

Monte Carlo Simulations of Spin Systems

Wolfhard Janke

Institut für Physik, Johannes Gutenberg-Universität, D-55099 Mainz, Germany

Abstract. This lecture gives a brief introduction to Monte Carlo simulations of classical $O(n)$ spin systems such as the Ising ($n = 1$), XY ($n = 2$), and Heisenberg ($n = 3$) model. In the first part I discuss some aspects of Monte Carlo algorithms to generate the raw data. Here special emphasis is placed on non-local cluster update algorithms which proved to be most efficient for this class of models. The second part is devoted to the data analysis at a continuous phase transition. For the example of the three-dimensional Heisenberg model it is shown how precise estimates of the transition temperature and the critical exponents can be extracted from the raw data. I conclude with a brief overview of recent results from similar high-precision studies of the Ising and XY model.

1 Introduction

The statistical mechanics of complex physical systems poses many hard problems which are difficult to solve by analytical approaches. Numerical simulation techniques will therefore be indispensable tools on our way to a better understanding of systems like (spin) glasses, disordered magnets, or proteins, to mention only a few classical problems. Quantum statistical problems in condensed matter or the broad field of elementary particles and quantum gravity in high-energy physics would fill many other volumes like this.

The numerical tools can roughly be divided into molecular dynamics (MD) and Monte Carlo (MC) simulations. With the ongoing advances in computer technology both approaches are expected to gain even more importance than they already have today. In the past few years the predictive power of especially the MC approach was considerably enhanced by the development of greatly improved simulation techniques. Not all of them are already well enough understood to be applicable to really complex physical systems. But as a first step it is gratifying to see that at least for relatively simple spin systems orders of magnitude of computing time can be saved by these refinements. The purpose of this lecture is to give a brief overview on what is feasible today.

From a theoretical view also spin systems are of current interest since on the one hand they provide the possibility to compare completely different approaches such as field theory, series expansions, and simulations, and on the other hand they are the ideal testing ground for conceptual considerations such as universality or finite-size scaling. And last but not least they have found a revival in slightly disguised form in quantum gravity and conformal field theory, where they serve as idealized “matter” fields on Feynman graphs or fluctuating manifolds.

The rest of the paper is organized as follows. In the next section I first recall the definition of $O(n)$ spin models and the definition of standard observables like the specific heat and the susceptibility. Then some properties of phase transitions are summarized and the critical exponents are defined. In Sec. 3, Monte Carlo methods are described, and Sec. 4 is devoted to an overview of reweighting techniques. In Sec. 5, applications to the three-dimensional classical Heisenberg model are discussed, and in Sec. 6, I conclude with a few comments on similar simulations of the Ising and XY model.

2 Spin Models and Phase Transitions

2.1 Models and Observables

In the following we shall confine ourselves to $O(n)$ symmetric spin models whose partition function is defined as

$$Z_n(\beta) = \sum_{\{\boldsymbol{\sigma}_i\}} \exp(-\beta H_n) , \quad (1)$$

with

$$H_n = -J \sum_{\langle ij \rangle} \boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_j; \quad \boldsymbol{\sigma}_i = (\sigma_i^{(1)}, \sigma_i^{(2)}, \dots, \sigma_i^{(n)}); \quad |\boldsymbol{\sigma}_i| = 1 . \quad (2)$$

Here $\beta = 1/k_B T$ is the inverse temperature, the spins $\boldsymbol{\sigma}_i$ live on the sites i of a D -dimensional cubic lattice of volume $V = L^D$, and the symbol $\langle ij \rangle$ indicates that the lattice sum runs over all 2D nearest-neighbor pairs. We always assume periodic boundary conditions.

Standard observables are the internal energy per site, $e = E/V$, with $E = -d \ln Z_n / d\beta \equiv \langle H_n \rangle$, and the specific heat,

$$C/k_B = \frac{de}{d(k_B T)} = \beta^2 (\langle H_n^2 \rangle - \langle H_n \rangle^2) / V . \quad (3)$$

On finite lattices the magnetization and susceptibility are usually defined as

$$m = M/V = \langle |\boldsymbol{\sigma}_{\text{av}}| \rangle; \quad \boldsymbol{\sigma}_{\text{av}} = \sum_i \boldsymbol{\sigma}_i / V , \quad (4)$$

$$\chi = \beta V (\langle \boldsymbol{\sigma}_{\text{av}}^2 \rangle - \langle |\boldsymbol{\sigma}_{\text{av}}| \rangle^2) . \quad (5)$$

In the high-temperature phase one often employs the fact that the magnetization vanishes in the infinite volume limit and defines

$$\chi' = \beta V \langle \boldsymbol{\sigma}_{\text{av}}^2 \rangle . \quad (6)$$

Similarly, the spin-spin correlation function can then be taken as

$$G(\mathbf{x}_i - \mathbf{x}_j) = \langle \boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_j \rangle . \quad (7)$$

At large distances, $G(\mathbf{x})$ decays exponentially and the correlation length ξ can be defined as

$$\xi = - \lim_{|\mathbf{x}| \rightarrow \infty} |\mathbf{x}| / \ln G(\mathbf{x}) . \quad (8)$$

For $n = 1$, the partition function $Z_n(\beta)$ describes the standard Ising model where the spins can take only the two discrete values, $\sigma_i \equiv \sigma_i^{(1)} = \pm 1$. For $n = 2, 3, \dots$ the spins vary continuously on the n -dimensional unit sphere. Particularly thoroughly studied cases are $n = 2$ (XY model) and $n = 3$ (Heisenberg model). The limit $n \rightarrow \infty$ is known to be equivalent to the spherical model of Berlin and Kac (1952). In three dimensions (3D) the model exhibits for all n a continuous phase transition from an ordered low-temperature phase to a disordered high-temperature phase. The associated critical exponents are generic for the so-called $O(n)$ universality classes. In two dimensions (2D) the situation is little more complex. For the 2D Ising model the exact solution by Onsager and later Yang predicts a continuous order-disorder phase transition similar to 3D. For all $n \geq 2$, however, the spin degrees of freedom are continuous and, as a consequence of the Mermin-Wagner-Hohenberg theorem, the magnetization vanishes for all temperatures. The 2D XY model nevertheless displays a very peculiar (infinite order) Kosterlitz-Thouless transition (Kosterlitz and Thouless 1973, Berezinskii 1971, Kosterlitz 1974). Due to the $O(2)$ symmetry this model admits point like topological defects (vortices) which are tightly bound to pairs at low temperatures. With increasing temperature isolated vortices are entropically favored, and the transition is usually pictured as the point where vortex pairs start to dissociate (for a review see, e.g., Kleinert (1989a)). For the 2D Heisenberg model and all other 2D $O(n)$ models with $n > 3$, on the other hand, it is commonly believed that there is no phase transition at finite temperature.¹

For later reference we also recall another generalization of the Ising model, the q -state Potts model (Potts 1952) with Hamiltonian

$$H_{\text{Potts}} = -J \sum_{\langle ij \rangle} \delta_{\sigma_i \sigma_j}; \quad \sigma_i \in 1, \dots, q . \quad (9)$$

This generalization has in 3D for all $q \geq 3$ a first-order transition; and in 2D it is exactly known to exhibit a second-order transition for $q \leq 4$, and a first-order transition for all $q \geq 5$ (Wu 1982, 1983).

2.2 Phase Transitions

In limiting cases like low and high temperatures (or fields, pressure, etc.) the physical degrees of freedom usually decouple and the statistical mechanics of even complex systems becomes quite manageable. Much more interesting is the region in between these extremes where strong cooperation effects may cause phase transitions, e.g., from an ordered phase at low temperatures to a disordered

¹ For an alternative view see, however, Patrascioiu and Seiler (1995) and references to earlier work therein.

phase at high temperatures. To predict the properties of this most difficult region of a phase diagram as accurately as possible is the most challenging objective of all statistical mechanics approaches, including numerical simulations.

The theory of phase transitions is a very broad subject described comprehensively in many textbooks. Here we shall be content with a rough classification in *first-order* and *second-order* (or, more generally, continuous) phase transitions, and a summary of those properties that are most relevant for numerical simulations. Some characteristic properties of first- and second-order phase transitions are sketched in Fig. 1.

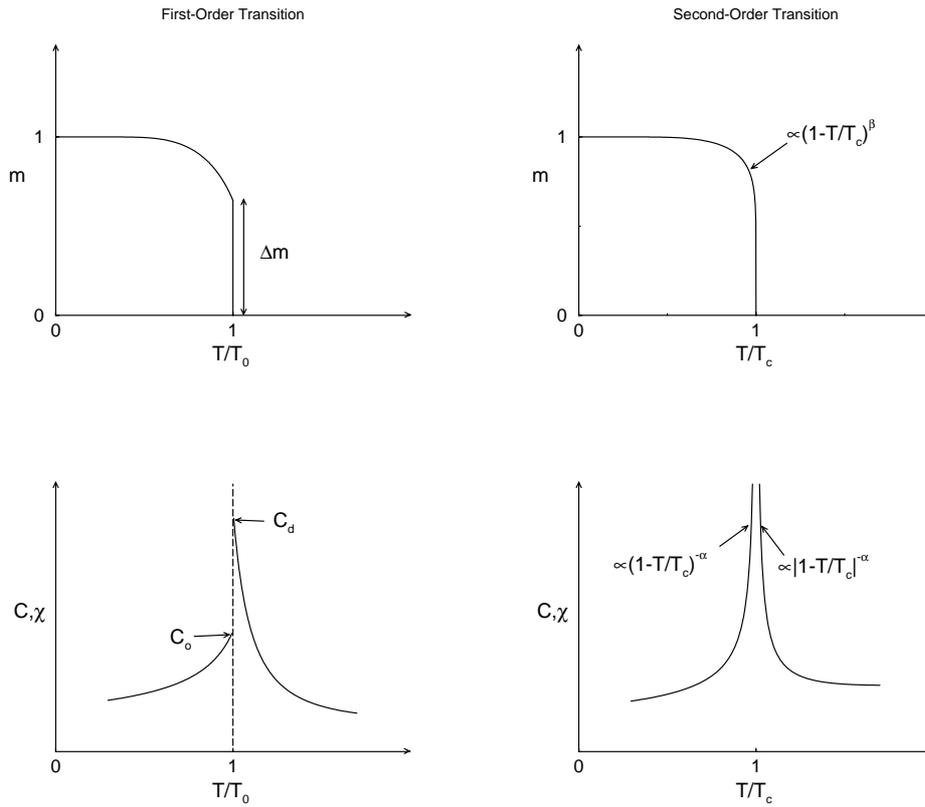


Fig. 1. The characteristic behaviour of the magnetization, m , specific heat, C , and susceptibility, χ , at first- and second-order phase transitions.

Most phase transitions in Nature are of first order (Gunton et al. 1983, Binder 1987, Herrmann et al. 1992, Janke 1994). The best known example is the field-driven transition in magnets at temperatures below the Curie point, while the paradigm of a temperature-driven first-order transition experienced every day

is ordinary melting (Kleinert 1989b, Janke and Kleinert 1986). In general, first-order phase transitions are characterized by *discontinuities* of the order parameter (the jump Δm of the magnetization m in Fig. 1), or the energy (the latent heat Δe), or both. This reflects the fact that, at the transition temperature T_0 , two (or more) phases can coexist. In the example of a magnet at low temperatures the coexisting phases are the phases with positive and negative magnetization, while at the melting transition they are the solid (ordered) and liquid (disordered) phases. The correlation length in the coexisting pure phases is finite.² Consequently also the specific heat, C , and the susceptibility, χ , do not diverge in the pure phases. Mathematically there are, however, superimposed delta function like singularities associated with the jumps of e and m .

In this lecture we will mainly consider second-order phase transitions, which are characterized by a *divergent* correlation length ξ at the transition temperature $T_c \equiv 1/\beta_c$. The growth of correlations as one approaches the critical region from high temperatures is illustrated in Fig. 2, where six typical configurations of the 2D Ising model at inverse temperatures $\beta/\beta_c = 0.50, 0.70, 0.85, 0.90, 0.95$, and 0.98 are shown. Because for an infinite correlation length fluctuations on all length scales are equally important, one expects power-law singularities in thermodynamic functions. The leading singularity of the correlation length is usually parametrized in the high-temperature phase as

$$\xi = \xi_{0+} |1 - T/T_c|^{-\nu} + \dots \quad (T \geq T_c) , \quad (10)$$

where the \dots indicate subleading corrections (analytical as well as confluent). This defines the critical exponent ν and the critical amplitude ξ_{0+} on the high-temperature side of the transition. In the low-temperature phase one expects a similar behaviour,

$$\xi = \xi_{0-} (1 - T/T_c)^{-\nu} + \dots \quad (T \leq T_c) , \quad (11)$$

with the same critical exponent ν but a different critical amplitude $\xi_{0-} \neq \xi_{0+}$.

