Journal of Statistical Physics, Vol. 82, Nos. 1/2, 1996

Multicanonical Recursions

Bernd A. Berg[1,2]

Received March 21, 1995; final June 14, 1995

The problem of calculating multicanonical parameters recursively is discussed. I describe in detail a computational implementation which has worked reasonably well in practice.

KEY WORDS: Monte Carlo sampling; multicanonical ensemble; algorithm; recursion; spectral density.

1. INTRODUCTION

Recently multicanonical Monte Carlo (MC) sampling and closely related methods have received much attention.[1–21] Considerable gains have been achieved in situations with "supercritical" slowing down, such as first-order transition[1,8,21] (for a recent review see ref. 12) and systems with conflicting constrains, for instance, spin glasses[2,4,17,18] or proteins.[15,16] In the multicanonical ensemble[1,5] one samples configurations such that exact reconstruction of canonical expectation values becomes feasible for a desired temperature range. This requires a broad energy distribution, and leaves innovative freedom concerning the optimal shape.[20] Considerable practical experience exists only for the uniform energy distribution, where one samples such that:

(a) The energy density is flat in a desired range,

$$P(E) = \text{const} \qquad \text{for} \quad E_{\min} \leqslant E \leqslant E_{\max} \tag{1}$$

(b) Each configuration of fixed energy E appears with the same likelihood.

[1] Department of Physics, Florida State University, Tallahassee, Florida FL 32306, E-mail: berg@hep.fsu.edu.

[2] Supercomputer Computations Research Institute, Tallahassee, Florida 32306.

822/82/1-2-21

It should be noted that condition (b) is nontrivial. A simple algorithm[22] exists to achieve (a), but which gives up (b). Exact connection to the canonical ensemble is then lost. Such algorithms are interesting for hard optimization problems, but unsuitable for canonical statistical physics. The present paper focuses on achieving (a) and (b).

A multicanonical investigation splits into two clearly distinct parts:

(i) Determination of the weight factors.

(ii) (Subsequent) Equilibrium simulations with fixed weight factors.

Whereas the computational gains and physical achievements of the equilibrium simulations (ii) are normally well documented, papers tend to be fairly sketchy about part (i). The reason is that in a multicanonical simulation one has to calculate the weight factors only once per system size. However, performing something once inhibits statistical conclusions concerning the efficiency of the employed method (if it was systematic at all). This seems to be the main reason it has been impossible to compare actually used recursions quantitatively. Unfortunately, this renders the method difficult, if not obscure, to the newcomer. To achieve a flat energy distribution (10), the appropriate unnormalized weight factor $w(E)$ is the inverse spectral density $w(E) = n^{-1}(E)$. Now (i) deals with the fact that the spectral density is *a priori* unknown. Otherwise we would have solved the problem in the first place. Presumably, reluctance about simulations with an *a priori* unknown weight factor is the main reason why earlier umbrella sampling[23] never became popular in statistical physics. It is the purpose of this paper to provide a first systematic and quantitative documentation for part (i) of multicanonical simulations.

One has to distinguish two rather different situations. For first-order phase transitions the problem of the *a priori* unknown weight factor is rather elegantly overcome by means of finite-size scaling (FSS) methods.[1, 8, 9, 11, 12] A sufficiently accurate estimate is obtained by extrapolation from the already simulated smaller lattices. The smallest lattices still allow for efficient canonical simulations.

For systems with conflicting constraints the situation is less satisfactory, because FSS does not work. For instance, for spin glasses one has to perform the additional average over quenched random variables (which are the exchange coupling constants). Different choices of these random variables define different realizations of the same system. For the Edward-Anderson Ising (EAI) spin glass it turned out[2, 17] that even for identical lattice sizes, different realizations need different weight factors. Each system requires a new estimate of the weights factors with no *a priori* information available. The present paper documents a yet unpublished recursive method that has worked satisfactorily for more than 1500 different realizations of the

3D EAI model.[17] No claim is made that the method is in any sense optimal. Actually the author is considering various improvements. It is supposed to provide a reasonable starting point.

The paper is organized as follows: In Section 2 generalized Ising models and related preliminaries are introduced. Mainly for pedagogical reasons, I focus on them for examples of this paper. It is clear that generalization to other systems is straightforward, although possibility tedious for continuous systems. In Section 3 I introduce the multicanonical method, and review recursions given in the literature.[2, 13] No quantitative documentation of their performance is available. As their more efficient variants need considerable attention by "hand," it also seems impractical to rerun them sufficiently often now as needed to draw such conclusions. Section 4 describes the recursion which I invented for the simulations of ref. 17, and Section 5 illustrates its performance. Summary and conclusions follow. The appendix gives and explains a Fortran implementation.

2. GENERALIZED ISING MODELS

Let us consider a d-dimensional hypercubic lattice of volume $V = N = L^d$ with periodic boundary conditions. Spins $s_i = \pm 1$ are located at the N sites, and exchange interactions $J_{ik} = \pm 1$ at the dN links of the lattice. The energy of generalized Ising models is given by

$$E = -\sum_{\langle ik \rangle} J_{ik} s_i s_j \tag{2}$$

where the sum is over the nearest neighbor pairs. For $J_{ik} \equiv 1$ the standard Ising ferromagnet (IF) is recovered. When the J_{ik} are quenched random variables, one obtains the EAI spin glass. I confine the subsequent discussion to these two situations. Let us further restrict the EAI spin glass to the situation $\Sigma_{\langle ik \rangle} J_{ik} = 0$. The partition function may be written as

$$Z(\beta) = \sum_E n(E)\, e^{-\beta E} \tag{3}$$

where $n(E)$ is the spectral density (24), more precisely the number of configurations (or states) with energy E. As the system has 2^N different states, this implies the normalization

$$\sum_E n(E) = 2^N \tag{4}$$

The lowest possible energy is $-dN$, reached when each link contributes $J_{ik} s_i s_k = 1$. For the IF this is achieved with either all spins up $(+1)$ or all

spins down (-1). The possible energy increments under the flip of a single spin are

$$\Delta E = 0, \pm 4, ..., \pm 4d \tag{5}$$

Consequently $n(E)$ may take nonzero values for

$$E = -dN, -dN+4, ..., 0, ..., dN-4, dN \tag{6}$$

For instance, for $d>1$ the IF has $n(-dN)=2$, $n(-dN+4)=0,...$, and $n(-dN+4d)=N$. For a typical EAI spin-glass configuration the ground-state energy $E_{\min}$ is considerably larger than $-dN$.

