^2 L$ [23]. (Alas, it is *extremely* difficult to distinguish numerically a small power from a logarithm.)

[2] Determining the connected components of an undirected graph is a classic problem of computer science. The depth-first-search and breadth-first-search algorithms [9] have a running time of order V, while the Fischer-Galler-Hoshen-Kopelman algorithm (in one of its variants) [10] has a worst-case running time of order $V \log V$, and an observed mean running time of order V in percolation-type problems [11].

L	χ	Swendsen–Wang $\tau_{\mathrm{int},\mathcal{E}}$	Metropolis $\tau_{\mathrm{exp},\mathcal{M}}$
4	12.183 ± 0.007	2.027 ± 0.010	
8	41.396 ± 0.008	2.590 ± 0.004	
16	139.584 ± 0.039	3.253 ± 0.008	285.6 ± 4.3
32	470.022 ± 0.140	4.012 ± 0.011	1258 ± 28
64	1581.319 ± 0.378	4.892 ± 0.011	5380 ± 140
128	5320.644 ± 1.680	5.875 ± 0.018	23950 ± 480
256	17899.581 ± 5.846	6.928 ± 0.022	
512	60184.698 ± 18.670	8.107 ± 0.025	

Table 1: Static and dynamic data for the two-dimensional Ising model at criticality on an $L \times L$ periodic lattice. Susceptibility χ and integrated autocorrelation time $\tau_{\mathrm{int},\mathcal{E}}$ ($\mathcal{E}$ = energy $\approx$ slowest mode) using the Swendsen–Wang algorithm [23, Table 6]; and exponential autocorrelation time $\tau_{\mathrm{exp},\mathcal{M}}$ ($\mathcal{M}$ = magnetization $\approx$ slowest mode), using the Metropolis algorithm [24]. Error bars are one standard deviation.

In either case, we can say that in practical terms, the critical slowing-down is almost completely eliminated.

By contrast, the conventional local algorithms for the two-dimensional Ising model have $\tau_{\mathrm{local}} \sim L^{\approx 2.13}$ [24], as can be seen in the last column of Table 1. So the advantage of Swendsen–Wang over conventional algorithms grows asymptotically (for this model) like $L^{\approx 2}$ — an enormous factor. To be sure, one iteration of the Swendsen–Wang algorithm may be a factor of ~ 10 more costly in CPU time than one iteration of a conventional algorithm (the exact factor depends on the efficiency of the cluster-finding subroutine). But the SW algorithm wins already for modest values of L.

For the other Potts models, the performance of the SW algorithm is less spectacular than for the two-dimensional Ising model, but it is still very impressive. Table 2 shows the current best estimates of the dynamic critical exponent z_{SW} for q-state Potts ferromagnets in d dimensions, as a function of q and d. All these exponents are much lower than the $z \gtrsim 2$ observed in the local algorithms.

2.3 How much do we understand about the SW dynamics?

Although the SW algorithm performs impressively well, we understand very little about *why* the dynamic critical exponents take the values they do. A few cases are easy:

1) If $q = 1$, then all spins are in the same state (the *only* state!), and all bonds are

	Estimates of $z_{\rm SW}$			
	$q = 1$	$q = 2$	$q = 3$	$q = 4$
$d = 1$	0	0	0	0
$d = 2$	0	0.222 ± 0.007	0.514 ± 0.006	1 ($\times \log^{??}$)
$d = 3$	0	0.46 ± 0.03	—	—
$d = 4$	0	1 ($\times \log^{??}$)	—	—

Table 2: Current best estimates of the dynamic critical exponent z for the Swendsen–Wang algorithm. Estimates are taken from [23] for $d = 2$, $q = 2$; [25] for $d = 2$, $q = 3$; [26, 27] for $d = 2$, $q = 4$; [28] for $d = 3$, $q = 2$; and [29–32] for $d = 4$, $q = 2$. Error bars are one standard deviation, and include statistical errors only.

thrown independently, so the autocorrelation time is zero. Here the SW algorithm just reduces to the standard *static* algorithm for independent bond percolation.

2) If $d = 1$ (more generally, if the lattice is a *tree*), the SW dynamics is exactly soluble: the behavior of each bond is independent of each other bond, and $\tau_{exp} \to -1/\log(1 - 1/q) < \infty$ as $\beta \to +\infty$.

But the remainder of our understanding of the SW algorithm is very murky, and we are only beginning to understand its dynamic critical behavior.