An important feature of second-order phase transitions is that due to the divergence of ξ the short-distance details of the Hamiltonian should not matter. This is the basis of the universality hypothesis which states that all systems with the same symmetries and same dimensionality should exhibit similar singularities governed by one and the same set of critical exponents. For the amplitudes this is not true, but certain amplitude ratios are also universal.

The singularities of the specific heat, magnetization, and susceptibility are similarly parametrized by the critical exponents α , β , and γ , respectively,

$$C = C_{\text{reg}} + C_0 |1 - T/T_c|^{-\alpha} + \dots , \quad (12)$$

$$m = m_0 (1 - T/T_c)^\beta + \dots , \quad (13)$$

$$\chi = \chi_0 |1 - T/T_c|^{-\gamma} + \dots , \quad (14)$$

² For the 2D q -state Potts model with $q \geq 5$, where many exact results are known, this is illustrated by the recent simulations of Janke and Kappler (1994, 1995a, 1995b, 1995c, 1995d); for details see Kappler (1995).

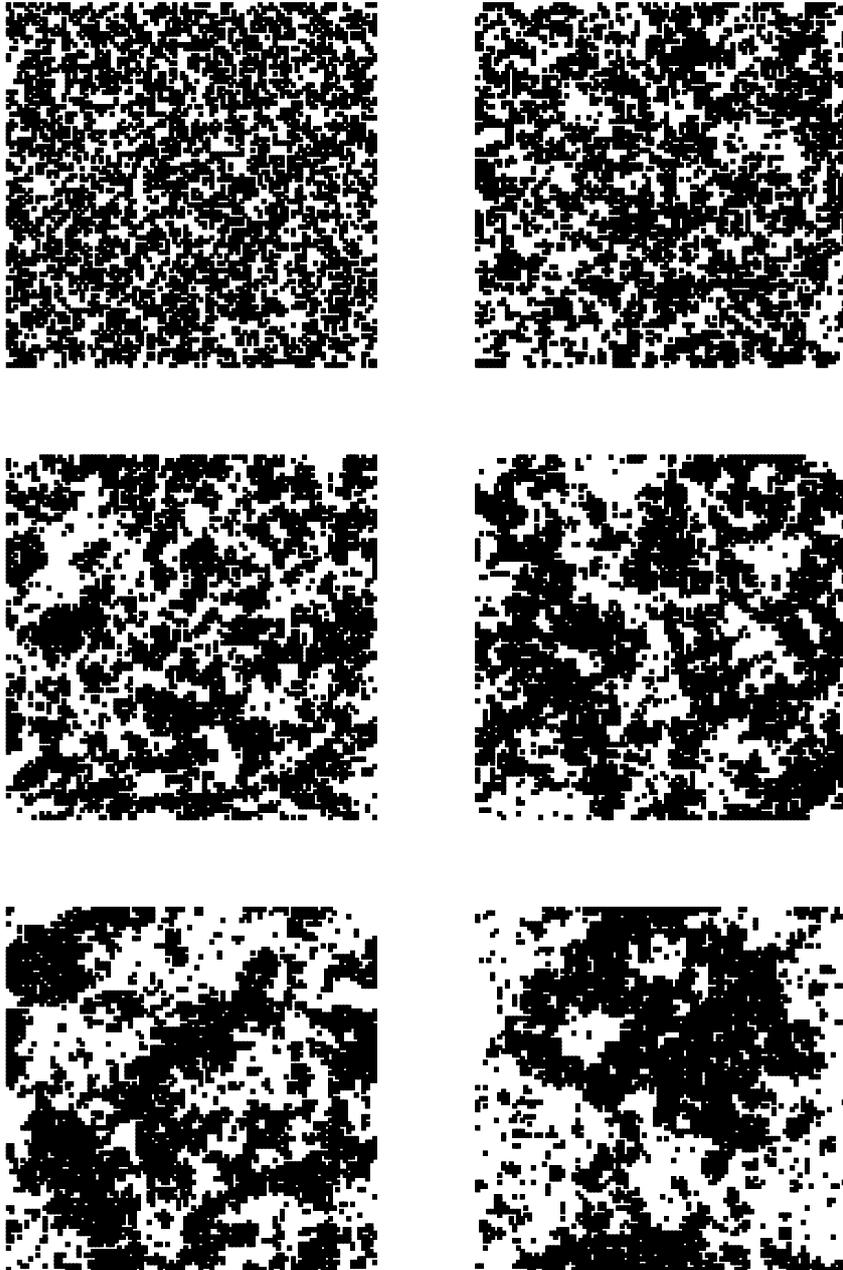


Fig. 2. From high temperatures (upper left) to the critical region (lower right), characterized by large spatial correlations. Shown are actual 2D Ising configurations for a 100×100 lattice at $\beta/\beta_c = 0.50, 0.70, 0.85, 0.90, 0.95,$ and 0.98 .

where C_{reg} is a regular background term, and the amplitudes are again different on the two sides of the transition, cp. Fig. 1.

Finite-Size Scaling. For systems of finite size, as in any numerical simulation, the correlation length cannot diverge, and also the divergencies in all other quantities are then rounded and shifted. This is illustrated in Fig. 3, where the specific heat of the 2D Ising model on various $L \times L$ lattices is shown. The curves are computed from the exact solution of Ferdinand and Fisher (1969) for any $L_x \times L_y$ lattice with periodic boundary conditions.

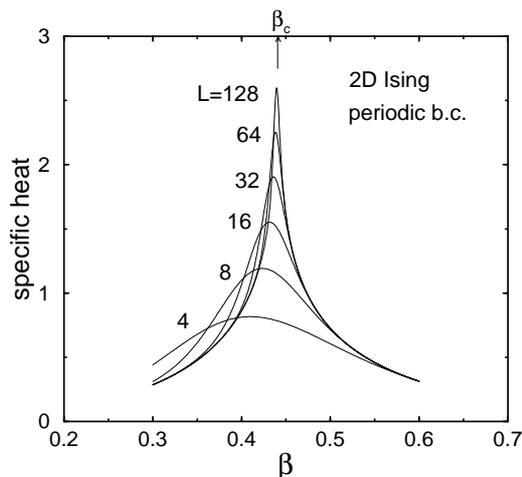


Fig. 3. Finite-size scaling behaviour of the specific heat of the 2D Ising model on $L \times L$ lattices. The critical point is indicated by the arrow on the top axis.

Near T_c the role of ξ in the scaling formulas is then taken over by the linear size of the system, L . By writing $|1 - T/T_c| \propto \xi^{-1/\nu} \rightarrow L^{-1/\nu}$, we see that at T_c the scaling laws (12)-(14) are replaced by the *finite-size scaling* (FSS) Ansätze,

$$C = C_{\text{reg}} + aL^{\alpha/\nu} + \dots, \quad (15)$$

$$m \propto L^{-\beta/\nu} + \dots, \quad (16)$$

$$\chi \propto L^{\gamma/\nu} + \dots. \quad (17)$$

More generally these scaling laws are valid in the vicinity of T_c as long as the scaling variable $x = (1 - T/T_c)L^{1/\nu}$ is kept fixed (Binder 1979, Barber 1983, Privman 1990, Binder 1992b). In particular this is true for the locations T_{max} of the (finite) maxima of thermodynamic quantities, which are expected to scale with the system size as $T_{\text{max}} = T_c(1 - x_{\text{max}}L^{-1/\nu} + \dots)$. In this more general

formulation the scaling law for, e.g., the susceptibility reads $\chi(T, L) = L^{\gamma/\nu} f(x)$. By plotting $\chi(T, L)/L^{\gamma/\nu}$ vs the scaling variable x , one thus expects that the data for different T and L fall onto a kind of master curve. While this is a nice way to demonstrate the scaling properties qualitatively, it is not particularly suited for quantitative analyses.

Since the goal of most simulation studies of spin systems are high-precision estimates of the critical temperature and the critical exponents, one therefore prefers fits either to the “thermodynamic” scaling laws (12)-(14) or to the FSS Ansätze (15)-(17).

Similar considerations for first-order phase transitions show that also the delta function like singularities, originating from the phase coexistence, are smeared out for finite systems (Fisher and Berker 1982, Privman and Fisher 1983, Binder and Landau 1984, Challa et al. 1986, Privman and Rudnik 1990). In finite systems they are replaced by narrow peaks whose height (width) grows proportional to the volume (1/volume) (Borgs and Kotecky 1990, 1992, Lee and Kosterlitz 1990, 1991, Borgs et al. 1991, Borgs and Janke 1992, Janke 1993).

3 The Monte Carlo Method

Let us now discuss how the expectation values in (3)-(7) can be computed numerically. A direct summation of the partition function is impossible, since already for the Ising model with only two possible states per site the number of terms would be enormous: $2^{2500} \approx 10^{753}$ for a 50×50 lattice! Also a naive random sampling of the spin configurations does not work. Here the problem is that the relevant region in the high-dimensional phase space is relatively narrow and hence too rarely hit by random sampling. The solution to this problem is known since long. One has to use the *importance sampling* technique (Hammersley and Handscomb 1965) which is designed to draw configurations according to their Boltzmann weight,

$$P^{\text{eq}}[\{\sigma_i\}] \propto \exp(-\beta H[\{\sigma_i\}]) . \quad (18)$$

In more mathematical terms one sets up a Markov chain,

$$\dots \xrightarrow{W} \{\sigma_i\} \xrightarrow{W} \{\sigma'_i\} \xrightarrow{W} \{\sigma''_i\} \xrightarrow{W} \dots ,$$

with a transition operator W satisfying the conditions

$$\text{a) } W(\{\sigma_i\} \longrightarrow \{\sigma'_i\}) \geq 0 \quad \text{for all } \{\sigma_i\}, \{\sigma'_i\} , \quad (19)$$

$$\text{b) } \sum_{\{\sigma'_i\}} W(\{\sigma_i\} \longrightarrow \{\sigma'_i\}) = 1 \quad \text{for all } \{\sigma_i\} , \quad (20)$$

$$\text{c) } \sum_{\{\sigma_i\}} W(\{\sigma_i\} \longrightarrow \{\sigma'_i\}) P^{\text{eq}}[\{\sigma_i\}] = P^{\text{eq}}[\{\sigma'_i\}] \quad \text{for all } \{\sigma'_i\} . \quad (21)$$

From (21) we see that P^{eq} is a fixed point of W . A somewhat simpler sufficient condition is detailed balance,

$$P^{\text{eq}}[\{\sigma_i\}] W(\{\sigma_i\} \longrightarrow \{\sigma'_i\}) = P^{\text{eq}}[\{\sigma'_i\}] W(\{\sigma'_i\} \longrightarrow \{\sigma_i\}) . \quad (22)$$

After an initial equilibration time, expectation values can then be estimated as an arithmetic mean over the Markov chain, e.g.,

$$\langle H \rangle = \sum_{\{\sigma_i\}} H[\{\sigma_i\}] P^{\text{eq}}[\{\sigma_i\}] \approx \frac{1}{N} \sum_{j=1}^N H[\{\sigma_i\}]_j . \quad (23)$$

A more detailed exposition of the basic concepts underlying any Monte Carlo algorithm can be found in many textbooks and reviews (Binder 1979, Binder and Heermann 1988, Heermann 1990, Binder 1992a).

