Fast flat-histogram method for generalized spin models

S. Reynal* and H. T. Diep

Laboratoire de Physique Théorique et Modélisation, CNRS-Université de Cergy-Pontoise (UMR 8089), 2 avenue A. Chauvin, F-95302 Cergy-Pontoise Cedex, France

We present a Monte Carlo method that efficiently computes the density of states for spin models having any number of interaction per spin. By combining a random walk in the energy space with collective updates controlled by the microcanonical temperature, our method yields dynamic exponents close to their ideal random-walk values, reduced equilibrium times, and very low statistical error in the density of states. The method can host any density of states estimation scheme, including the Wang-Landau algorithm and the transition matrix method. Our approach proves remarkably powerful in the numerical study of models governed by long-range interactions, where it is shown to reduce the algorithm complexity to that of a short-range model with the same number of spins. We apply the method to the q-state Potts chains ($3 \leq q \leq 12$) with power-law decaying interactions in their first-order regime; we find that conventional local-update algorithms are outperformed already for sizes above a few hundred spins. By considering chains containing up to 2^{16} spins, which we simulated in fairly reasonable time, we obtain estimates of transition temperatures correct to five-figure accuracy. Finally, we propose several efficient schemes aimed at estimating the microcanonical temperature.

PACS number(s): 05.10.Ln, 64.60.Cn, 75.10.Hk

I. INTRODUCTION

Long-range spin models have drawn increasing interest in the last decade, both in the microscopic modeling of a variety of systems ranging from model alloys [1] to spin glasses [2] to neural networks [3], and as a powerful laboratory frame to investigate fundamental issues in the physics of critical phenomena. These include, e.g., the effect of dimensionality [4], the crossover from short-range to long-range behavior [5–7], mean-field driven phase transitions [8], and possible connections with Tsallis's nonextensive thermodynamics [9–11]. Monte Carlo (MC) methods have now gained a prominent role in the investigation of phase transitions in these models [12–18]. In particular, a major breakthrough was recently initiated by the introduction of a (canonical) cluster algorithm able to overcome the algorithm complexity inherent to long-range (LR) models, namely, the need to take a huge number of interactions into account at each Monte Carlo step (MCS) [12]. In a recent article, we proposed a generalization of this algorithm to simulations in the multicanonical ensemble [19]. It is the goal of the present work to introduce a general and versatile method aimed at embedding any cluster update scheme in a flat histogram algorithm, with special emphasis given to LR spin models.

Whether short- or long-range interactions are considered, canonical MC simulations of long-range spin models suffer indeed from severe shortcomings, the use of cluster updates notwithstanding. First and foremost, models exhibiting first-order phase transitions or complicated energy landscapes experience supercritical slowing down [20]: the time needed

for the dynamics to tunnel through free energy barriers grows exponentially with the lattice size, leading to quasi-ergodicity-breaking and unreliable statistics. Second, the computation of free energies and related thermodynamic quantities is highly involved, and a precise determination of the order of the transition is often intractable. In practice, these shortcomings preclude the use of canonical MC algorithms at first-order transitions except at modest lattice sizes and in the case of weakly first-order transitions.

An efficient approach aimed at overcoming this limitation is the simulation in generalized ensembles [21,22], in particular its multicanonical flavor initially proposed by Berg [20,23,24], reconsidered in the context of transition matrix dynamics [25,26] and recently revisited by Wang and Landau [27,28]. The key idea here is to artificially enhance rare events corresponding to local maxima in the free energy, by feeding the Markov chain with an appropriate distribution $w(E)$. In the *multicanonical ensemble*, $w(E)$ is set to the inverse of the density of states, so that the resulting dynamics is a random walk in the energy space that yields a flat histogram of the energy. Other ensembles have been proposed in the last decade, including the $1/k$ ensemble, which enhances low-energy states [29], and very recently, the optimal ensemble, which aims at optimizing the distribution $w(E)$ with respect to the local diffusivity of the random walker, so that tunneling times are minimized [30,31]. While still broad, histograms engendered by these last ensembles are no longer flat; in the optimal ensemble for instance, the histogram is slightly peaked around the critical region, so that the larger time spent by the random walker inside the critical region compensates the lower diffusivity in this region.