3. MULTICANONICAL SAMPLING

In the pedagogical review of ref. 5 I emphasized that the inverse spectral density is the appropriate weight factor to obtain a flat energy density

$$w(E) = n^{-1}(E) = e^{-\beta(E)E + \alpha(E)} \tag{7}$$

Here $\beta(E)$, $\alpha(E)$ is the multicanonical parameterization.[1,2,26] The rationale of this notation is that $\beta(E)$ relates to the temperature, as will be shown. It should be noted that MC calculations are insensitive to an overall independent factor, i.e., against replacing $w(E)$ by $cw(E)$. In the following I will exploit this property from time to time, and not trace back the corresponding multiplicative or additive constants. If necessary, they may be obtained by introducing a convenient normalization. The spectral density may be written as

$$n(E) = e^{S(E)} \tag{8}$$

where $S(E)$ is the microcanonical entropy.[24] The thermodynamic relation for the inverse temperature ($\beta = T^{-1}$, where my Boltzmann constant convention is $k=1$) reads

$$\beta = \frac{\partial S}{\partial E} \tag{9}$$

For models with discrete energy values this may be translated into

$$\beta(E) = \frac{S(E+\varepsilon) - S(E)}{\varepsilon} \tag{10}$$

where ε is the smallest possible energy increment such that $n(E+\varepsilon)$ and $n(E)$ are both nonzero. Typically we have $\varepsilon = 4$ for the model of Section 2 (special care is needed for the IF close to its ground state). Note that

Eq. (10) is in part convention. Other valid options would be $\beta(E) = [S(E) - S(E-\varepsilon)]/\varepsilon$ or $\beta(E) = [S(E+\varepsilon) - S(E-\varepsilon)]/(2\varepsilon)$. For consistency with refs. 2 and 5 I stay with (10).

Once $\beta(E)$ is given, $\alpha(E)$ may be determined recursively. The equality of $e^{-S(E)}$ and $e^{-\beta(E)E+\alpha(E)}$ implies

$$S(E) - S(E-\varepsilon) = \beta(E)E - \beta(E-\varepsilon)(E-\varepsilon) - \alpha(E) + \alpha(E-\varepsilon)$$

Using (10) to eliminate the term $\varepsilon\beta(E-\varepsilon)$, we find for $\alpha(E)$ the recursion relation

$$\alpha(E-\varepsilon) = \alpha(E) + [\beta(E-\varepsilon) - \beta(E)]E, \qquad \alpha(E_{\max}) = 0 \tag{11}$$

Here $\alpha(E_{max}) = 0$ is a choice of the overall multiplicative constant needed to start off the recursion.

To perform a multicanonical simulation, we do not need to know the exact weight factor (7). Instead, a working estimate $\bar{w}(E)$ of $w(E)$ is sufficient, such that the sampled energy histogram $H(E)$ is approximately flat in the desired energy range[1]. In the subsequent discussion I use the notation $\bar{n}(E)$, $\bar{S}(E)$, $\bar{\beta}(E)$, and $\bar{\alpha}(E)$ for estimators of the corresponding quantities $n(E)$, $S(E)$, $\beta(E)$, and $\alpha(E)$. The technical feasibility of multicanonical sampling depends on the existence of efficient methods to obtain an acceptable estimate $\bar{w}(E)$. It seems that different workers in the field have tried various approaches. I am only familiar with two of them, as follows:

(a) Methods which work in one or two steps.[1, 8, 11] Employing FSS, a reasonably good approximation $\bar{w}^{(1)}(E)$ is obtained by extrapolation from previously simulated, smaller lattices. With $\bar{w}^{n}(E)$ a first multicanonical simulation is carried out. Its results give an improved estimate $\bar{w}^{(2)}(E)$ with which additional simulations may be done. This approach works well for first-order phase transitions, but failed badly for some disordered systems.

(b) Recursive calculations $\bar{w}^{n}(E) \to \bar{w}^{n+1}(E)$ have been employed. They are subject of the following subsection.

3.1. Recursive Multicanonical Calculations

Let $H^{n}(E)$ be the unnormalized histogram obtained from a (short) multicanonical simulation with $\bar{w}^{n}(E)$. At energy values for which $H^{n}(E)$ is *reliable*, the new estimate is

$$\bar{w}^{n+1}(E) = \frac{\bar{w}^{n}(E)}{H^{n}(E)} \tag{12}$$

Clearly Eq. (12) fails for energy values for which $H^n(E)=0$, and also values like $H^n(E)=1$ or 2 are of course statistically unreliable. Worse, situations can be encountered where even large values like $H^n(E)=10^6$ or more entries are all correlated. Let us assume that the starting point for the recursion is

$$\bar{w}^0(E) \equiv 1 \tag{13}$$

In general this is a reasonable choice, which will allow us to recover the normalization (4) when desired. For some applications other choices, like a canonical simulation at a certain temperature, may be more convenient.

In the Berg and Celik[2] Eq. (12) was stated in the multicanonical notation (7). It reads then (note $\varepsilon=4$ in ref. 2)

$$\bar{\beta}^{n+1}(E) = \bar{\beta}^n(E) + \varepsilon^{-1} \ln[H^n(E+\varepsilon)/H^n(E)] \tag{14a}$$

The function $\bar{\alpha}^{n+1}(E)$ is then determined by Eq. (11). In addition to (14a), specific rules were given about how to exclude unreliable histogram entries. Namely,

$$\bar{\beta}^{n+1}(E) = \begin{cases} \bar{\beta}^n(E) & \text{for} \quad E \geqslant E^n_{\text{median}} \\ \bar{\beta}^{n+1}(E^n_{\text{cutoff}}) & \text{for} \quad E < E^n_{\text{cutoff}} \end{cases} \tag{14b}$$

Here E^n_{median} is the median of the nth energy distribution, and $E^n_{\text{cutoff}} < E^n_{\text{median}}$ is an energy cutoff, such that in simulation n the temperature is kept constant for $E < E^n_{\text{cutoff}}$. Further, note that the starting condition (13) becomes $\bar{\beta}^0(E) \equiv 0$, $\bar{\alpha}^0(E) \equiv 0$.