One source of insight is an exact (non-rigorous but presumably rigorizable) calculation yielding $z_{\rm SW} = 1$ for the Ising model on the complete graph (also called the Curie–Weiss or mean-field model) [30, 32]. This result suggests (but of course does not prove) that $z_{\rm SW} = 1$ for Ising models ($q = 2$) in dimensions $d \geq 4$.

A second source of insight is a rigorous lower bound on the Swendsen–Wang autocorrelation time, proved more than 15 years ago by Li and Sokal [33] (see also [25]):

$$\tau_{\rm exp} \gtrsim \tau_{{\rm int},\mathcal{E}} \geq \tau_{{\rm int},\mathcal{N}} \geq \text{const} \times C_H \tag{2.12}$$

where $\mathcal{E}$ is the energy, $\mathcal{N}$ is the bond density, and C_H is the specific heat. This immediately implies the lower bound

$$z_{\rm SW} \geq \alpha/\nu \tag{2.13}$$

for the SW dynamic critical exponent in terms of static exponents. The bound (2.13) is fairly close to sharp — indeed, it may even be sharp modulo a logarithm — for the $q = 2, 3, 4$ Potts models in two dimensions.[3] But it is extremely far from sharp for the

[3]See [23, 25–27] for numerical studies of the possible sharpness of the Li–Sokal bound for two-dimensional Potts models with $q = 2, 3, 4$. These data are consistent with sharpness modulo a

Ising models in dimensions 3 and 4, where $z_{\rm SW} \approx 0.46$ [28] and $z_{\rm SW} = 1$, respectively, compared to $\alpha/\nu \approx 0.1756$ [34] and $\alpha/\nu = 0$ ($\times \log^{1/3}$).

Clearly we are failing to grasp something extremely fundamental about the SW dynamics. Physically, the bound (2.12) expresses the slow equilibration of the bond density $\mathcal{N}$, and more generally of "energy-like" observables — a fact that can be captured mathematically in a Rayleigh–Ritz proof using $\mathcal{N}$ as a trial vector [1,25,33]. The large discrepancy $z_{\rm SW} - \alpha/\nu$ for the 3- and 4-dimensional Ising models suggests that in these cases the SW dynamics is dominated by a different (and much slower) mechanism, possibly one having to do with "susceptibility-like" observables. Some fairly wild speculations in this direction can be found in [28, Section 6]. But no one has yet been able to find a trial vector that allows a useful bound other than (2.12) to be proven. Nor has anyone found a plausible heuristic argument that would allow $z_{\rm SW}$ to be predicted, even roughly, in terms of static exponents.

One important recent contribution is the invention of a new SW-like algorithm by Chayes and Machta [35]. Whereas the SW algorithm starts from the Potts model (2.3) and introduces auxiliary variables $n_{ij} \in \{0,1\}$ living on the bonds, the Chayes–Machta algorithm starts from the random-cluster model (2.8) and introduces auxiliary variables $\sigma_i \in \{1,2,\ldots,k\}$ living on the sites; here k is an arbitrary positive integer $\le q$. The Chayes–Machta algorithm thus applies to all random-cluster models with parameter $q \ge 1$, integer or noninteger. If q is an integer and we choose $k = q$, then the auxiliary variables σ_i are precisely the Potts spins, and the Chayes–Machta algorithm is identical to the Swendsen–Wang algorithm. If q is an integer and we choose $k < q$, then the Chayes–Machta algorithm corresponds to a minor variant of the Swendsen–Wang algorithm (which almost certainly has the same dynamic critical exponent as the standard SW algorithm). But the key fact about the Chayes–Machta algorithm is that it handles noninteger q on an equal footing with integer q. Indeed, the Chayes–Machta algorithm deserves to be thought of as the "natural" interpolation of the Swendsen–Wang algorithm to noninteger q (although unfortunately only for $q \ge 1$). This is important, both conceptually and numerically, because it is difficult to understand the SW dynamics based on only five nontrivial data points ($d = 2$, $q = 2,3,4$ and $d = 3,4$, $q = 2$). The Chayes–Machta algorithm allows us to study the whole range $1 \le q \le 4$ for two-dimensional lattices, and the range $1 \le q \le q_c(\mathcal{L})$ for three- and four-dimensional lattices [here $q_c(\mathcal{L})$ is the value of q at which the phase transition in the random-cluster model on lattice $\mathcal{L}$ changes from second-order to first-order; note that it can depend on the lattice and not only on the dimension]. Further information about the Chayes–Machta algorithm, both theoretical and numerical, will hopefully be available soon [36,37].