3.1 Estimators and Autocorrelation Times

In principle there is no limitation on the choice of observables. The expectation value $\langle \mathcal{O} \rangle$ of any observable \mathcal{O} can be estimated in a MC simulation as a simple arithmetic mean over the Markov chain, $\bar{\mathcal{O}} = \frac{1}{N} \sum_{j=1}^N \mathcal{O}_j$, where $\mathcal{O}_j \equiv \mathcal{O}[\{\sigma_i\}]_j$ is the measurement after the j -th iteration. Conceptually it is important to distinguish between the expectation value $\langle \mathcal{O} \rangle$, which is an ordinary number, and the estimator $\bar{\mathcal{O}}$, which is a *random* number fluctuating around the theoretically expected value. Of course, in practice one does not probe the fluctuations of the estimator directly (which would require repeating the whole MC simulation many times), but rather estimates its variance $\sigma_{\bar{\mathcal{O}}}^2 = \langle [\bar{\mathcal{O}} - \langle \bar{\mathcal{O}} \rangle]^2 \rangle = \langle \bar{\mathcal{O}}^2 \rangle - \langle \bar{\mathcal{O}} \rangle^2$ from the distribution of \mathcal{O}_j . If the N measurements \mathcal{O}_j were all uncorrelated then the relation would simply be $\sigma_{\bar{\mathcal{O}}}^2 = \sigma_{\mathcal{O}_j}^2 / N$, with $\sigma_{\mathcal{O}_j}^2 = \langle \mathcal{O}_j^2 \rangle - \langle \mathcal{O}_j \rangle^2$. For correlated measurements one obtains after some algebra

$$\sigma_{\bar{\mathcal{O}}}^2 = \frac{\sigma_{\mathcal{O}_k}^2}{N} 2\tau_{\bar{\mathcal{O}},\text{int}} , \quad (24)$$

where the *integrated* autocorrelation time,

$$\tau_{\bar{\mathcal{O}},\text{int}} = \frac{1}{2} + \sum_{j=1}^N A(j) \left(1 - \frac{j}{N} \right) , \quad (25)$$

turns out to be a sum (“integral”) over the the autocorrelation function,

$$A(j) = \frac{\langle \mathcal{O}_i \mathcal{O}_{i+j} \rangle - \langle \mathcal{O}_i \rangle \langle \mathcal{O}_i \rangle}{\langle \mathcal{O}_i^2 \rangle - \langle \mathcal{O}_i \rangle \langle \mathcal{O}_i \rangle} . \quad (26)$$

For large time separations the autocorrelation function decays exponentially,

$$A(j) \xrightarrow{j \rightarrow \infty} a e^{-j/\tau_{\bar{\mathcal{O}},\text{exp}}} , \quad (27)$$

with a being a constant. This defines the *exponential* autocorrelation time $\tau_{\bar{\mathcal{O}},\text{exp}}$. Due to the exponential decay of $A(j)$, in any meaningful simulation with $N \gg \tau_{\bar{\mathcal{O}},\text{exp}}$, the correction term in parentheses in (25) can safely be neglected. Notice that only if $A(j)$ is a pure exponential, the two autocorrelation times, $\tau_{\bar{\mathcal{O}},\text{int}}$ and

$\tau_{\bar{O},\text{exp}}$, coincide (up to minor corrections for small $\tau_{\bar{O},\text{int}}$) (Madras and Sokal 1988).

The important point is that for correlated measurements the statistical error $\epsilon_{\bar{O}} \equiv \sqrt{\sigma_{\bar{O}}^2}$ on the MC estimator \bar{O} is enhanced by a factor of $\sqrt{2\tau_{\bar{O},\text{int}}}$. This can be rephrased by writing the statistical error similar to the uncorrelated case as $\epsilon_{\bar{O}} = \sqrt{\sigma_{\bar{O}_j}^2/N_{\text{eff}}}$, but now with an effective statistics parameter $N_{\text{eff}} = N/2\tau_{\bar{O},\text{int}}$. This shows more clearly that only every $2\tau_{\bar{O},\text{int}}$ iteration the measurements are approximately uncorrelated and gives a better idea of the relevant effective size of the statistical sample. Since some quantities (e.g., the specific heat or susceptibility) can severely be underestimated if the effective statistics is too small (Ferrenberg et al. 1991), any serious simulation should therefore provide at least a rough order-of-magnitude estimate of autocorrelation times.

Unfortunately, it is very difficult to give reliable a priori estimates, and an accurate numerical analysis is often too time consuming (as a rough estimate it is about ten times harder to get precise information on dynamic quantities than on static quantities like critical exponents). To get at least an idea of the orders of magnitude, it is useful to record the “running” autocorrelation time

$$\tau_{\bar{O},\text{int}}(k) = \frac{1}{2} + \sum_{j=1}^k A(j) , \quad (28)$$

which approaches $\tau_{\bar{O},\text{int}}$ for large k . Approximating the tail end of $A(j)$ by a single exponential as in (27), one derives (Janke and Sauer 1995)

$$\tau_{\bar{O},\text{int}}(k) = \tau_{\bar{O},\text{int}} - \frac{a}{e^{1/\tau_{\bar{O},\text{exp}}} - 1} e^{-k/\tau_{\bar{O},\text{exp}}} . \quad (29)$$

The latter expression may be used for a numerical estimate of both the exponential and integrated autocorrelation times.

To summarize this subsection, any realization of a Markov chain (i.e., MC update algorithm) is characterized by autocorrelation times which enter directly in the statistical errors of MC estimates. Since correlations always increase the statistical errors, it is a very important issue to develop MC update algorithms that keep autocorrelation times as small as possible. In the next subsection we first discuss the classical Metropolis algorithm as an example of an update algorithm which near criticality is plagued by huge temporal correlations. The discussion of cluster updates in the next subsection then demonstrates that there indeed exist clever ways to overcome this critical slowing down problem.

3.2 Metropolis Algorithm

In the standard Metropolis algorithm (Metropolis et al. 1953) the Markov chain is realized by *local* updates of single spins. If E and E' denote the energy before

and after the spin flip, respectively, then the probability to accept the proposed spin update is given by

$$W(\{\sigma_i\} \longrightarrow \{\sigma'_i\}) = \begin{cases} \exp[-\beta(E' - E)] & E' \geq E \\ 1 & E' < E \end{cases} . \quad (30)$$

If the energy is lowered, the spin flip is always accepted. But even if the energy is increased, the flip has to be accepted with a certain probability to ensure the proper treatment of entropic contributions. In thermal equilibrium the *free* energy is minimized and not the energy. Only at zero temperature ($\beta \rightarrow \infty$) this probability tends to zero and the MC algorithm degenerates to a minimization algorithm for the energy functional. With some additional refinements, this is the basis of the simulated annealing technique also discussed in this volume, which is usually applied to hard optimization and minimization problems.

There are many ways how to choose the spins to be updated. The lattice sites may be picked at random or according to a random permutation, which can be updated every now and then. But also a simple fixed lexicographical order is permissible. Or one updates first all odd and then all even sites, which is the usual choice in vectorized codes. A so-called lattice sweep is completed when on the average³ for all spins an update was proposed.

The advantage of this simple algorithm is its flexibility which allows the application to a great variety of physical systems. The great disadvantage, however, is that this algorithm is plagued by large autocorrelation times, as most other *local* update algorithms (one exception is the overrelaxation method (Creutz 1987, Adler 1988, Neuberger 1988, Gupta et al. 1988)). Empirically one finds that the autocorrelation time grows proportional to the spatial correlation length,

$$\tau \propto \xi^z , \quad (31)$$

with a dynamical critical exponent $z \approx 2$. Heuristically this can be understood by assuming that local excitations diffuse through the system like a random walk. Since ξ diverges at criticality, the Metropolis algorithm thus severely suffers from *critical slowing down*. Of course, in finite systems ξ cannot diverge. Then ξ is replaced by the linear lattice size L , yielding $\tau \propto L^z$.

The problem of critical slowing down can be overcome by *non-local* update algorithms. In the past few years several different types of such algorithms have been proposed. Quite promising results were reported with Fourier acceleration (Doll et al. 1985) and multigrid techniques (Goodman and Sokal 1986, 1989, Mack 1988, Kandel et al. 1988, 1989, Mack and Meyer 1990). A very nice and pedagogical introduction to these techniques is given by Sokal (1989, 1992). For the $O(n)$ spin models considered here, however, the best performance was achieved with cluster update algorithms which will be described in the next subsection in more detail.

³ This is only relevant in the case where the lattice sites are picked at random.

3.3 Cluster Algorithms

As we shall see below, cluster update algorithms (Swendsen and Wang 1987, Wolff 1989a) are much more powerful than the Metropolis algorithm. Unfortunately, however, they are less general applicable. We therefore consider first only the Ising model, where the prescription for cluster update algorithms can easily be read off from the equivalent Fortuin-Kasteleyn representation (Kasteleyn and Fortuin 1969, Fortuin and Kasteleyn 1972, Fortuin 1972a, 1972b),

$$Z = \sum_{\{\sigma_i\}} \exp \left(\beta \sum_{\langle ij \rangle} \sigma_i \sigma_j \right) \quad (32)$$

$$= \sum_{\{\sigma_i\}} \prod_{\langle ij \rangle} e^{\beta} [(1-p) + p\delta_{\sigma_i\sigma_j}] \quad (33)$$

$$= \sum_{\{\sigma_i\}} \sum_{\{n_{ij}\}} \prod_{\langle ij \rangle} e^{\beta} [(1-p)\delta_{n_{ij},0} + p\delta_{\sigma_i\sigma_j}\delta_{n_{ij},1}] \quad , \quad (34)$$

with

$$p = 1 - e^{-2\beta} \quad . \quad (35)$$

Here the n_{ij} are bond variables which can take the values $n_{ij} = 0$ or 1 , interpreted as “deleted” or “active” bonds. In the first line we used the trivial fact that the product $\sigma_i\sigma_j$ of two Ising spins can only take the two values ± 1 , so that $\exp(\beta\sigma_i\sigma_j) = x + y\delta_{\sigma_i\sigma_j}$ can easily be solved for x and y . And in the second line we made use of the “deep” identity $a + b = \sum_{n=0}^1 (a\delta_{n,0} + b\delta_{n,1})$.

Swendsen-Wang Cluster. According to (34) a cluster update sweep then consists of alternating updates of the bond variables n_{ij} for given spins with updates of the spins σ_i for a given bond configuration. In practice one proceeds as follows:

1. Set $n_{ij} = 0$ if $\sigma_i \neq \sigma_j$, or assign values $n_{ij} = 1$ and 0 with probability p and $1 - p$, respectively, if $\sigma_i = \sigma_j$, cp. Fig. 4.
2. Identify clusters of spins that are connected by “active” bonds ($n_{ij} = 1$).
3. Draw a random value ± 1 independently for each cluster (including one-site clusters), which is then assigned to all spins in a cluster.

Technically the cluster identification part is the most complicated step, but there are by now quite a few efficient algorithms available which can even be used on parallel computers. Vectorization, on the other hand, is only partially possible.

Notice the difference between the just defined *stochastic* clusters and *geometric* clusters whose boundaries are defined by drawing lines through bonds between unlike spins. In fact, since in the stochastic cluster definition also bonds between like spins are “deleted” with probability $p_0 = 1 - p = \exp(-2\beta)$, stochastic clusters are on the average smaller than geometric clusters. Only at zero temperature ($\beta \rightarrow \infty$) p_0 approaches zero and the two cluster definitions coincide.

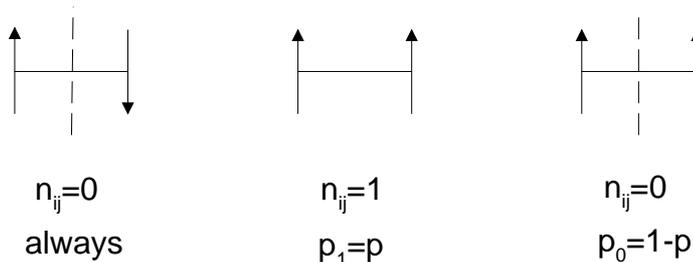


Fig. 4. Illustration of the bond variable update. The bond between unlike spins is always “deleted” as indicated by the dashed line. A bond between like spins is only “active” with probability $p = 1 - \exp(-2\beta)$. Only at zero temperature ($\beta \rightarrow \infty$) stochastic and geometric clusters coincide.

As described above, the cluster algorithm is referred to as Swendsen-Wang (SW) or multiple-cluster update (Swendsen and Wang 1987). The distinguishing point is that the *whole* lattice is decomposed into stochastic clusters whose spins are assigned a random value $+1$ or -1 . In one sweep one thus attempts to update all spins of the lattice.

Wolff Cluster. Shortly after the original discovery of cluster algorithms, Wolff (1989a) proposed a somewhat simpler variant in which only a single cluster is flipped at a time. This variant is therefore sometimes also called single-cluster algorithm. Here one chooses a lattice site at random, constructs only the cluster connected with this site, and then flips all spins of this cluster. A typical example is shown in Fig. 5. In principle, one could also here choose a value $+1$ or -1 at random, but then nothing at all would be changed if one hits the current value of the spins. Here a sweep consists of $V/\langle C \rangle$ single cluster steps, where $\langle C \rangle$ denotes the average cluster size. With this definition autocorrelation times are directly comparable with results from the Metropolis or Swendsen-Wang algorithm. Apart from being somewhat easier to program, Wolff’s single-cluster variant is usually more efficient than the Swendsen-Wang multiple-cluster algorithm, especially in 3D. The reason is that with the single-cluster method, on the average, larger clusters are flipped.