Lee[13] states his recursion in two parts:

$$\bar{S}^{n+1}(E) = \bar{S}^n(E) + \ln H^n(E) \qquad \text{for} \quad H^n(E) \geqslant 1 \tag{15a}$$

and

$$\bar{S}^{n+1}(E) = \bar{S}^n(E) \qquad \text{for} \quad H^n(E) = 0 \tag{15b}$$

The first part is obviously Eq. (12), as follows from $\bar{w}^n(E) = \exp[-\bar{S}^n(E)]$. The identity[14] of (15a) and (14a) follows from (10). Obviously (15a) is an intermediate step to derive (14a). The second part (15b) is a specific prescription about how to handle $H^n(E)=0$. The other unreliable $H^n(E)$ are included into the recursion (12). Let us note the following:

(a) Aside from minor notational differences, both approaches handle the reliable part of the data identically. One should note that the equivalent equations (12), (14a), and (15a) are all nonlocal. The $\beta=0$ transition

probabilities $w[E \to E']$ relate different energies. They form a (sparse) matrix whose eigenvector with eigenvalue one is finally supposed to become the spectral density, i.e., determines the weight factors. This diagonalization (implicitly carried out by the MC simulation) is a nonlocal process, and may induce certain instabilities.

(b) As the recursions (12) and (15a) stand, the statistical accuracy of estimate $n+1$ is entirely determined by MC simulation n. With increasing n the covered energy range gets larger and larger. One needs longer and longer simulations just to regain the previously reached statistical accuracy (on the appropriate energy subrange). It is possible, but tedious, to combine the statistics of simulations $n, n-1, \ldots, 1, 0$.

(c) The median rule of (14b) freezes estimates on some part of the already covered energy range, but one should improve on it by using subsequent statistics when available. In ref. 2 it was suggested to combine the median rule with upper bounds on the energy, such that the energy range gets reasonably restricted. However, it is then difficult to ensure ergodicity.

(d) A central difficulty of the recursions is the handling of energy regions for which reliabe statistical information is not yet available. I elaborate on this now.

Lee's proposal (15b) looks attractive because of its simplicity. It works for the very small systems considered in his paper, but for many realistic situations it will lead to an unacceptable slowing down. The reason is that (15b) is equivalent to simulating with a constant weight factor (7). Now, at low temperatures one typically encounters

$$n(E-\varepsilon)/n(E) \sim V^{-1} \tag{16}$$

Therefore, for a not yet covered energy range $E \leqslant E_0$ one will need of order V attempts just to achieve once the transition $E_0 \to E_0 - \varepsilon$.

The rule $\bar{\beta}^{n+1} = \bar{\beta}^{n+1}(E_{\text{cutoff}})$ for $E < E_{\text{cutoff}}$ from (14b) achieves a far better performance for this situation. Assume that $\beta(E)$ is monotonically increasing toward lower energies (exceptions are first-order phase transitions). A canonical simulation with $\beta^{n+1}(E_{\text{cutoff}})$ will have its maximum energy density at $E = E_{\text{cutoff}}$, because its first derivative with respect to the energy is zero there. The width of its energy distribution is of order $\sqrt{V}$, and by this amount the recursion is supposed to proceed forward toward lower energies. In practice one has to use estimators $\bar{\beta}^{n+1}(E)$. One would like to choose E_{cutoff} as low as possible, but one encounters noise problem when the cutoff energy is shifted too far toward the edge of the reliably covered energy range. With some experience a good "pick" for E_{cutoff} can be achieved by just inspecting the function $\bar{\beta}^{n+1}(E)$. Alternatively, one

may use a fit $\bar{\beta}^{n+1}_{\max}$ from several energy values instead of $\bar{\beta}^{n+1}(E_{\text{cutoff}})$, or even fit the continuation of the entire function $\bar{\beta}^{n+1}(E)$ for $E < E_{\text{cutoff}}$ (with the penalty of spurious instabilities). In any case, in energy regions where (16) holds, one expects a performance increase by at least a volume factor over using (15b). On the other hand, it is precisely this part of the recursion (14) which requires annoying attention by hand.

How the recursion (14) slows down with volume thus depends on the details of its implementation. Typically, one has to cover a macroscopic energy range, i.e., $E_{\max} - E_{\min} \sim V$. The optimal slowing down of a single multicanonical simulation on this range is $\sim V^2$, corresponding to a random walk in the energy.[1] Of order $V^{0.5}$ simulations are needed to iterate from an initial canonical distribution up to covering the entire energy range multicanonically. This leads to an optimal slowing down $\sim V^{2.5}$ for the recursion.

4. ACCUMULATIVE RECURSION

I now introduce a recursion which calculates $\bar{\beta}^{n+1}(E)$ on the basis of the statistics accumulated in all previous runs $n, n-1, \ldots, 1$. For this purpose let us first rewrite (14a) as

$$\bar{\beta}^{n+1}(E) = \varepsilon^{-1} \ln[H^n(E+\varepsilon)/H^n_{\beta}(E)] \tag{17}$$

where

$$H^n_{\beta}(E) = H^n(E) \exp[-i\bar{\beta}^n(E)\varepsilon] \tag{18}$$

Equation (17) still holds when $H^n(E)$ and $H^n_{\beta}(E)$ are replaced by nonzero linear combinations $\hat{H}^n(E)$ and $\hat{H}^n_{\beta}(E)$:

$$\hat{H}^n(E) = \sum_{m=0}^{n} W^m(E)\, H^m(E) \tag{19a}$$

$$\hat{H}^n_{\beta}(E) = \sum_{m=0}^{n} W^m(E)\, H^m_{\beta}(E) \tag{19b}$$

The accumulated statistics can be presented by suitable choice of the weight factors $W^m(E)$. The optimal choice is not clear, as it may depend nontrivially on the dynamics. In practice

$$W^m(E) = \frac{\min[H^n(E+\varepsilon), H^m(E)]}{\max[H^n(E+\varepsilon), H^m(E)]} \tag{20}$$

has worked well. It relies on the conservative assumption that each contribution to the estimate

$$\bar{\beta}^{n+1}(E) = \varepsilon^{-1} \ln[\hat{H}^n(E+\varepsilon)/\hat{H}^n_\beta(E)] \tag{21}$$

will be as good as its weakest part. This equation is supplemented by

$$\bar{\beta}^{n+1}(E) = \bar{\beta}^{n+1}(E+\varepsilon) \tag{22}$$

for the case that either $\hat{H}^n(E+\varepsilon)$ or $\hat{H}^n_\beta(E)$ has insufficient statistics. To provide some feeling for the estimator (21), let me discuss two special cases.