The Potts model with q *large* behaves very differently. Instead of a critical point,

---

logarithm, i.e. $\tau_{\rm SW}/C_H \sim \log L$. But they are also consistent with a small difference in the exponents, i.e. $z_{\rm SW} - \alpha/\nu$ of order 0.1. Once again, it is extremely difficult to distinguish numerically a small power from a logarithm.

the model undergoes a *first-order* phase transition: on the standard two-dimensional lattices, this occurs when $q > 4$, while in three or more dimensions, it is believed to occur already when $q \geq 3$ [4]. At a first-order transition, *both* the conventional algorithms and the Swendsen–Wang algorithm have an extremely severe slowing-down (*much* more severe than the slowing-down at a critical point): right at the transition temperature, we expect $\tau_{\mathrm{SW}} \sim \exp(cL^{d-1})$. This is because sets of configurations typical of the ordered and disordered phases are separated by free-energy barriers of order L^{d-1}, i.e. by sets of intermediate configurations that contain interfaces of surface area $\sim L^{d-1}$ and therefore have an equilibrium probability $\sim \exp(-cL^{d-1})$. Indeed, a slightly weaker lower bound has been proven rigorously by Borgs *et al.* [38], namely $\tau_{\mathrm{SW}} \gtrsim \exp[cL/(\log L)^2]$.[4]

Acknowledgments

Many of the ideas reported here have grown out of joint work with my colleagues Robert Edwards, Xiao-Jian Li, Neal Madras, Giovanni Ossola and Jesús Salas. I thank them for many pleasant and fruitful collaborations. In addition, I would like to thank Mal Kalos for introducing me to Monte Carlo work many years ago, and for patiently answering my first naive questions.

The author's research was supported in part by the U.S. National Science Foundation grant PHY–0424082.

References

[1] A.D. Sokal, Monte Carlo methods in statistical mechanics: Foundations and new algorithms, in C. DeWitt-Morette, P. Cartier and A. Folacci, eds., *Functional Integration: Basics and Applications* (1996 Cargèse summer school), pp. 131–192 (Plenum, New York, 1997).

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

[3] R.B. Potts, Proc. Camb. Phil. Soc. **48**, 106 (1952).

[4] F.Y. Wu, Rev. Mod. Phys. **54**, 235 (1982); **55**, 315 (E) (1983).

[4] Earlier, Gore and Jerrum [39] proved that on the complete graph K_n, we have $\tau_{\mathrm{SW}} \gtrsim \exp(cn)$ at the first-order transition temperature ($q \geq 3$).

[5] F.Y. Wu, J. Appl. Phys. **55**, 2421 (1984).

[6] K. Krickeberg, *Probability Theory* (Addison-Wesley, Reading, Mass., 1965), Sections 4.2 and 4.5.

[7] R.H. Schonmann, J. Stat. Phys. **52**, 61 (1988).

[8] A.D. Sokal, unpublished.

[9] E.M. Reingold, J. Nievergelt and N. Deo, *Combinatorial Algorithms: Theory and Practice* (Prentice-Hall, Englewood Cliffs, N.J., 1977), Chapter 8; A. Gibbons, *Algorithmic Graph Theory* (Cambridge University Press, 1985), Chapter 1; S. Even, *Graph Algorithms* (Computer Science Press, Potomac, Maryland, 1979), Chapter 3.

[10] B.A. Galler and M.J. Fischer, Commun. ACM **7**, 301, 506 (1964); D.E. Knuth, *The Art of Computer Programming*, vol. 1, 2^{nd} ed., (Addison-Wesley, Reading, Massachusetts, 1973), pp. 353–355, 360, 572; J. Hoshen and R. Kopelman, Phys. Rev. **B14**, 3438 (1976).

[11] K. Binder and D. Stauffer, in *Applications of the Monte Carlo Method in Statistical Physics*, ed. K. Binder (Springer-Verlag, Berlin, 1984), section 8.4; D. Stauffer, *Introduction to Percolation Theory* (Taylor & Francis, London, 1985).

[12] P.W. Kasteleyn and C.M. Fortuin, J. Phys. Soc. Japan **26** (Suppl.), 11 (1969); C.M. Fortuin and P.W. Kasteleyn, Physica **57**, 536 (1972); C.M. Fortuin, Physica **58**, 393 (1972); C.M. Fortuin, Physica **59**, 545 (1972).