Embedded Cluster. While it is quite easy to generalize the derivation (32) – (35) to q -state Potts models, for the $O(n)$ spin models with $n \geq 2$ one needs a new strategy (Wolff 1989a, 1989b, 1990, Hasenbusch 1990). Here the basic idea is to isolate Ising degrees of freedom by projecting σ_i onto a randomly chosen unit vector \mathbf{r} ,

$$\sigma_i = \sigma_i^{\parallel} + \sigma_i^{\perp}; \quad \sigma_i^{\parallel} = \epsilon |\sigma_i \cdot \mathbf{r}| \mathbf{r}; \quad \epsilon = \text{sign}(\sigma_i \cdot \mathbf{r}) . \quad (36)$$

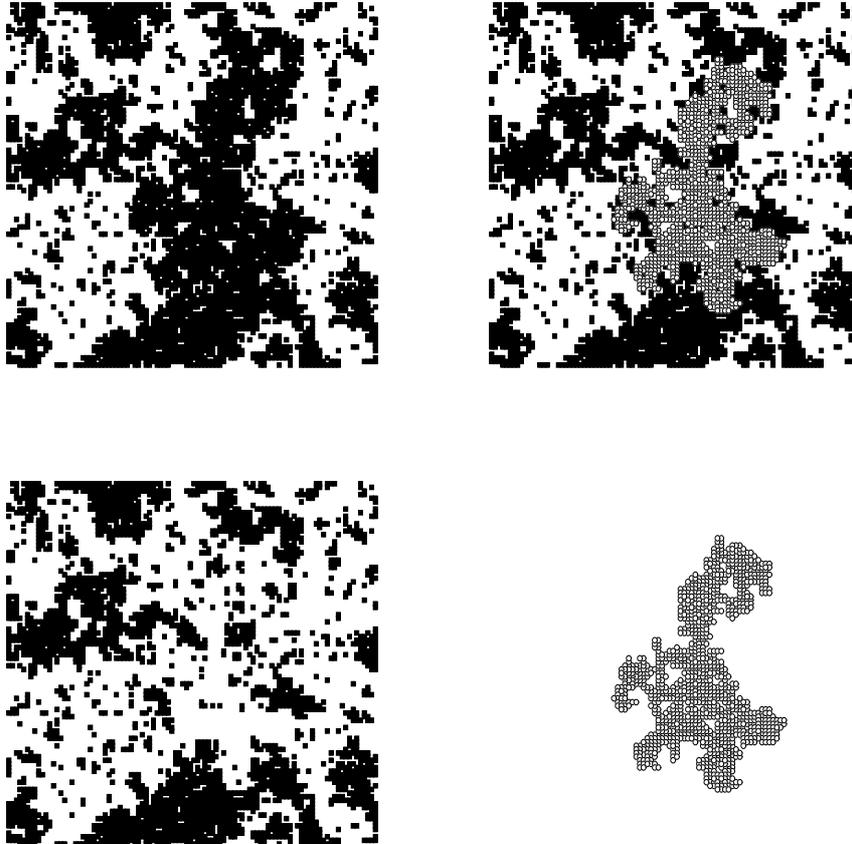


Fig. 5. Illustration of the Wolff cluster update. Upper left: Initial configuration. Upper right: The stochastic cluster is marked. Lower left: Final configuration after flipping the spins in the cluster. Lower right: The flipped cluster. The shown spin configuration is from an actual simulation of the 2D Ising model at $0.97 \times \beta_c$ on a 100×100 lattice.

If this is inserted in the original Hamiltonian one ends up with an effective Hamiltonian

$$H = - \sum_{\langle ij \rangle} J_{ij} \epsilon_i \epsilon_j + \text{const} , \quad (37)$$

with positive random couplings,

$$J_{ij} = J |\boldsymbol{\sigma}_i \cdot \mathbf{r}| |\boldsymbol{\sigma}_j \cdot \mathbf{r}| \geq 0 , \quad (38)$$

whose Ising degrees of freedom ϵ_i can be updated with a cluster algorithm as described above.

For $O(n)$ spin models the performance of both types of cluster algorithms is excellent. As is demonstrated in Table 1 and Fig. 6, critical slowing down is drastically reduced. We see that especially in three dimensions the Wolff cluster algorithm performs better than the Swendsen-Wang algorithm. Compared with the Metropolis algorithm, factors of up to 10 000 in CPU time have been saved in realistic simulations (Baillie 1990, Swendsen et al. 1992).

Table 1. Dynamical critical exponents z for the 2D and 3D Ising model ($\tau \propto L^z$)

algorithm	D=2	D=3	observable	authors
Metropolis	2.125	2.03		
Swendsen-Wang cluster	0.35(1)	0.75(1)	$z_{E,\text{exp}}$	Swendsen and Wang (1987)
	0.27(2)	0.50(3)	$z_{E,\text{int}}$	Wolff (1989c)
	0.20(2)	0.50(3)	$z_{\chi,\text{int}}$	Wolff (1989c)
	$0(\log L)$	–	$z_{M,\text{exp}}$	Heermann and Burkitt (1990)
	0.25(5)	–	$z_{M,\text{rel}}$	Tamayo (1993)
Wolff cluster	0.26(2)	0.28(2)	$z_{E,\text{int}}$	Wolff (1989c)
	0.13(2)	0.14(2)	$z_{\chi,\text{int}}$	Wolff (1989c)
	0.25(5)	0.3(1)	$z_{E,\text{rel}}$	Ito and Kohring (1993)

Improved Estimators. A further advantage of cluster algorithms is that they lead quite naturally to so-called improved estimators which are designed to further reduce the statistical errors. Suppose we want to measure the expectation value $\langle \mathcal{O} \rangle$ of an observable \mathcal{O} . Then any estimator $\hat{\mathcal{O}}$ satisfying $\langle \hat{\mathcal{O}} \rangle = \langle \mathcal{O} \rangle$ is permissible. This does not determine $\hat{\mathcal{O}}$ uniquely since there are infinitely many other possible choices, $\hat{\mathcal{O}}' = \hat{\mathcal{O}} + \hat{\mathcal{X}}$, where the added estimator $\hat{\mathcal{X}}$ has zero expectation, $\langle \hat{\mathcal{X}} \rangle = 0$. The variances of the estimators $\hat{\mathcal{O}}'$, however, can be quite different and are not necessarily related to any physical quantity (contrary to the standard mean-value estimator of the energy whose variance is proportional to the specific heat). It is exactly this freedom in the choice of $\hat{\mathcal{O}}$ which allows the construction of improved estimators.

For the single-cluster algorithm an improved “cluster estimator” for the spin-spin correlation function in the high-temperature phase, $G(\mathbf{x}_i - \mathbf{x}_j) \equiv \langle \boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_j \rangle$, is given by (Wolff 1990)

$$\hat{G}(\mathbf{x}_i - \mathbf{x}_j) = n \frac{V}{|C|} \mathbf{r} \cdot \boldsymbol{\sigma}_i \mathbf{r} \cdot \boldsymbol{\sigma}_j \theta_C(\mathbf{x}_i) \theta_C(\mathbf{x}_j) , \quad (39)$$

where \mathbf{r} is the normal of the mirror plane used in the construction of the cluster of size $|C|$ and $\theta_C(\mathbf{x})$ is its characteristic function (=1 if $\mathbf{x} \in C$ and 0 other-

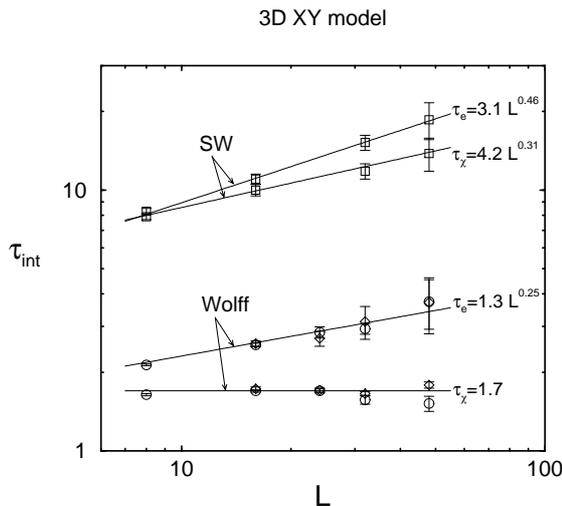


Fig. 6. Double logarithmic plot of the integrated autocorrelation times for the Swendsen-Wang (SW) and Wolff algorithm of the 3D XY model near criticality. The squares ($\beta = 0.45421$) are taken from Hasenbusch and Meyer (1990), the circles ($\beta = 0.4539$) and diamonds ($\beta = 0.4543$) from Janke (1990).

wise). For the Fourier transform, $\tilde{G}(\mathbf{k}) = \sum_{\mathbf{x}} G(\mathbf{x}) \exp(-i\mathbf{k} \cdot \mathbf{x})$, this implies the improved estimator

$$\hat{G}(\mathbf{k}) = \frac{n}{|C|} \left[\left(\sum_{i \in C} \mathbf{r} \cdot \boldsymbol{\sigma}_i \cos \mathbf{kx}_i \right)^2 + \left(\sum_{i \in C} \mathbf{r} \cdot \boldsymbol{\sigma}_i \sin \mathbf{kx}_i \right)^2 \right], \quad (40)$$

which, for $\mathbf{k} = \mathbf{0}$, reduces to an improved estimator for the susceptibility χ' in the high-temperature phase,

$$\hat{G}(\mathbf{0}) = \chi'/\beta = \frac{n}{|C|} \left(\sum_{i \in C} \mathbf{r} \cdot \boldsymbol{\sigma}_i \right)^2. \quad (41)$$

For the Ising model ($n = 1$) this reduces to $\chi'/\beta = \langle |C| \rangle$, i.e., the improved estimator of the susceptibility is just the average cluster size of the single-cluster update algorithm. For the XY and Heisenberg model one finds empirically that in two as well as in three dimensions $\langle |C| \rangle \approx 0.81\chi'/\beta$ for $n = 2$ (Wolff 1989b, Janke 1990) and $\langle |C| \rangle \approx 0.75\chi'/\beta$ for $n = 3$ (Wolff 1990, Holm and Janke 1993b), respectively.

It should be noted that by means of the estimators (39)-(41) a significant reduction of variance should only be expected outside the FSS region where the average cluster size is small compared with the volume of the system.

3.4 Multicanonical Algorithms for First-Order Transitions

Let us finally make a few brief comments on numerical simulations of first-order phase transitions (Janke 1994). Since here the correlation lengths in the pure phases are finite, the numerical problems are completely different than in the case of second-order phase transitions. Here the origin of numerical difficulties near the transition point can be traced back to the coexistence of two phases which, for finite systems, is reflected by a double-peak structure of the corresponding order-parameter or energy distribution. The minimum between the two peaks is governed by mixed phase configurations which are strongly suppressed by an additional Boltzmann factor $\propto \exp(-2\sigma L^{D-1})$. Here σ denotes the interface tension at the phase boundaries, L^{D-1} is the cross-section of the (cubic) system of size $V = L^D$, and the factor 2 takes into account the usually employed periodic boundary conditions (Binder 1981, 1982). The problem of numerical simulations is to achieve equilibrium between the two phases. Most of the time the system spends in the pure phases. Only very rarely it “tunnels” through the exponentially suppressed mixed-phase region from one phase to the other. These rare tunneling events, however, are necessary to achieve equilibrium between the pure phases. The relevant time scale of equilibrium simulations is thus given by the inverse of the additional Boltzmann factor, i.e., the characteristic time τ grows exponentially with the system size, $\tau \propto \exp(2\sigma L^{D-1})$ (Billoire et al. 1991). Since for an accurate numerical study the simulation (and thus computing) time must be much larger than τ , this phenomenon has been termed the *exponential or super-critical slowing down* problem.

A surprisingly simple solution to this problem was discovered by Berg and Neuhaus (1991). In what they call “multicanonical” simulations one determines iteratively artificial weight factors which modify the original Hamiltonian in such a way that the order-parameter or energy distribution is approximately flat between the two peaks of the canonical distribution. Since then the system has no longer to pass through an exponentially suppressed region one expects in multicanonical simulations a drastic reduction of the characteristic time scale τ . A simple random walk argument suggests a power-law behaviour, $\tau \propto V^\alpha$, which has indeed been confirmed in numerical simulations of 2D Potts models (Berg and Neuhaus 1992, Janke et al. 1992, Rummukainen 1993).

The multicanonical technique is strictly speaking not an update algorithm but a reweighting procedure as discussed in detail in the next section. In principle, it can therefore be combined with any legitimate update algorithm. The earlier studies all employed the Metropolis or heat-bath algorithm. In more recent work it was shown that also combinations with multigrid techniques (Janke and Sauer 1994, 1995, Sauer 1994) and cluster update algorithms (Rummukainen 1993, Janke and Kappler 1995e, 1995f, Kappler 1995) are feasible and can further reduce autocorrelation times.

4 Reweighting Techniques

Even though the physics underlying reweighting techniques is extremely simple and the basic idea has been known since long (see the list of references by Ferrenberg and Swendsen (1989a)), their power in practice has been realized only quite recently (Ferrenberg and Swendsen 1988, 1989a). The best performance is achieved *near* criticality, and in this sense reweighting techniques are complementary to improved estimators.