(a) When the desired, flat distribution is already reached, the weight factors (20) equal 1 up to statistical fluctuations. Let us ignore fluctuations for the moment. Then $\hat{H}^{n-1}(E+\varepsilon) = \hat{H}^{n-1}(E)$ holds before the nth run, which uses $\bar{\beta}^n(E)$ as defined by Eq. (21). In the nth recursion $H^n(E+\varepsilon) = H^n(E)$ is obtained by assumption. This leads to $\hat{H}^n(E+\varepsilon) = \hat{H}^{n-1}(E+\varepsilon) + H^n(E+\varepsilon)$ and $\hat{H}^n_\beta(E) = \hat{H}^{n-1}_\beta(E) + H^n(E)\exp(-\bar{\beta}^n\varepsilon)$. Equations (19), (21) yield $\bar{\beta}^{n+1}(E) = \bar{\beta}^n(E)$, i.e., the $\beta(E)$ function is a fixed point when the sampled distribution is flat.

(b) Consider the first recursion, carried out with $\bar{\beta}^0(E) \equiv 0$. The sampling results will be $H^0(E+\varepsilon)/H^0(E) = n(E+\varepsilon)/n(E)$, again up to statistical fluctuations. Recursion (21) yields $\bar{\beta}^1(E) = \varepsilon^{-1} \ln[n(E+\varepsilon)/n(E)]$, which is already the final multicanonical answer due to the fact that we have neglected statistical fluctuations. Quite generally it can be shown that the desired multicanonical function $\beta(E)$ is an attractive fixed point of the recursion.

In practice there may be severe statistical fluctuations due to only few, correlated entries in $H^n(E+\varepsilon)$, $H^n(E)$, or both. If the number of entries in both arrays is small, but approximately equal [$W^n(E) \approx 1$], Eqs. (19) guarantee that increase from $H^{n-1} \to H^n$ is in proportion the generated statistics (assuming similar autocorrelation times in runs $n-1, n-2, \ldots$). If the number of entries is only small in either $H^n(E+\varepsilon)$ or $H^n(E)$, the weight factor (20) corrects for the asymmetry. The larger statistics is reduced to the smaller one, and the smaller is even more suppressed. As the ratio $\hat{H}^n(E+\varepsilon)/\hat{H}(E)$ determines the estimate $\bar{\beta}^n(E)$, it is clear that a large statistical fluctuation in either the numerator or the denominator is sufficient to destroy the entire estimate. The weight factor prevents this.

In ref. 17 we did not supplement the present recursion by a median restriction of the type (14b), although this might lead to further improvements. Without such a restriction, typically the recursion leads quickly to rather high $\bar{\beta}$ values, and works its way back from the corresponding low-

energy values through the entire energy range. Occasionally this has led to "hang-up" situations, for which a simple "retreat" strategy has turned out to be sufficient. For the case of the generalized Ising model, the appendix gives and explains an actual program listing. A generalization of my recursion to nonflat distributions, for instance, those proposed in ref. 20, would be straightforward.

5. NUMERICAL TESTS

Results for the $3d$ IF and the $3d$ EAI spin glass are reported. To measure the performance quantitatively, I use the "tunneling time" as defined in refs 1 and 2. This is the average computer time τ, measured in updates, which it takes to proceed from $E_{\min}$ to $E_{\max}$ and back. For instance, in the IF the degenerate ground states characterized by all spins ups and all spins down are separated by a large free energy barrier. In canonical simulations a tunneling process may lead from the up to the down state or vice versa. Therefore the notion "tunneling time." But it should be noted that the multicanonical method overcomes free energy barriers actually not through a tunneling process, but by moving along valleys, which are connected to the disordered phase.

To keep the relation to the program listing in the appendix close, I shall use

$$I_A = \tfrac{1}{4}(-E + dN) \qquad \text{with} \quad N = L^d \tag{23}$$

instead of the energy defined by (2). The rationale of I_A is its range:

$$I_A = 0, 1, 2, \ldots, dN/2 \tag{24}$$

in typical increments of 1. For comparison, we had $-dN \leqslant E \leqslant dN$ in typical increments of 4. Consequently, for the purposes of programing I_A is far more convenient. Functions of E are now interpreted as functions of I_A in the obvious way, i.e., $\beta[E(I_A)] \to \beta(I_A)$, and so on.

5.1. Three-Dimensional Ising Ferromagnet

The first few terms of the low-temperature expansion on a finite (but sufficiently large) lattice are collected in Table I. The present computer program is unsuitable to cope with $n(I_A) = 0$ for $I_A = (3N/2) - 1$, $(3N/2) - 2$, and $(3N/2) - 4$. I just bypass[3] the problem by restricting the updating to the range $I_A \leqslant N_{\max} = (3N/2) - 5$. Proposals with $I_A > N_{\max}$ are simply rejected.

[3] It is straightforward but tedious to modify the subroutine UPMUCA of the appendix such that it can cope with $n(I_A)$ for (isolated) I_A values.

Table I. Finite-Lattice Low Temperature Expansion for the 3D IF ($N=L^3$)

E	I_A	$n(I_A)$
$-3N$	$3N/2$	2
$-3N+4$	$(3N/2)-1$	0
$-3N+8$	$(3N/2)-2$	0
$-3N+12$	$(3N/2)-3$	$2N$
$-3N+16$	$(3N/2)-4$	0
$-3N+20$	$(3N/2)-5$	$6N$
$-3N+24$	$(3N/2)-6$	$2N^2-14N$
$-3N+28$	$(3N/2)-7$	$30N$
$-3N+32$	$(3N/2)-8$	$6N^2-66N$
$-3N+36$	$(3N/2)-9$	$(2N^3-42N^2+1252N)/6$

We want to calculate multicanonical parameters for the temperature range infinity down to zero. Simulations with $\beta \equiv 0$ are peaked around $I_A = N_{\min} = 3N/4$. We therefore fix the function β to $\beta(I_A) = 0$ for $I_A \leqslant N_{\min}$. For $I_A > N_{\min}$ we perform the multicanonical recursion of Section 4. The covered range of lattices was $4 \leqslant L \leqslant 16$. In a first set of runs the recursion was applied until the system tunneled at least 60 times. The (expected) experience from these runs is that the recursion remained stable after the first tunneling. The tunneling time $\bar{\tau}$ is then measured after the first tunneling has occured, while continuing to update the parameters. Table II collects the measured tunneling times $\bar{\tau}$, and states on how many tunneling events n_τ the estimates rely.