[13] C.-K. Hu, Phys. Rev. B **29**, 5103 (1984); T.A. Larsson, J. Phys. A: Math. Gen. **19**, 2383 (1986) and **20**, 2239 (1987).

[14] R.G. Edwards and A.D. Sokal, Phys. Rev. D **38**, 2009 (1988).

[15] J. Goodman and A.D. Sokal, Phys. Rev. Lett. **56**, 1015 (1986).

[16] J. Goodman and A.D. Sokal, Phys. Rev. D **40**, 2035 (1989).

[17] R.G. Edwards, J. Goodman and A.D. Sokal, Nucl. Phys. B **354**, 289 (1991).

[18] R.G. Edwards, S.J. Ferreira, J. Goodman and A.D. Sokal, Nucl. Phys. B **380**, 621 (1992).

[19] T. Mendes and A.D. Sokal, Phys. Rev. D **53**, 3438 (1996), hep-lat/9503024.

[20] T. Mendes, A. Pelissetto and A.D. Sokal, Nucl. Phys. B **477**, 203 (1996), hep-lat/9604015.

[21] G. Mana, A. Pelissetto and A.D. Sokal, Phys. Rev. D **55**, 3674 (1997), hep-lat/9610021.

[22] G. Parisi, in *Progress in Gauge Field Theory* (1983 Cargèse lectures), ed. G. 't Hooft *et al.* (Plenum, New York, 1984); G.G. Batrouni, G.R. Katz, A.S. Kronfeld, G.P. Lepage, B. Svetitsky and K.G. Wilson, Phys. Rev. D **32**, 2736 (1985); C. Davies, G. Batrouni, G. Katz, A. Kronfeld, P. Lepage, P. Rossi, B. Svetitsky and K. Wilson, J. Stat. Phys. **43**, 1073 (1986); J.B. Kogut, Nucl. Phys. B **275** [FS17], 1 (1986); S. Duane, R. Kenway, B.J. Pendleton and D. Roweth, Phys. Lett. B **176**, 143 (1986); E. Dagotto and J.B. Kogut, Phys. Rev. Lett. **58**, 299 (1987) and Nucl. Phys. B **290** [FS20], 451 (1987).

[23] J. Salas and A.D. Sokal, Universal amplitude ratios in the critical two-dimensional Ising model on a torus, cond-mat/9904038v1, see Section 6. For space reasons, this material was deleted from the published version of this paper [J. Stat. Phys. **98**, 551 (2000)].

[24] N. Ito, M. Taiji and M. Suzuki, J. Phys. Soc. Japan **56**, 4218 (1987).

[25] J. Salas and A.D. Sokal, J. Stat. Phys. **87**, 1 (1997), hep-lat/9605018.

[26] J. Salas and A.D. Sokal, J. Stat. Phys. **85**, 297 (1996), hep-lat/9511022.

[27] J. Salas and A.D. Sokal, J. Stat. Phys. **88**, 567 (1997), hep-lat/9607030.

[28] G. Ossola and A.D. Sokal, Nucl. Phys. B **691**, 259 (2004), hep-lat/0402019.

[29] W. Klein, T. Ray and P. Tamayo, Phys. Rev. Lett. **62**, 163 (1989).

[30] T. Ray, P. Tamayo and W. Klein, Phys. Rev. A **39**, 5949 (1989).

[31] P.D. Coddington and C.F. Baillie, Phys. Rev. Lett. **68**, 962 (1992).

[32] N. Persky, R. Ben-Av, I. Kanter and E. Domany, Phys. Rev. E **54**, 2351 (1996), cond-mat/9603134.

[33] X.-J. Li and A.D. Sokal, Phys. Rev. Lett. **63**, 827 (1989).

[34] M. Hasenbusch, K. Pinn and S. Vinti, Phys. Rev. B **59**, 11471 (1999), hep-lat/9806012.

[35] L. Chayes and J. Machta, Physica A **254**, 477 (1998), see Section 2.1.

[36] G. Ossola, M. Polin and A.D. Sokal, in preparation.

[37] Y. Deng, T.M. Garoni, J. Machta and A.D. Sokal, in preparation.

[38] C. Borgs, J.T. Chayes, A. Frieze, J.H. Kim, P. Tetali, E. Vigoda and V.H. Vu, in *40th Annual Symposium on Foundations of Computer Science* (IEEE Computer Society, Los Alamitos, CA, 1999), pp. 218–229.

[39] V.K. Gore and M.R. Jerrum, J. Stat. Phys. **97**, 67 (1999).