If we denote the number of states (spin configurations) that have the same energy by $\Omega(E)$, the partition function at the simulation point β_0 can always be written as⁴

$$Z(\beta_0) = \sum_E \Omega(E) e^{-\beta_0 E} . \quad (42)$$

This shows that the energy distribution $\mathcal{P}_{\beta_0}(E)$ (normalized to unit area) is given by

$$\mathcal{P}_{\beta_0}(E) = \Omega(E) e^{-\beta_0 E} / Z(\beta_0) . \quad (43)$$

It is then easy to see that, given $\mathcal{P}_{\beta_0}(E)$, the energy distribution is actually known for any β ,

$$\mathcal{P}_\beta(E) = c \mathcal{P}_{\beta_0}(E) e^{-(\beta - \beta_0)E} , \quad (44)$$

where c is a normalization constant (which in practice is trivially determined by enforcing the condition $\sum_E \mathcal{P}_\beta(E) = 1$. Formally, one easily finds that $c = 1 / \sum_E \mathcal{P}_{\beta_0}(E) e^{-(\beta - \beta_0)E} = Z(\beta_0) / Z(\beta)$.) Knowing $\mathcal{P}_\beta(E)$, expectation values of the form $\langle f(E) \rangle$ are easy to compute,

$$\langle f(E) \rangle(\beta) = \sum_E f(E) \mathcal{P}_\beta(E) . \quad (45)$$

Since the relative statistical errors increase in the wings of $\mathcal{P}_{\beta_0}(E)$ one expects (44), (45) to give reliable results only for β near β_0 . If β_0 is near criticality, the distribution is relatively broad and the method works best. In this case reliable estimates from (44) can be expected for β values in an interval around β_0 of width $\propto L^{-1/\nu}$, i.e., just in the FSS region. As a rule of thumb the reweighting range can be determined by the condition that the peak location of the reweighted distribution should not exceed the energy values at which the input distribution had decreased to one third of its maximum value (Alves et al. 1992). In most applications this range is wide enough to locate from a single simulation, e.g., the specific-heat maximum by using any standard maximization routine.

This is illustrated in Figs. 7 and 8, again for simplicity for the 2D Ising model. In Fig. 7 the filled circle shows the result of a MC simulation at $\beta_c = \log(1 + \sqrt{2})/2 \approx 0.440686\dots$, using the Swendsen-Wang cluster algorithm with 5000 sweeps for equilibration and 50 000 sweeps for measurements. The results

⁴ For simplicity we consider here only models with *discrete* energies. If the energy varies continuously, sums have to be replaced by integrals, etc.. Also lattice size dependencies are suppressed to keep the notation short.

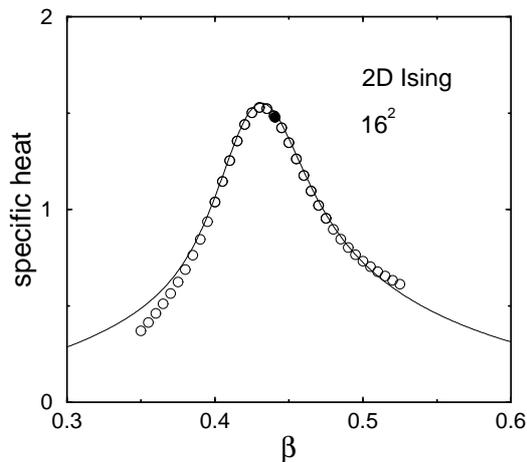


Fig. 7. Specific heat of the 2D Ising model computed by reweighting (\circ) from a single MC simulation at $\beta_0 = \beta_c$ (\bullet). The continuous line shows for comparison the exact result of Ferdinand and Fisher (1969).

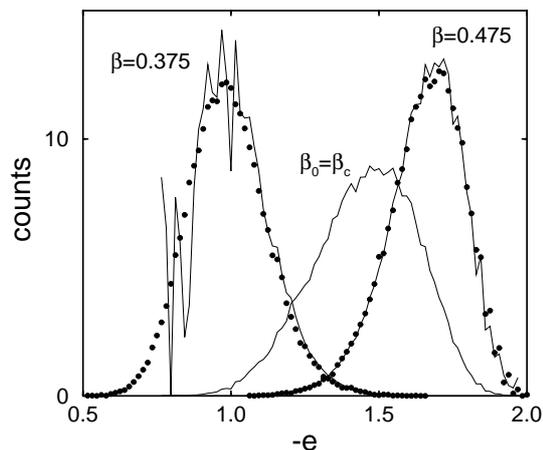


Fig. 8. The energy histogram at the simulation point $\beta_0 = \beta_c$, and reweighted to $\beta = 0.375$ and $\beta = 0.475$. The black dots show histograms obtained in additional simulations at these temperatures.

of the reweighting procedure are shown as open circles (recall that the spacing between the circles can be made as small as desired, here it was chosen quite large for clarity of the plot) and compared with the exact curve (Ferdinand and Fisher 1969). We see that even with this rather modest statistics the whole specific-heat peak can be obtained with reasonable accuracy from a single simulation. But we also notice significant deviations in the tails of the peak. To understand the origin of the deviations it is useful to have a look at the energy histograms in Fig. 8. The

curve labeled $\beta_0 = \beta_c$ is the histogram of the MC data at the simulation point, and the other two histograms at $\beta = 0.375$ and $\beta = 0.475$ are computed from this input histogram by reweighting. For comparison we have also included the histograms obtained from additional MC simulations (with the same statistics) at the two β -values, indicated by the black dots. We see that the reweighted histogram at $\beta = 0.475$ looks smooth to the eye - and indeed agrees very well with the “direct” result of the additional MC simulation at this temperature. In Fig. 7 this is reflected by the still very good agreement of the numerical and the exact result. For $\beta = 0.375$, on the other hand, already visually one would not trust the reweighted histogram. While the tail on the right-hand side is still in reasonable agreement with the “direct” simulation, the left tail is obviously hopelessly wrong. By recalling that the reweighted histograms are computed by multiplying the input histogram with exponential factors, this is no surprise at all: For $-e \lesssim 1$ there are hardly any entries in the input histogram and hence the relative statistical errors ($\propto 1/\sqrt{\text{counts}}$) are huge. This is the source for the large deviations from the exact curve in Fig. 7 for $\beta \lesssim 0.4$.

The information stored in $P_{\beta_0}(E)$ is not yet sufficient to calculate also the magnetic moments $\langle m^k \rangle(\beta)$ as a function of β from a single simulation at β_0 . Conceptually, the simplest way to proceed is to record the two-dimensional histogram $P_{\beta_0}(E, M)$, where $M = mV$ is the total magnetization. Because of disk space limitations one sometimes prefers to measure instead the “microcanonical averages”

$$\langle\langle m^k \rangle\rangle(E) = \sum_M P_{\beta_0}(E, M) m^k / P_{\beta_0}(E) , \quad (46)$$

where the relation $\sum_M P_{\beta_0}(E, M) = P_{\beta_0}(E)$ was used. In practice one simply accumulates the measurements of m^k in different slots or bins according to the energy of the configuration and normalizes at the end by the total number of hits of each energy bin. Clearly, once $\langle\langle m^k \rangle\rangle(E)$ is determined, this is a special case of $f(E)$ in (45), so that

$$\langle m^k \rangle(\beta) = \frac{\sum_E \langle\langle m^k \rangle\rangle(E) P_{\beta_0}(E) e^{-(\beta - \beta_0)E}}{\sum_E P_{\beta_0}(E) e^{-(\beta - \beta_0)E}} . \quad (47)$$

Similar to $\Omega(E)$ in (42), theoretically also the microcanonical averages $\langle\langle m^k \rangle\rangle(E)$ do not depend on the temperature at which the simulation is performed. Due to the limited statistics in the wings of $P_{\beta_0}(E)$, however, there is only a finite range around $E_0 \equiv \langle E \rangle(\beta_0)$ where one can expect reasonable results for $\langle\langle m^k \rangle\rangle(E)$. Outside of this range it simply can happen (and does happen) that there are no events to be averaged. This is illustrated in Fig. 9, where $\langle\langle m^2 \rangle\rangle(E)$ is plotted for the 3D Heisenberg model as obtained from three runs at different temperatures. We see that the function looks smooth only in the range where the statistics of the corresponding energy histogram is high enough. To take full advantage of the histogram reweighting technique it is therefore advisable to perform a few simulations at slightly different inverse temperatures β_i . Instead of spending

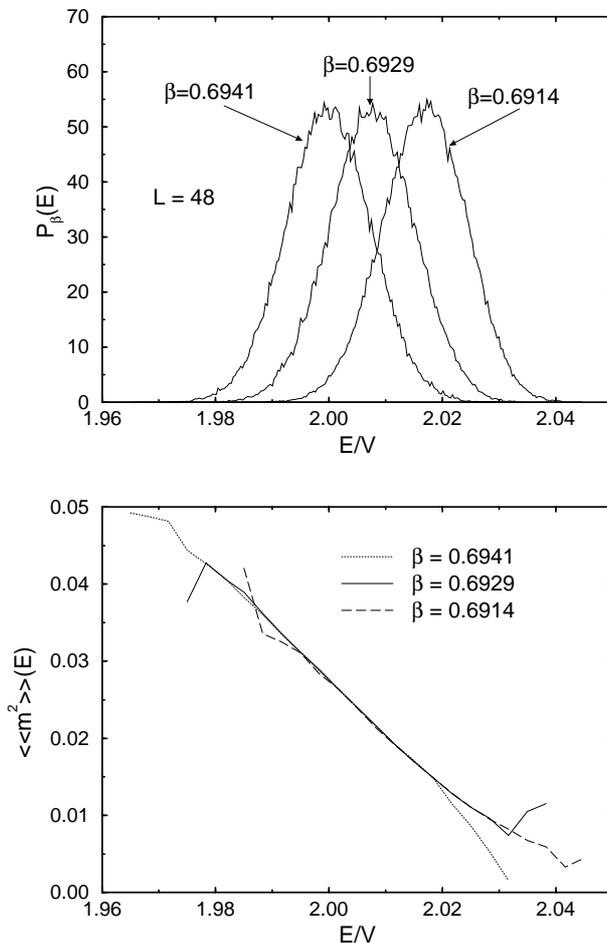


Fig. 9. Energy histograms and microcanonical magnetization squared of the 3D Heisenberg model at the three simulation points for $L = 48$.

all computer time in a single long run, it is usually more efficient to perform three or four shorter runs. To find the best solution, however, is a very difficult optimization problem, which depends on many details of the model under study!

Now the question arises how to combine the data from different runs most efficiently. A very clear way is to compute for each simulation (at β_i) the β -dependence of $O_i \equiv O^{(\beta_i)}(\beta)$ plus the associated statistical error ΔO_i , using jack-knife techniques (Miller 1974, Efron 1982), say. A single optimal expression

for $O \equiv O(\beta)$ is then obtained by combining the values O_i in such a way that the relative error $\Delta O/O$ is minimized (Holm and Janke 1993b),

$$O = \left(\frac{O_1}{(\Delta O_1)^2} + \frac{O_2}{(\Delta O_2)^2} + \frac{O_3}{(\Delta O_3)^2} \right) (\Delta O)^2, \quad (48)$$

with

$$\frac{1}{(\Delta O)^2} = \frac{1}{(\Delta O_1)^2} + \frac{1}{(\Delta O_2)^2} + \frac{1}{(\Delta O_3)^2}. \quad (49)$$

A different procedure at the level of distribution functions was discussed by Ferrenberg and Swendsen (1989b). For the specific heat the two methods were found (Holm and Janke 1993b) to give comparable results within the statistical errors. The optimization at the level of observables, however, is simpler to apply to quantities involving constant energy averages such as $\langle\langle m \rangle\rangle(E)$, and, more importantly, minimizes the error on each observable of interest separately.

5 Applications to the 3D Heisenberg Model

Let us now turn to applications of the just described techniques to the 3D classical Heisenberg model ($n = 3$), focussing on an accurate determination of the transition temperature and the critical exponents. To this end the cluster update algorithm proved to be a very important tool. Previous studies (Peczak and Landau 1990a, 1990b, 1993, Peczak et al. 1991) employing the Metropolis algorithm reported for the magnetization an exponential autocorrelation time of $\tau_{m,\text{exp}} = aL^z$, with $a \approx 3.76$ and $z = 1.94(6)$. In simulations with the single-cluster algorithm we obtained for the susceptibility values of $\tau_{\chi,\text{int}} \approx 1.5 - 2.0$ (Holm and Janke 1993a, 1993b, 1994). As for the 3D Ising and 3D XY model critical slowing down is thus almost completely eliminated. Compared to the Metropolis algorithms this implies for a 80^3 lattice an acceleration of the simulation by about four orders of magnitude.

In the paper by Holm and Janke (1993b) two sets of MC simulations are reported. The first set of data consists of 18 simulations for $T > T_c$ in the range $\beta = 0.650$ to $0.686 \approx 0.99\beta_c$, with typically about 10^5 (almost uncorrelated) measurements. The correlation length varies in this β -range from $\xi \approx 3$ to $\xi \approx 12$ (see Fig. 11 below). Here the use of improved estimators was very useful and led to a further reduction of the statistical errors by a factor of about 2 – 3.

The second set of simulations was performed near criticality. For each lattice size typically three independent runs with more than 10^5 measurements each at different β values around β_c were combined using the optimized reweighting technique, which is the most important additional tool in the finite-size scaling region.