By τ_0 I denote the time (as always in updates) it takes until the first tunneling has taken place. This is essentially the time our recursion needs to provide a reliable estimate of the multicanonical parameters, and it will therefore be called the *recursion time*. Two estimates, $\bar{\tau}_0^a$ and $\bar{\tau}_0^b$ are given in Table II. They differ by the number of sweeps performed before the multicanonical parameters are updated (i.e., the subroutine UPMUCA of the

Table II. Tunneling and Recursion Times for the 3D IF

L	n_τ	$\bar{\tau}$	$n_{\tau_0^a}$	$\bar{\tau}_0^a$	$n_{\tau_0^b}$	$\bar{\tau}_0^b$	n_{τ_1}	$\bar{\tau}_1$
4	548	719 (19) E01	126	661 (41) E02	126	557 (46) E02	111	731 (58) E01
6	354	126 (05) E03	252	195 (20) E04	126	219 (23) E04	145	129 (10) E03
8	559	881 (23) E03	126	311 (55) E05	126	253 (50) E05	125	839 (69) E03
12	322	118 (06) E05	140	95 (32) E07	164	90 (15) E07	141	127 (11) E05
16	577	760 (30) E05	180		2	14(big) E09	180	746 (54) E05

appendix is called). A sweep is defined by updating $N = L^d$ spins. For $\bar{\tau}_0^a$, UPMUCA was called every 120 sweeps, whereas for $\bar{\tau}_0^b$, it was called every N sweeps. Within the (still large) statistical errors there is no noticeable difference.

The values $n_{\tau_0^a}$ and $n_{\tau_0^b}$ are the numbers of $\beta(I_A) \equiv 0$ restarts on which the respective estimates rely. As the average time needed per recursion is substantially higher than the average tunneling time τ, I have limited the τ_0 analysis to $L \leqslant 12$. The given error bars are somewhat unreliable, as the obtained distributions have long tails toward large τ_0 values. Figure 1 employs a logscale for τ_0 to show the histograms for τ_0^b. The distributions for the tunneling times τ themselves are more reasonably (Poisson-like) behaved.

Figure 2 shows the increase of τ and τ_0^b with volume on a log–scale. The straight lines correspond to the fits $\tau = cV^\delta$ and to $\tau_0^b = c_0 V^{\delta_0}$. The results for the fit parameters are

$$\ln(c) = -0.53 \pm 0.16, \qquad \delta = 2.249 \pm 0.021 \qquad (Q = 0.18) \qquad (25)$$

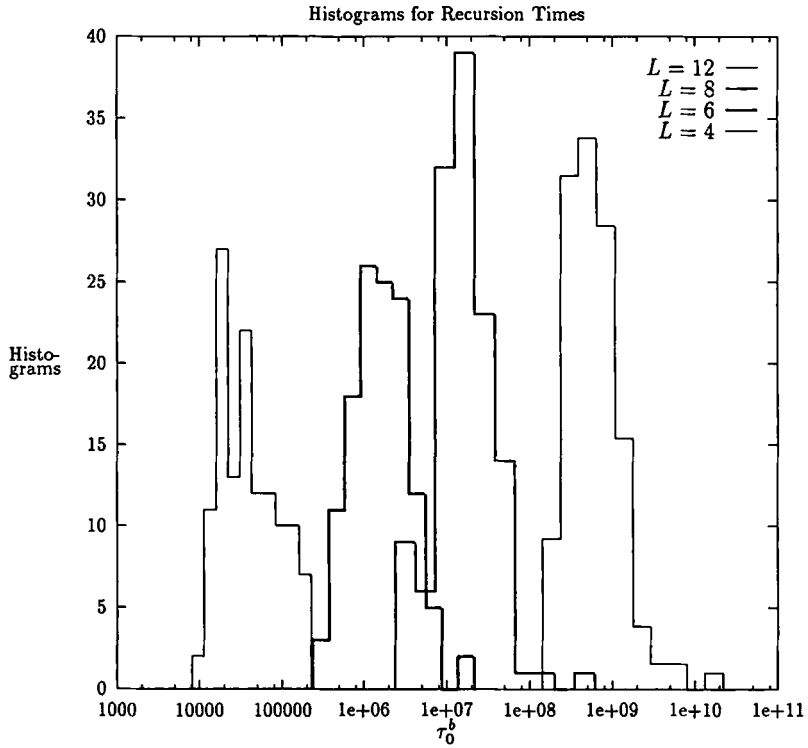

Fig. 1. Histograms for the 3D IF recursion time τ_0^b on lattices of size L^3.

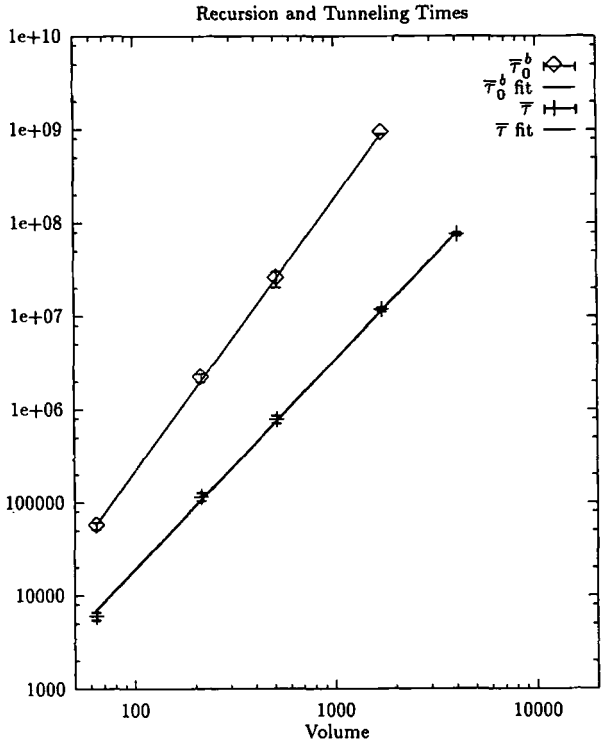

Fig. 2. Estimates for the recursion time $\bar{\tau}_0^b$ and the tunneling time $\bar{\tau}$ for the 3D IF.

and

$$\ln(c_0) = -1.24 \pm 0.17, \qquad \delta_0 = 2.931 \pm 0.023 \qquad (Q = 0.70) \qquad (26)$$

where Q is the goodness of fit[27]. It should be remembered that the lower bounds are $\delta = 2$ (see ref. 1) and $\delta_0 = 2.5$ (see Section 3).

To demonstrate that after a few tunneling events the multicanonical parameters are indeed already useful, I have also measured a tunneling time τ_1 obtained by fixing the multicanonical parameters after the first four tunneling events. Table II contains also the corresponding estimates $\bar{\tau}_1$. Within the statistical errors, there is no difference with the estimates $\bar{\tau}$.