5.1 Simulations for $T \gtrsim T_c$

The conceptually easiest way to measure critical exponents are simulations in the high-temperature phase. In principle one simply has to fit the MC data for ξ , C , m , χ , ... with the expected power-law divergencies (10), (12)-(14) at criticality. For high-precision estimates, however, the procedure is far from being trivial. The problem is to locate the temperature range in which a simple power-law Ansatz like (10) is valid. Clearly, since the omitted correction terms are positive powers of $T/T_c - 1$, at first sight, one would like to perform the simulations as close to T_c as possible. But very close to T_c the correlation length gets very large, and on finite lattices one starts seeing finite-size corrections. The only way around these correction terms are large enough lattice sizes. In many models one finds empirically that the thermodynamic limit is approached when the linear lattice size satisfies $L \gtrsim (6 - 8)\xi$. But since the amplitude ξ_{0+} is non-universal, this estimate is not guaranteed to be always true. Therefore this question must be investigated very carefully for each model separately. With increasing temperature the correlation length decreases and finite-size corrections are no longer a problem, but then it is for a different reason again not clear if the simple power-law Ansatz is valid. Very far away from T_c the lattice structure becomes important and the observables show a completely different behaviour. In an intermediate range one sees the confluent and analytic correction terms which are very difficult to take into account in the fits. So in essence the problem is to locate a temperature window in which $1 \ll \xi \ll L$.

There are many ways to extract the correlation length ξ from the asymptotic decay of the spatial correlation function,

$$G(\mathbf{x}_i - \mathbf{x}_j) = \langle \boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_j \rangle \propto \exp(-|\mathbf{x}_i - \mathbf{x}_j|/\xi) . \quad (50)$$

One way is measuring the Fourier transform, $\tilde{G}(\mathbf{k}) = \sum_{\mathbf{x}} G(\mathbf{x}) \exp(-i\mathbf{k} \cdot \mathbf{x})$, for a few long-wavelength modes and performing least-square fits to

$$\tilde{G}(\mathbf{k})^{-1} = c \left[\sum_{i=1}^3 2(1 - \cos k_i) + (1/\xi)^2 \right] \approx c [\mathbf{k}^2 + (1/\xi)^2] , \quad (51)$$

where c is a constant and $k_i = (2\pi/L)n_i$, $n_i = 1, \dots, L$. Recall that for zero momentum the susceptibility is recovered, $\tilde{G}(\mathbf{k} = \mathbf{0}) = \chi'/\beta$. As an example Fig. 10 shows the data for the 3D Heisenberg model [$\mathbf{n} = (0,0,0)$, $(1,0,0)$, $(1,1,0)$, $(1,1,1)$, $(2,0,0)$, and $(2,1,0)$] at $\beta = 0.686 \approx 0.99\beta_c$. By repeating this analysis for different temperatures one obtains the data shown in Fig. 11. The solid lines are fits according to

$$\xi(T) = \xi_{0+}(T/T_c - 1)^{-\nu} , \quad (52)$$

with

$$\beta_c = 0.69281 \pm 0.00004 , \quad (53)$$

$$\nu = 0.698 \pm 0.002 , \quad (54)$$

$$\xi_{0+} = 0.484 \pm 0.002 , \quad (55)$$

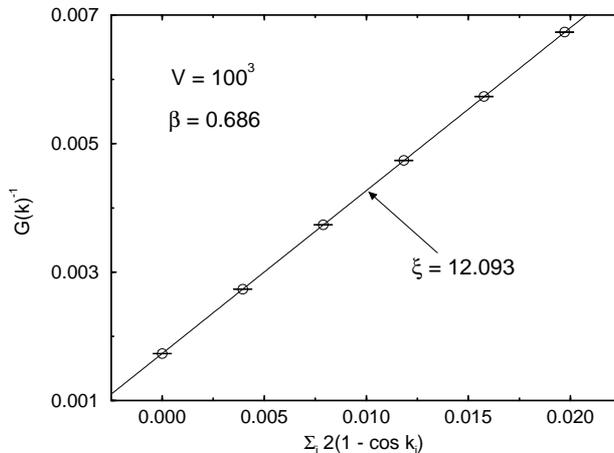


Fig. 10. Fit to the inverse of the Fourier transformed correlation function to compute the correlation length ξ .

and a goodness-of-fit parameter $Q = 0.92$, and

$$\chi'(\beta)/\beta = \bar{\chi}'_0(1 - \beta/\beta_c)^{-\gamma} , \quad (56)$$

with

$$\beta_c = 0.69294 \pm 0.00003 , \quad (57)$$

$$\gamma = 1.391 \pm 0.003 , \quad (58)$$

$$\bar{\chi}'_0 = 0.955 \pm 0.006 , \quad (59)$$

and $Q = 0.93$. Notice that the correlation length in (52) is written as a function of T and the susceptibility in (56) as a function of β . The high quality of the fits a posteriori justifies this choice and indicates that in the chosen temperature range confluent as well as analytic correction terms are negligible. If we rewrite $T/T_c - 1 = 1 - \beta/\beta_c + (1 - \beta/\beta_c)^2 + \dots$ and consider $\xi(\beta)$ instead of $\xi(T)$, we expect (and indeed confirmed) an analytical correction to asymptotic scaling. There is, however, so far no theoretical understanding why for the particular choice of arguments the analytical correction terms should vanish.

5.2 Simulations near T_c

The second set of data consists of simulations near T_c on lattices of size up to 48^3 (Holm and Janke 1993a, 1993b). In a later study focussing on topological defects the maximal size could even be increased to 80^3 (Holm and Janke 1994). In the vicinity of T_c , finite-size corrections are dominant and one has to employ

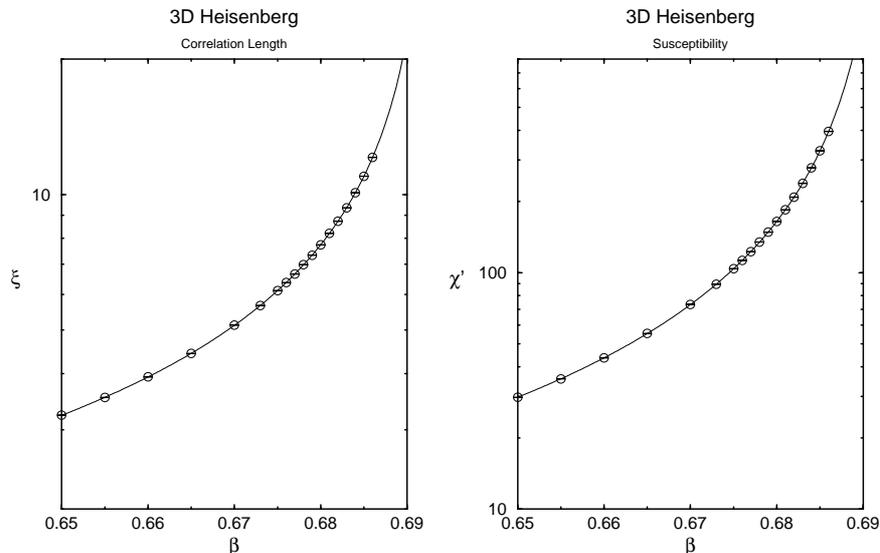


Fig. 11. Scaling behaviour of the correlation length and susceptibility. The solid lines are fits to $\xi(T)$ and $\chi'(\beta)$ according to the asymptotic power laws.

finite-size scaling (FSS) concepts to analyze the data. Usually one starts with an analysis of ratios of magnetization moments (Binder 1981), e.g.,

$$U_L(\beta) = 1 - \frac{1}{3} \frac{\langle m^4 \rangle}{\langle m^2 \rangle^2}, \quad (60)$$

which leads to estimates of β_c and the critical exponent ν . In the spontaneously broken low-temperature phase the magnetization distribution develops a double-peak structure with peaks at $\pm m_0 \neq 0$. Since the width of the peaks decreases with increasing lattice size one expects that U_L approaches $\frac{2}{3}$ for all $T < T_c$. In the disordered high-temperature phase the magnetization vanishes and the moments are determined by fluctuations alone. In the infinite volume limit the fluctuations become Gaussian and a simple calculation yields $U_L \rightarrow 2(n-1)/3n = 4/9$ for $n = 3$. Only at the transition point one expects a non-trivial limiting value which has been estimated by field theoretical methods (Brézin and Zinn-Justin 1985) to be $U^* = 0.59684\dots$ for $n = 3$. For finite systems, FSS predicts that the curves $U_L(\beta)$ for different L intersect around (β_c, U^*) with slopes $\propto L^{1/\nu}$, apart from confluent corrections explaining small systematic deviations. This allows an almost unbiased estimate of β_c , U^* and the critical exponent of the correlation length ν .

The data for the 3D Heisenberg model in Fig. 12 clearly confirm the theoretical expectations with a pronounced intersection point at $(\beta_c, U^*) = (0.6930(1), 0.6217(8))$. The final numbers are actually obtained from a little more elaborate

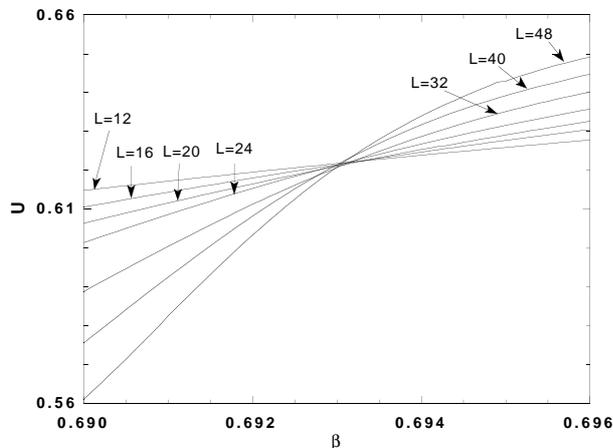


Fig. 12. The Binder parameter of the 3D Heisenberg model for various lattice sizes. The intersection points determine the inverse critical temperature $\beta_c = 0.6930(1)$.

analysis taking into account also the confluent corrections to asymptotic FSS (Holm and Janke 1993a, 1993b).

Also the slopes $dU_L/d\beta$ at $\beta = 0.6930 \approx \beta_c$ in Fig. 13 show the expected behaviour and a fit to the FSS prediction $dU_L/d\beta \propto L^{1/\nu}$ yields $\nu = 0.704(6)$, consistent with (54) and in very good agreement with estimates obtained by field theoretical methods or from series expansions, cp. Table 2.

In the analysis of the magnetization and susceptibility one proceeds similarly. The fit to the magnetization data reweighted to $\beta = 0.6930 \approx \beta_c$ shown in Fig. 14 yields $\beta/\nu = 0.514(1)$, and from the FSS of the susceptibility one reads off $\gamma/\nu = 1.9729(17)$. By multiplying the exponent ratios with the estimate of $\nu = 0.704(6)$, we finally arrive at the values for the critical exponents β and γ given in Table 2.

The analysis of the specific heat is much more complicated since the critical exponent α is usually quite small and therefore the singularity in C not very pronounced. For the 3D Heisenberg model α is actually negative, so that we do not expect a divergence at all. By using the additional data on lattices up to 80^3 (Holm and Janke 1994), we obtained from the fit $C = C_{\text{reg}} + C_0 L^{\alpha/\nu}$ shown in Fig. 15 an estimate of $\alpha/\nu = -0.225(80)$, resulting in $\alpha = -0.158(59)$. Due to the rather large statistical error this estimate is still consistent with the value obtained from hyperscaling, $\alpha = 2 - 3\nu = -0.112(18)$. Actually a much more precise result was obtained from the corresponding FSS fit to the energy at β_c . Using the Ansatz $e = e_{\text{reg}} + e_0 L^{(\alpha-1)/\nu}$, we obtained $\alpha/\nu = -0.166(31)$, translating into $\alpha = -0.117(23)$. Obviously this value is in a much better agreement with the hyperscaling prediction.

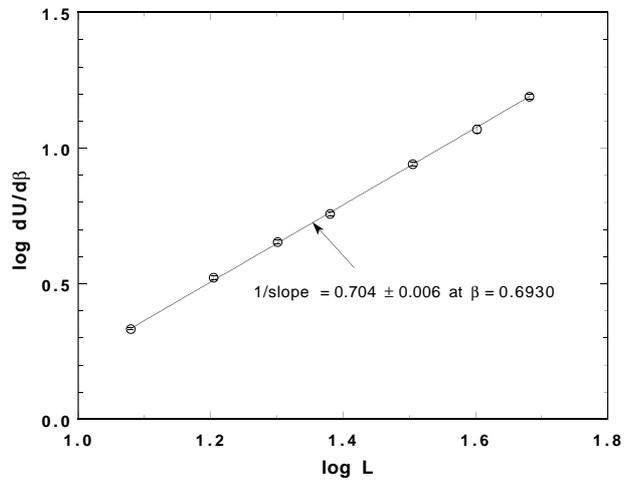


Fig. 13. FSS of the Binder parameter slopes at $\beta = 0.6930 \approx \beta_c$. The linear fit yields an estimate of the correlation length exponent, $\nu = 0.704(6)$.