5.2. Three-Dimensional Edwards–Anderson Ising Spin Glass

First I present some results from ref. 17 which are not contained in this reference. For fixed lattice size L, tunneling times τ are found to vary

greatly for different J_{ik} realizations. For each lattice size, Fig. 3 connects the tunneling times, sorted in decreasing order. For $L=4$–8 there are 512 different realizations per lattice. For $L=12$ there are only seven realizations, depicted at $64(i-1)$, $i=1, ..., 7$. The lines are drawn to guide the eye. Figure 4 depicts histograms for the $L=4$–8 tunneling times. In both figures a logarithmic scale is used for τ. The worst realization have dramatically larger tunneling times than *typical* ones, defined by the median value $\tau_{0.5}$. This leads to large differences between the mean value $\bar{\tau}$, which determines the needed computer time, and the median value $\bar{\tau}_{0.5}$. These values are collected in Table III. With increasing lattice size the discrepancy between mean and median increases dramatically (the $L=12$ data have to be considered unreliable for this purpose). This lack of self-averaging of the spin glass with respect to the multicanonical tunneling time is somewhat surprising, and needs to be better understood. Also collected in the table are the smallest $\tau_{0.0}$ and largest $\tau_{1.0}$ tunneling times found on the investigated realizations.

For $L \leqslant 8$ typical spin-glass realizations, i.e., the realizations corresponding to the median $\tau_{0.5}$ tunneling times of Table III, I have performed

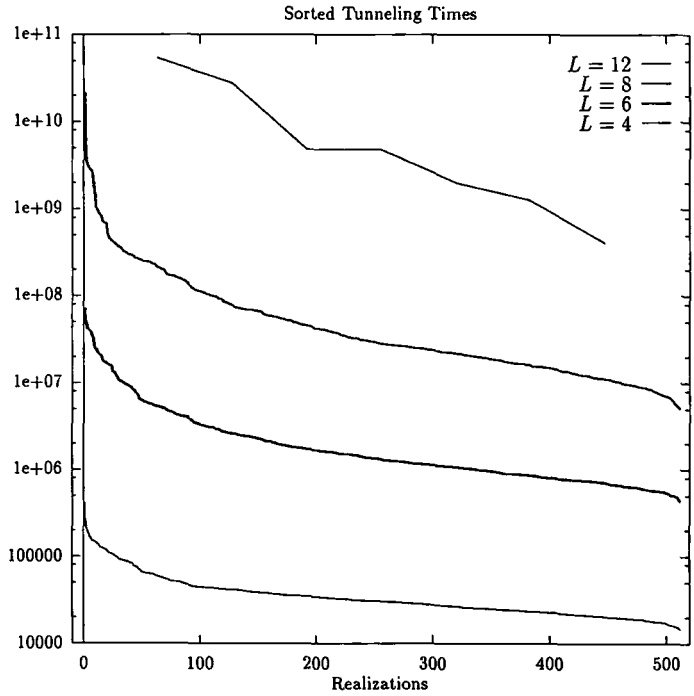

Fig. 3. Sorted 3D EAI tunneling times for various lattices of sizes L^3.

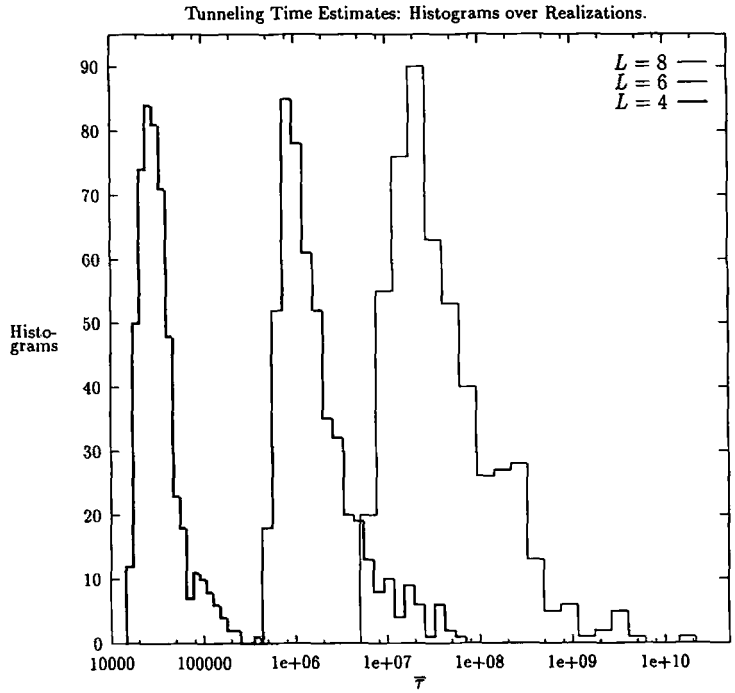

Fig. 4. Histograms with respect to realizations for the 3D EAI tunneling time estimates $\bar{\tau}$ on lattices of size L^3.

the same analysis as for the $3d$ IF in the previous subsection. The results are collected in Table IV. An interesting and unexpected result is that I find $\bar{\tau}_1$ systematically smaller than $\bar{\tau}$, i.e., further applications of the recursion relation make the tunneling worse. My tentative interpretation is that the flat distribution is not optimal. Due to statistical fluctuations, one can then

Table III. Mean $\bar{\tau}$ and Some q-Tiles τ_q for the 3D EAI Tunneling Time

L	$\bar{\tau}$	$\bar{\tau}_{0.0}$	$\bar{\tau}_{0.5}$	$\bar{\tau}_{1.0}$
4	398 (15) E02	144 E02	304 E02	411 E03
6	336 (30) E04	436 E03	131 E04	670 E05
8	171 (46) E06	505 E04	282 E05	213 E08
12	139 (77) E08	408 E06	481 E07	544 E08

Table IV. Tunneling and Recursion Times for Typical 3D EAI Spin-Glass Realizations

L	n_τ	$\bar{\tau}$	$n_{\tau_0^a}$	$\bar{\tau}_0^a$	$n_{\tau_0^b}$	$\bar{\tau}_0^b$	n_{τ_1}	$\bar{\tau}_1$
4	185	332 (25) E02	126	523 (19) E02	126	409 (22) E02	270	228 (14) E02
6	256	181 (12) E04	126	150 (10) E04	126	172 (11) E04	357	131 (08) E04
8	134	272 (26) E05	252	245 (13) E05	252	253 (16) E05	207	203 (16) E05

imagine that immediately after one of the first few tunneling events the generated multicanonical parameters are positively correlated toward a more optimal choice. As for the IF, the recursion times $\bar{\tau}_0^a$ and $\bar{\tau}_0^b$ are practically identical. However, a second unexpected result is that now the recursion times take the same order of magnitude as the tunneling times.

Runs were also performed on an $L = 12$ realization, but they did not allow statistically reliable conclusions. A lesson to be learned from these runs is that it is advantageous to perform several independent starts when applying the recursion to systems on the edge of what can be done.

The subsequent results are obtained from straight-line fits to the equations $\tau = cV^\delta$, $\tau_1 = c_1 V^{\delta_1}$, and $\tau_0^b = c_0 V^{\delta_0}$:

$$\ln(c) = -3.04 \pm 0.29, \qquad \delta = 3.24 \pm 0.06 \qquad (Q = 0.40) \tag{27a}$$

$$\ln(c_1) = -3.61 \pm 0.24, \qquad \delta_1 = 3.28 \pm 0.05 \qquad (Q = 0.39) \tag{27b}$$

and

$$\ln(c_0) = -2.23 \pm 0.24, \qquad \delta_0 = 3.09 \pm 0.04 \qquad (Q = 0.78) \tag{28}$$

Here, as well as in the previous section, the routine GFIT from ref. 27 gives results perfectly compatible with the linear-fit results. A figure corresponding to (27) and (28) looks similar to Fig. 2, but is not very instructive, as all three fits lines are almost on top of one another. The exponent δ is smaller than the one reported in ref. 17. The reason is that it is defined differently. In ref. 17 the tunneling time was averaged over all realization, whereas here I have picked single, typical realizations. There is evidence that for the worst realizations, the tunneling time slows down exponentially with L.

6. SUMMARY AND CONCLUSIONS

For the 3D Ising ferromagnet it is clear that the FSS methods employed in refs. 1 and 8 provide reliable estimates of the multicanonical

parameters more efficiently than the recursion of this paper. On the other hand, the FSS approach breaks down[2] for the important class of disordered systems. Then recursions like the one of this paper become crucial, and the Ising ferromagnet is still a suitable testing ground to set quantitative performance scales. These are now given, for the first time, by Tables III and IV. Table IV corresponds to the important case of a typical Edwards–Anderson Ising spin glass. Future investigations will have to cope with these standards. It is my hope that they will bring improvements in the constant factor, and possibly toward a V^2 power-law behavior, which is optimal for any kind of local random walk behavior.

APPENDIX

In this appendix I describe the actually used computer implementation for the accumulative recursion of the multicanonical parameters. The relevant Fortran subroutine is listed as Table V. Parameters (to be set) are the dimension ND and the lattice size NL.

The argument IRPT keeps track of the number of repeated calls to UPMUCA. In an outside DO-loop IRPT runs from 1 to NRPT. Inside our subroutine NRPT is only needed to dimension the LOGICAL array LRTRT, which keeps track of the number of "retreats". A parameter not needed at all in our subroutine is NSW. It denotes the number of update sweeps performed in between the calls to UPMUCA. NAMIN sets the lower bound on the IA range (I_A of Section 5) to which the recursion is applied.

Most arguments are passed through COMMON blocks. On entry the array HA contains the newly assembled statistics, i.e., the histogram of the number of times a certain IA value has been visited during the last NSW sweeps. (The information is collected after each single spin update.) Further arguments passed by the COMMON block MEAH (measurements) are: IAMIN, the smallest IA value encountered so far (not used in UPMUCA); IAMAX, the largest IA value encountered so far; ITMIN, the smallest IA value encountered during the last NSW sweeps; and ITMAX, the largest IA value encountered during the last NSW sweeps. The meaning of the array(s) HAMU is explained by the comments. Central for the code are the lines

```
W1=HAMIN/HAMAX
HAMU(IA, 3)=HAMU(IA, 3)+W1*HA(IA)
HAMU(IA, 4)=HAMU(IA, 4)+W1*HA(IAP1)*EXP(-4.0D0*B(IAP1))
```

Table V. Fortran Subroutine for Update of Multicanonical Parameters

```
      SUBROUTINE UPMUCA (IRPT)
C HAMUA(*, 1):over-all sum (record keeper only).
C HAMUA(*, 2): LRTRT adjusted over-all sum (record keeper only).
C HAMUA(*, 3): 1. weighted sum; HAMUA(*, 4): 2. weighted sum.
      IMPLICIT REAL*8 (A-H, O-Z)
      IMPLICIT LOGICAL (L)
      PARAMETER (ND=3, NL=08, NS=NL**ND, NRPT=100, NSW=NS)
      PARAMETER (NNH=(ND*NS)/2, NAMIN=NNH/2, FRTRT=3.DO, EPS=1.D-8)
      PARAMETER (HMIN=1.ODOO*FLOAT(NS)*FLOAT(NSW))
      COMMON/MEAH/HA(O:NNH), IAMIN, IAMAX, ITMIN, ITMAX
      COMMON/MUCA/B(O:NNH), A(O:NNH), HAMU(O:NNH, 4), LRTRT(NRPT)
      DO IA=ITMIN, ITMAX
      HAMU(IA, 1)=HAMU(IA, 1)+HA(IA)
      HAMU(IA, 2)=HAMU(IA, 2)+HA(IA)
      END DO
C Retreat strategy (below) implies: range up to IAMAX.GE.ITMAX
      IAMAM1=IAMAX-1
      DO IA=NAMIN, IAMAM1
      IAP1=IA+1
      HAMIN=MIN(HA(IA), HA(IAP1))
      HAMAX=MAX(HA(IA), HA(IAP1))
      IF(HAMIN.GT.0.5DOO)                                    THEN
      W1=HAMIN HAMAX
      HAMU(IA, 3)=HAMU(IA, 3)+W1*HA(IA)
      HAMU(IA, 4)=HAMU(IA, 4)+W1*HA(IAP1)*EXP(-4.ODOO*B(IAP1))
                                                                END IF
C BETA update (after retreat HAMIN.LE.0.5 possible):
      HAMUMIN=MIN(HAMU(IA, 3), HAMU(IA, 4))
      B(IAP1)=B(IA)
      IF(HAMUMIN.GT.EPS) B(IAP1)=-0.25DOO*LOG(HAMU(IA, 4)/HAMU(IA, 3))
      END DO
C Retreat strategy for hung-up situations:
      LRTRT(IRPT)=.FALSE.
C Besides retreat, update of MUCA A-array is performed
C (range up to IAMAX.GE.ITMAX is needed for this reason).
      DO IA=NAMIN, IAMAM1
      IAP1=IA+1
      IF(HAMU(IAP1, 2).GT.HMIN.
      AND.HAMU(IAP1, 2).GT.FRTRT*HAMU(IA, 2))          THEN
C The program may need mofications, if there are
C energy values without states in the .LE.IMAX range
      IF(HAMU(IA, 2).EQ.O) PRINT*, 'UPMUCA Warning: IA=', IA
      IF(.NOT.LRTRT(IRPT)) PRINT*
     6'RETREAT! IRPT, IA, HAMUs:', IRPT, IA HAMU(IAP1, 2), HAMU(IA, 2)
      LRTRT(IRPT)=.TRUE.
                                                                END IF
      IF(LRTRT(IRPT))                                    THEN
      HAMU(IAP1, 2)=HAMU(IAP1, 2)/FRTRT
      B(IAP1)=0.ODOO
      HAMU(IAP1, 3)=HAMU(IAP1, 3)/FRTRT
      HAMU(IAP1, 4)=HAMU(IAP1, 4)/FRTRT
                                                              END IF
      A(IAP1)=A(IA)-4.ODOO*(B(IAP1)-B(IA))*FLOAT(IA)
      END DO
      DO IA=(IAMAX+1), NNH
      B(IA)=B(IA-1)
      A(IA)=A(IA-1)
      END DO
      RETURN
      END
```