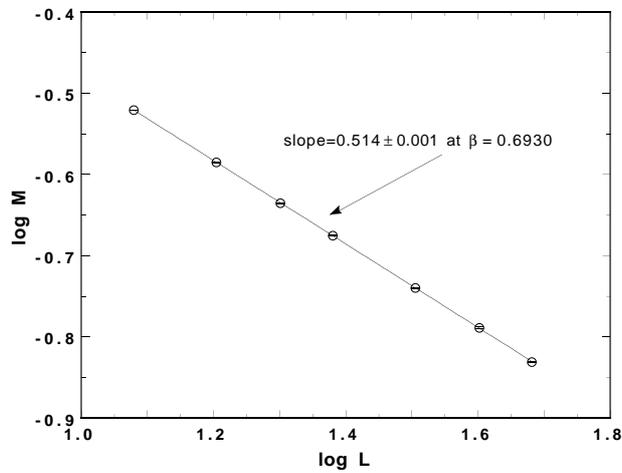


Fig. 14. FSS of the magnetization at $\beta = 0.6930 \approx \beta_c$. The linear fit yields an estimate of the exponent ratio, $\beta/\nu = 0.514(1)$.

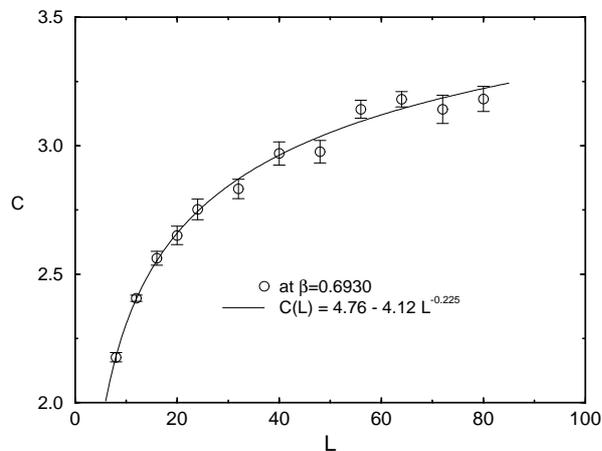


Fig. 15. FSS of the specific heat at $\beta = 0.6930 \approx \beta_c$. The non-linear fit taking into account a regular background term yields an estimate of the exponent ratio, $\alpha/\nu = -0.225(80)$.

Table 2. Critical coupling and critical exponents of the 3D classical Heisenberg ($n = 3$) model

method	β_c	ν	γ	β	α	δ
g -expansion ^a -		0.705(3)	1.386(4)	0.3645(25)	-0.115(9)	4.802(37)
ϵ -expansion ^b -		0.710(7)	1.390(10)	0.368(4)	-0.130(21)	4.777(70)
MC ^c	0.6929(1)	0.706(9)	1.390(23)	0.364(7)	-0.118(27)	4.819(36)
MC ^d	0.6930(1)	0.704(6)	1.389(14)	0.362(4)	-0.112(18)	4.837(11)
MC ^e	0.693035(37)	0.7036(23)	1.3896(70)	0.3616(31)	-0.1108(69)	-
series ^f	0.6929(1)	0.712(10)	1.400(10)	0.363(10)	-0.136(30)	4.86(10)

^a Le Guillou and Zinn-Justin 1977, 1980

^b Le Guillou and Zinn-Justin 1985

^c Peczak et al. 1991

^d Holm and Janke 1993a, 1993b

^e Chen et al. 1993

^f Adler et al. 1993

6 Concluding Remarks

The intention of these lecture notes was to give an elementary introduction to the basic concepts of modern Monte Carlo simulations and to illustrate their usefulness by applications to one typical model. Since the choice of the 3D Heisenberg model was obviously biased by my own work in this field, I want to conclude with at least a few remarks on the 3D Ising and 3D XY model, for which also quite a few high-precision simulations have been performed.

Due to its relative simplicity, the 3D Ising model is the best studied model of all $O(n)$ spin systems. Apart from finite-size scaling analysis as described here, also many other techniques have been applied, including the Monte Carlo renormalization group (MCRG) and the finite-size scaling of partition function zeros. This has led to quite a few very accurate estimates of critical exponents. Some of them are compiled in Table 3, where for comparison also field theory and series expansion estimates are given. As an amusing side remark it is worth mentioning the Rosengren (1986) conjecture that the critical coupling of the 3D Ising model is given by $\beta_c = \tanh^{-1}[(\sqrt{5} - 2) \cos(\pi/8)] = 0.221\,658\,637\dots$ – a value which is indeed in impressive agreement with the most precise Monte Carlo estimates!

Table 3. Critical coupling and selected critical exponents of the 3D Ising ($n = 1$) model

method	β_c	ν	γ
g -expansion ^a -		0.6300(15)	1.241(2)
ϵ -expansion ^b -		0.6310(15)	1.2390(25)
MCRG ^c	0.221 652(3)	0.624(2)	-
MC ^d	0.221 6595(26)	0.6289(8)	1.239(7)
series ^e	0.221 655(5)	0.631(4)	1.239(3)

^a Le Guillou and Zinn-Justin 1977, 1980

^b Le Guillou and Zinn-Justin 1985

^c Baillie et al. 1992

^d Ferrenberg and Landau 1991

^e Adler 1983

A similar accuracy could also be reached for the 3D XY model, the simplest model of the $O(2)$ universality class which governs the critical behaviour of the λ -transition in liquid helium. Some recent results of Monte Carlo simulations employing the Swendsen-Wang and Wolff cluster update algorithm, as well as estimates using series expansions and field theory methods are compiled in Table 4.

Table 4. Critical coupling and selected critical exponents of the 3D classical XY ($n = 2$) model

method	β_c	ν	γ
^4He experiment ^a	-	0.6705(6)	-
g -expansion ^b	-	0.669(2)	1.3160(25)
ϵ -expansion ^c	-	0.671(5)	1.315(7)
MC ^d	0.45421(8)	-	1.327(8)
MC ^e	0.45408(8)	0.670(2)	1.316(5)
MC ^f	0.45417(1)	0.662(7)	1.324(1)
series ^g	0.45406(5)	0.67(1)	1.315(9)
series ^h	0.45414(7)	≈ 0.673	≈ 1.325

^a Goldner and Ahlers 1992

^b Le Guillou and Zinn-Justin 1977, 1980

^c Le Guillou and Zinn-Justin 1985

^d Hasenbusch and Meyer 1990

^e Janke 1990

^f Gottlob and Hasenbusch 1993

^g Butera et al. 1993

^h Adler et al. 1993

By comparing the various Monte Carlo estimates collected in Tables 2-4 with results from field theory and series expansions it is fair to conclude that for $O(n)$ spin models modern Monte Carlo techniques are at the present time superior to series expansion analyses. The recently derived critical exponents are in fact competitive in accuracy with estimates obtained with the best and very elaborate methods of field theory.

Acknowledgements

The results for the 3D Heisenberg model are taken from joint work with C. Holm. It is a pleasure to thank him for a very fruitful and enjoyable collaboration. I am also grateful to S. Kappler and T. Sauer for numerous stimulating discussions and collaborations on various aspects of Monte Carlo simulations described here. Along the way, discussions with B. Berg, D. Johnston, and R. Villanova and their comments helped a lot. Last but not least I would like to thank K. Binder and D.P. Landau for sharing their experience and insight in many of the problems touched in these lecture notes. Finally, I gratefully acknowledge a Heisenberg fellowship by the Deutsche Forschungsgemeinschaft.

Appendix: Program Codes

The accompanying diskette contains five FORTRAN codes (*is_clu.f*, *reweight.f*, *rew_his.f*, *1dis_ex.f*, *2disnm_ex.f*) and two data files (*3d_e004.plo*, *3d_c004.plo*) which can be used to reproduce Figs. 3, 7, and 8, and to generate an Ising model analog of Fig. 9. The output files denoted by *...plo* are kept as simple as possible to allow easy plotting, e.g. with the standard utility *gnuplot*.

is_clu.f is a Monte Carlo simulation program for the nearest-neighbor Ising model with subroutines for the Wolff single-cluster (*sc*), Swendsen-Wang multiple-cluster (*sw*), Metropolis (*me*), and heat-bath (*hb*) update algorithm, which can be selected in the main parameter statement. All subroutines are set up to work for general D-dimensional (hyper-) cubic lattices of size L^D with periodic boundary conditions. The dimension, lattice size, simulation temperature, and the statistics parameters are also defined in the main parameter statement (the dimension and lattice size have to be changed globally in all subroutines as well). With the choice of parameters as given in *is_clu.param.example*, the 2D Ising (multiple-cluster) Monte Carlo data of Figs. 7 and 8 can be reproduced within about one minute run-time. Of course the detailed timing depends on the type of PC or workstation, but it should always be possible to run the simulation interactively. The average energy, magnetization, specific heat, susceptibility, Binder parameter, and cluster averages are written on standard output. Furthermore, the energy and magnetization histograms, and the “microcanonical” magnetization averages $\langle\langle|m|\rangle\rangle$ and $\langle\langle m^2\rangle\rangle$ (cp. (46)) at the simulation temperature are saved in the files *ehis_b0.plo*, *mhis_b0.plo*, *malis_b0.plo*, and *m2lis_b0.plo* for easy plotting, and in *e_b0.his* for further reweighting analyses (containing the necessary parameter informations). With these output files it is straightforward to produce for the Ising model a plot similar to Fig. 9. If desired the time evolution of the energy and magnetization measurements can be saved in the files *e_series.plo* and *m_series.plo*, respectively, by turning on the corresponding logical switches in the parameter statement.

For simplicity the standard UNIX random number generator `RAND()` is called in the MC program. For illustration purposes this generator is good enough, but for a serious simulation study it should at any rate be replaced by a more reliable random number generator. Again only for simplicity all statistical error analysis subroutines are omitted. For a sensible MC study this is clearly unacceptable.

reweight.f takes as input the energy histogram and the “microcanonical” magnetization lists stored in *e_b0.his* and computes by reweighting the energy and specific heat, the susceptibilities χ/β and χ'/β , and the Binder parameter as a function of inverse temperature β . The desired β -range can be set in the parameter statement, and the dimension and lattice size parameters must be the same as in *is_clu.f*. The results are written into the files *e016_mc.plo*, *c016_mc.plo*,

sus016.mc.plo, *chi016.mc.plo*, and *U016.mc.plo*, respectively, where *016* indicates the linear lattice size.

The MC data for the one-dimensional Ising chain can be tested against the exact results provided by *1dis_ex.f* for any chain length. The two-dimensional MC results can be used together with the output from *2disnm_ex.f* to reproduce Fig. 7. Further comparison data for a 16^2 lattice from high-statistics single-cluster simulations at $\beta_c = \ln(1+\sqrt{2})/2$ are $2+e = 0.546\,85(10)$, $C = 1.4978(10)$, $\chi = 139.669(31)$, $\langle|C|\rangle = 139.656(29)$, and $U = 0.611\,537(50)$. In three dimensions the MC data can be compared with the exact energy and specific heat curves for a 4^3 lattice contained in the data files *3d_e004.plo* and *3d_c004.plo*.

rew_his.f reads again as input the energy histogram stored in *e_b0.his*, computes reweighted histograms, and stores them in, e.g., *ehis_b4750.plo*. The dimension and lattice size parameters must be the same as in *is.clu.f*. Here *b4750* indicates that the histogram is reweighted to $\beta = 0.4750$. The new inverse temperature is inquired interactively by the program. In this way Fig. 8 can be reproduced. If *gnuplot* is used for plotting the histograms, then by using the escape character "!" the reweighting program can be called and the new histogram immediately displayed without leaving the plot session.

1dis_ex.f computes the exact temperature dependence of the energy, specific heat, and susceptibility of the one-dimensional Ising chain with periodic boundary conditions of arbitrary length L . For e.g. $L = 16$, the results are stored in the files *1d_e016.plo*, *1d_c016.plo*, and *1d_chi016.plo*.

2disnm_ex.f implements the exact solution of Ferdinand and Fisher (1969) for the 2D nearest-neighbor interaction Ising model on finite $L_x \times L_y$ lattices ($L_x, L_y = \text{even}$) with periodic boundary conditions. The desired lattice size and inverse temperature range can be chosen in the parameter statement of the main program. The output are two data files, e.g. *2d_e016.plo* and *2d_c016.plo* for $L_x = L_y = 16$, containing minus the internal energy per site, $-E/V$, and the specific heat per site, C , as a function of the inverse temperature β . By running this code for various lattice sizes, Fig. 3 can be reproduced.