which implement Eqs. (19) and (20) recursively. Next, the arrays A and B correspond to the multicanonical functions β and α. The lines

```
B(IAP1)=-0.25D0*LOG(HAMU(IA, 4)/HAMU(IA, 3))
```

and

```
A(IAP1)=A(IA)-4.0D0*(B(IAP1)-B(IA))*FLOAT(IA)
```

implement Eqs. (21) and (11). Of course, A(NAMIN)=0. The parameter EPS prevents the β-recursion from taking place without sufficient statistics, and otherwise Eq. (22) is chosen.

Some complications arise, mainly because a "retreat" strategy has been implemented to get out of certain "hang-up" situations. An extreme difference between HAMU(IA+1,2) and HAMU(IA,2) can turn out to be artificial, such that its statistics is better not trusted. "Extreme" is defined by the parameter FRTRT, put to 3 in the presented code. When the limit thus defined is exceeded, the assembled statistics is reduced in weight by the factor 1/FRTRT and $\beta(I_A)$ is put in the corresponding energy region to $\beta(I_A)=0$ for the next recursion. However, one has to choose FRTRT to be very large (around 200) if one wishes to calculate multicanonical parameters for an 24^3 IF in the range described in Section 5.1. The reason is the peculiar IF density of states anomaly from $I_A=(3N/2)-7$ to $I_A=(3N/2)-6$ (see Table I).

ACKNOWLEDGMENTS

I would like to thank Wolfhard Janke and Claus Vohwinkel for useful discussions. The manuscript was partly written at the Institut für Physik, Johannes Gutenberg Universität, Mainz. The author thanks Kurt Binder and his group for their hospitality. This research was partially funded by the Department of Energy under contracts DE-GF05-87ERH0319 and DE-FC05-85ER2500.

REFERENCES

1. B. Berg and T. Neuhaus, *Phys. Lett. B* **267**:249 (1991); *Phys. Rev. Lett.* **68**:9 (1992).
2. B. Berg and T. Celik, *Phys. Rev. Lett.* **69**:2292 (1992); *Int. J. Mod. Phys. C* **3**:1251 (1992).
3. A. P. Lyubartsev, A. A. Martsinowski, S. V. Shevkunov, and P. N. Vorontsov-Velyaminow, *J. Chem. Phys.* **96**:1176 (1992).
4. E. Marinari and G. Parisi, *Europhys. Lett.* **19**:451 (1992).
5. B. Berg, *Int. J. Mod. Phys. C* **3**:311 (1992).
6. A. Hüller, *Z. Phys. B* **88**:79 (1992).
7. C. Borgs and S. Kappler, *Phys. Lett. A* **171**:2011 (1992).

8. B. Berg, U. Hansmann, and T. Neuhaus, *Z. Phys. B* **90**:229 (1993).
9. B. Grossman and M. L. Laursen, *Nucl. Phys. B* **408**:637 (1993).
10. K. Rummukainen, *Nucl. Phys. B* **390**:621 (1993).
11. A. Billoire, T. Neuhaus, and B. Berg, *Nucl. Phys. B* **396**:779 (1993); **413**:795 (1994).
12. W. Janke, Recent developments in Monte-Carlo simulations of first-order phase transitions, in *Computer Simulations Studies in Condensed Matter Physics VII*, D. P. Landau, K. K. Mon, and H.-B. Schüttler, eds. (Springer, Berlin, 1994), pp. 29–44.
13. J. Lee, *Phys. Rev. Lett.* **71**:211 (1993); Erratum **71**:2353 (1993).
14. B. Berg, unpublished; B. Berg, U. Hansmann, and Y. Okamoto, *J. Phys. Chem.* **99**:2236 (1995).
15. U. Hansmann and Y. Okamoto, *J. Comput. Chem.* **14**:1333 (1993).
16. M.-H. Hao and H. A. Scheraga, *J. Phys. Chem.* **98**:4940 (1994).
17. B. Berg, U. Hansmann, and T. Celik, *Phys. Rev. B* **50**:16444 (1994).
18. W. Kerler and P. Rehberg, *Phys. Rev. E* **50**:4220 (1994).
19. W. Janke and S. Kappler, *Phys. Rev. Lett.* **74**:212 (1995).
20. B. Hesselbo and R. Stinchcombe, *Phys. Rev. Lett.* **74**:2151 (1995).
21. N. B. Wilding and M. Müller, *J. Chem. Phys.* **102**:2562 (1995).
22. B. Berg, *Nature* **361**:708 (1993).
23. G. M. Torrie and J. P. Valleau, *J. Comput. Phys.* **23**:187 (1977).
24. Kerson Huang, *Statistical Mechanics* (Wiley, New York, 1987), p. 134.
25. M. Mézard, G. Parisi, and M. A. Virasoro, *Spin Glasses and Beyond* (World Scientific, Singapore, 1987).
26. B. Baumann, *Nucl. Phys. B* **285**:391 (1987).
27. W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes* (Cambridge University Press, Cambridge, 1988).