References

- Adler, J. (1983): J. Phys. **A16**, 3585
 Adler, J., Holm, C., Janke, W. (1993): Physica **A201**, 581
 Adler, S.L. (1988): Phys. Rev. **D37**, 458
 Alves, N.A., Berg, B.A., Sanielevici, S. (1992): Nucl. Phys. **B376**, 218
 Baillie, C.F. (1990): Int. J. Mod. Phys. **C1**, 91
 Baillie, C.F., Gupta, R., Hawick, K.A., Pawley, G.S. (1992): Phys. Rev. **B45**, 10438
 Barber, M.E. (1983): in *Phase Transitions and Critical Phenomena*, Vol. 8, edited by C. Domb and J.L. Lebowitz (Academic, New York)
 Berg, B.A., Neuhaus, T. (1991): Phys. Lett. **B267**, 249
 Berg, B.A., Neuhaus, T. (1992): Phys. Rev. Lett. **68**, 9
 Berezinskii, V.L. (1971): Zh. Eksp. Teor. Fiz. **61**, 1144 [Sov. Phys.–JETP **34**, 610 (1972)]
 Berlin, T.H., Kac, M. (1952): Phys. Rev. **86**, 821
 Billoire, A., Lacaze, R., Morel, A., Gupta, S., Irbäck, A., Petersson, B. (1991): Nucl. Phys. **B358**, 231
 Binder, K., (1979): in *Monte Carlo Methods in Statistical Physics*, edited by K. Binder (Springer, Berlin)
 Binder, K., (1981): Z. Phys. **B43**, 119
 Binder, K., (1982): Phys. Rev. **A25**, 1699
 Binder, K., (1987): Rep. Prog. Phys. **50**, 783
 Binder, K. (ed.), (1992a): *The Monte Carlo Method in Condensed Matter Physics* (Springer, Berlin)
 Binder, K. (1992b): in *Computational Methods in Field Theory*, p. 59, Schladming 1992 lecture notes, edited by H. Gausterer and C.B. Lang (Springer, Berlin)
 Binder, K., Heermann, D.W. (1988): *Monte Carlo Simulations in Statistical Physics: An Introduction* (Springer, Berlin)
 Binder, K., Landau, D.P. (1984): Phys. Rev. **B30**, 1477
 Borgs, C., Janke, W. (1992): Phys. Rev. Lett. **68**, 1738
 Borgs, C., Kotecky, R. (1990): J. Stat. Phys. **61**, 79
 Borgs, C., Kotecky, R. (1992): Phys. Rev. Lett. **68**, 1734.
 Borgs, C., Kotecky, R., Miracle-Solé, S. (1991): J. Stat. Phys. **62**, 529.
 Brézin, E., Zinn-Justin, J. (1985): Nucl. Phys. **B257**[FS14], 867
 Butera, P., Comi, M., Guttmann, A.J. (1993): Phys. Rev. **B48**, 13987
 Challa, M.S.S., Landau, D.P., Binder, K. (1986): Phys. Rev. **B34**, 1841
 Chen, K., Ferrenberg, A.M., Landau, D.P. (1993): Phys. Rev. **B48**, 3249
 Creutz, M. (1987): Phys. Rev. **D36**, 515
 Dimitrović, I., Hasenfratz, P., Nager, J., Niedermayer, F. (1991): Nucl. Phys. **B350**, 893
 Doll, J.D., Coalsen, R.D., Freeman, D.L. (1985): Phys. Rev. Lett. **55**, 1
 Efron, B. (1982): *The Jackknife, the Bootstrap and other Resampling Plans* (SIAM, Philadelphia)
 Ferdinand, A.E., Fisher, M.E. (1969): Phys. Rev. **185**, 832
 Ferrenberg, A.M., Landau, D.P. (1991): Phys. Rev. **B44**, 5081
 Ferrenberg, A.M., Swendsen, A.M. (1988): Phys. Rev. Lett. **61**, 2635
 Ferrenberg, A.M., Swendsen, A.M. (1989a): Phys. Rev. Lett. **63**, 1658(E)
 Ferrenberg, A.M., Swendsen, A.M. (1989b): Phys. Rev. Lett. **63**, 1195
 Ferrenberg, A.M., Landau, D.P., Binder, K. (1991): J. Stat. Phys. **63**, 867
 Fortuin, C.M., Kasteleyn, P.W. (1972): Physica **57**, 536

- Fisher, M.E., Berker, A.N. (1982): Phys. Rev. **B26**, 2507
- Fortuin, C.M. (1972a): Physica **58**, 393
- Fortuin, C.M. (1972b): Physica **59**, 545
- Goldner, L.S., Ahlers, G. (1992): Phys. Rev. **B45**, 13129
- Goodman, J., Sokal, A.D. (1986): Phys. Rev. Lett. **56**, 1015
- Goodman, J., Sokal, A.D. (1989): Phys. Rev. **D40**, 2035
- Gottlob, A.P., Hasenbusch, M. (1993): Physica **A201**, 593
- Gunton, J.D., Miguel, M.S., Sahni, P.S. (1983): in *Phase Transitions and Critical Phenomena*, Vol. 8, eds. C. Domb and J.L. Lebowitz (Academic Press, New York)
- Gupta, R., DeLapp, J., Battrouni, G.G., Fox, G.C., Baillie, C.F., Apostolakis, J. (1988): Phys. Rev. Lett. **61**, 1996
- Hammersley, J.M., Handscomb, D.C. (1965): *Monte Carlo Methods* (London)
- Hasenbusch, M. (1990): Nucl. Phys. **B333**, 581
- Hasenbusch, M., Meyer, S. (1990): Phys. Lett. **B241**, 238
- Heermann, D.W. (1990): *Computer Simulation Methods in Theoretical Physics*, 2nd ed., (Springer, Berlin)
- Heermann, D.W., Burkitt, A.N. (1990): Physica **A162**, 210
- Herrmann, H.J., Janke, W., Karsch, F. (eds.) (1992): *Dynamics of First Order Phase Transitions* (World Scientific, Singapore)
- Holm, C., Janke, W. (1993a): Phys. Lett. **A173**, 8
- Holm, C., Janke, W. (1993b): Phys. Rev. **B48**, 936
- Holm, C., Janke, W. (1994): J. Phys. **A27**, 2553
- Ito, N., Kohring, G.A. (1993): Physica **A201**, 547
- Janke, W. (1990): Phys. Lett. **A148**, 306
- Janke, W. (1993): Phys. Rev. **B47**, 14757
- Janke, W. (1994): *Recent Developments in Monte Carlo Simulations of First-Order Phase Transitions*, in *Computer Simulations in Condensed Matter Physics VII*, p. 29, eds. D.P. Landau, K.K. Mon, and H.-B. Schüttler (Springer, Berlin)
- Janke, W., Kappler, S. (1994): Nucl. Phys. **B** (Proc. Suppl.) **34**, 674
- Janke, W., Kappler, S. (1995a): Phys. Lett. **A197**, 227
- Janke, W., Kappler, S. (1995b): Nucl. Phys. **B** (Proc. Suppl.) **42**, 770
- Janke, W., Kappler, S. (1995c): Europhys. Lett. **31**, 345
- Janke, W., Kappler, S. (1995d): Mainz preprint KOMA-95-65 (September 1995), hep-lat/9509057, to appear in the proceedings of the conference *Lattice '95*, Melbourne, Australia, July 11 – 15, 1995
- Janke, W., Kappler, S. (1995e): Nucl. Phys. **B** (Proc. Suppl.) **42**, 876
- Janke, W., Kappler, S. (1995f): Phys. Rev. Lett. **74**, 212
- Janke, W., Kleinert, H. (1986): Phys. Rev. **B33**, 6346
- Janke, W., Sauer, T. (1994): Phys. Rev. **E49**, 3475
- Janke, W., Sauer, T. (1995): J. Stat. Phys. **78**, 759
- Janke, W., Berg, B.A., Katoot, M. (1992): Nucl. Phys. **B382**, 649
- Kandel, D., Domany, E., Brandt, A. (1989): Phys. Rev. **B40**, 330
- Kandel, D., Domany, E., Ron, D., Brandt, A., Loh, Jr., E. (1988): Phys. Rev. Lett. **60**, 1591
- Kappler, S. (1995): Ph.D. Thesis, Johannes Gutenberg-Universität Mainz
- Kasteleyn, P.W., Fortuin, C.M. (1969): J. Phys. Soc. Japan **26** (Suppl.), 11
- Kleinert, H. (1989a): *Gauge Fields in Condensed Matter*, Vol. I (World Scientific, Singapore)

- Kleinert, H. (1989b): *Gauge Fields in Condensed Matter*, Vol. II (World Scientific, Singapore)
- Kosterlitz, J.M. (1974): J. Phys. **C7**, 1046
- Kosterlitz, J.M., Thouless, D.J. (1973): J. Phys. **C6**, 1181
- Lee, J., Kosterlitz, J.M. (1990): Phys. Rev. Lett. **65**, 137
- Lee, J., Kosterlitz, J.M. (1991): Phys. Rev. **B43**, 3265
- Le Guillou, J.C., Zinn-Justin, J. (1977): Phys. Rev. Lett. **39**, 95
- Le Guillou, J.C., Zinn-Justin, J. (1980): Phys. Rev. **B21**, 3976
- Le Guillou, J.C., Zinn-Justin, J. (1985): J. Physique Lett. **46**, L137
- Mack, G. (1988): in *Nonperturbative quantum field theory*, Cargèse lectures 1987, edited by G.'t Hooft et al. (Plenum, New York)
- Mack, G., Meyer, S. (1990): Nucl. Phys. **B** (Proc. Suppl.) **17**, 293
- Madras, N., Sokal, A.D. (1988): J. Stat. Phys. **50**, 109
- Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.W., Teller, E. (1953): J. Chem. Phys. **21**, 1087
- Miller, R.G. (1974): Biometrika **61**, 1
- Neuberger, H. (1988): Phys. Lett. **B207**, 461
- Patrascioiu, A., Seiler, E. (1995): Munich preprint MPI-PhT/95-73, hep-lat/9508014
- Peczak, P., Landau, D.P. (1990a): Bull. Am. Phys. Soc. **35**, 255
- Peczak, P., Landau, D.P. (1990b): J. Appl. Phys. **67**, 5427
- Peczak, P., Landau, D.P. (1993): Phys. Rev. **B47**, 14260
- Peczak, P., Ferrenberg, A.M., Landau, D.P. (1991): Phys. Rev. **B43**, 6087
- Potts, R.B. (1952): Proc. Camb. Phil. Soc. **48**, 106
- Privman, V. (editor) (1990): *Finite-Size Scaling and Numerical Simulations of Statistical Systems* (World Scientific, Singapore)
- Privman, V., Fisher, M.E. (1983): J. Stat. Phys. **33**, 385
- Privman, V., Rudnik, J. (1990): J. Stat. Phys. **60**, 551
- Rosengren, A. (1986): J. Phys. **A19**, 1709
- Rummukainen, K. (1993): Nucl. Phys. **B390**, 621
- Sauer, T. (1994): Ph.D. Thesis, Freie Universität Berlin
- Sokal, A.D. (1989): *Monte Carlo Methods in Statistical Mechanics: Foundations and New Algorithms*, Cours de Troisième Cycle de la Physique en Suisse Romande, Lausanne
- Sokal, A.D. (1992): *Bosonic Algorithms*, in *Quantum Fields on the Computer*, p. 211, edited by M. Creutz (World Scientific, Singapore)
- Swendsen, R.H., Wang, J.-S. (1987): Phys. Rev. Lett. **58**, 86
- Swendsen, R.H., Wang, J.-S., Ferrenberg, A.M. (1992): in *The Monte Carlo Method in Condensed Matter Physics*, edited by K. Binder (Springer, Berlin)
- Tamayo, P. (1993): Physica **A201**, 543
- Wolff, U. (1989a): Phys. Rev. Lett. **62**, 361
- Wolff, U. (1989b): Nucl. Phys. **B322**, 759
- Wolff, U. (1989c): Phys. Lett. **A228**, 379
- Wolff, U. (1990): Nucl. Phys. **B334**, 581
- Wu, F.Y. (1982): Rev. Mod. Phys. **54**, 235
- Wu, F.Y. (1983): Rev. Mod. Phys. **55**, 315(E)

2D Ising model exact solution, 7, 19

Binder parameter, 25

critical exponent, 5
critical slowing down, 11

energy distribution, 18
exponential autocorrelation time, 9
exponential slowing down, 17

finite-size scaling Ansätze, 7
first-order phase transition, 4, 17

importance sampling, 8
improved cluster estimator, 15
integrated autocorrelation time, 9

Markov chain, 8
microcanonical averages, 20
multiple-cluster update, 13

non-local update algorithms, 11

$O(n)$ spin models, 2

second-order phase transition, 5
single-cluster algorithm, 13
statistical error, 10

universality, 5