When implemented through local (i.e., single-spin) updates [7], simulations in the multicanonical ensemble suffer, however, from two serious hurdles. First, while tunneling times—measured in Monte Carlo steps (MCSs)—are re-

―――――
*Permanent address: E.N.S.E.A., 6 avenue du Ponceau, 95014 Cegry Cedex, France. Electronic address: reynal@ensea.fr; URL: http://www-reynal.ensea.fr

duced from an exponential to a power law $\tau \sim L^z$ of the lattice size, the dynamic exponents z are still substantially higher than the ideal value $z \sim D$ that should be expected from the dynamics of a random walker [32,33]. This observation, as we will witness in this paper, applies equally well to effective autocorrelation times and to equilibrium times; this represents a serious hindrance in terms of scalability, in particular whenever a higher precision is desired and large amounts of decorrelated data need to be gathered. In this respect, it is important to mention that correlations between successive measurements do not only have an impact on the statistical efficiency of multicanonical production runs, yet also represent a source of systematic error regarding the estimation of the density of states [34]. A second impediment to the scalability of local-update implementations specifically relates to long-range models. Here, the very presence of long-range interactions makes the computation of the energy—an essential ingredient of multicanonical methods—a very time consuming operation, namely, one associated with an $O(L^{2D})$ algorithm complexity. As a result, the demand on CPU time needed to generate perfectly decorrelated statistics grows as L^{z+2D}, with $z > D$.

In this paper, we present a Monte Carlo method which successfully tackles these issues by performing simulations in the multicanonical ensemble using collective updates. Our methods combine the fast-decorrelating capabilities of cluster algorithms with the versatility of flat-histogram methods in an efficient and straightforward way, and with wide applicability in view. In particular, it can be readily combined with any iteration scheme dedicated to the estimation of the density of states, e.g., Wang-Landau's method [27] or transition matrix algorithms [25]. Additionally, while our method is presented here in the context of long-range spin models, where it gives drastic improvements over commonly used methods, it is perfectly applicable to any class of models for which a cluster algorithm exists in the canonical ensemble.

Noteworthy enough, embedding a collective update scheme in a multicanonical algorithm is not straightforward, however, due to the fundamentally *nonlocal* nature of the multicanonical weight $w(E)$. Indeed, cluster algorithms depend heavily upon particular symmetries of the model Hamiltonian, which $w(E)$ does not keep track of; in particular, there is no longer a canonical temperature. With simulations of spin models with nearest-neighbor interactions in view, several attempts have been made at combining cluster updates with multicanonical methods in some way or another during the last decade: the multibond algorithm [32,35–37] or variants thereof targeting Wang-Landau's algorithm [31,38] simulate the model in its spin-bond representation; Rummukainen's hybridlike two-step algorithm lumps together a microcanonical cluster algorithm and a multicanonical daemon refresh [39]. As opposed to these, however, our method relies on a cluster-building process which simply depends on the microcanonical temperature of the current configuration—a quantity that may be readily derived from the estimated density of states—in order to determine appropriate bond probabilities. In particular, it does not require prior knowledge of the transition temperature, as is the case in the multibond method. We further show that our approach makes it particularly straightforward to incorporate two op-

timization schemes dedicated to LR models [12,40], which cut down the algorithm complexity from $O(L^{2D})$ to $O(L^D \ln L^D)$. As a result, the total demand on CPU time with respect to uncorrelated data is reduced to approximately $L^{2D} \ln L^D$, since cluster updates also lower z to around D; where LR models are concerned, the benefit of cluster updates is thus clearly twofold. Let us also mention that, as a by-product, using cluster updates provides improved estimators for the statistical moments of the order parameter [41] and for spin-spin correlation functions; for instance, the last quantity can be better estimated by counting the fraction of time two given sites belong to the same cluster [42,43]. Further interesting information, including information connected with fractal geometry, may also be gleaned from cluster statistics [44,45].

Overall, the sharp reduction of the computer load brought about by our method allowed us to study q-state Potts chains with $1/r^{1+\sigma}$ interactions containing up to 2^{16} spins in a few days on a modern Intel-based computer. It must be noted that, with standard multicanonical methods based on single-spin updates, such huge sizes are simply intractable, since the largest size of 2^{16} investigated in this work would demand several months of computation. As regards dynamic performance, we obtain a substantial reduction in the dynamic exponent, from e.g., $z \sim 1.35(3)$ to $z \sim 1.05(1)$ for $q = 6$ and $\sigma = 0.7$. We also show that our method produces faster equilibration, lower effective autocorrelation times, and—where implementations based on the Wang-Landau algorithm are concerned—lower statistical errors in the estimate of the density of states, e.g., of nearly an order of magnitude for $q = 6$; $\sigma = 0.9$, and $L = 512$ spins. As a result, we obtain estimates of transition temperatures that have a noticeably higher precision than those obtained using local updates [7] or standard canonical methods [13,14]. Finally, in order to check that our method did not produce systematic errors, we performed several simulations of the two-dimensional seven- and ten-state Potts models with nearest-neighbor interactions and sizes up to $L = 256 \times 256$. We obtain dynamical exponents close to the ideal random-walk value $z \sim 2$. Although computed from rather modest statistics, our estimate of the interfacial free energy for the largest size reaches a precision of four digits. In this respect, our method compares perfectly well with other methods operating in the multicanonical ensemble, and represents an alternative way for short-range spin models.

The layout of this paper is as follows. In Sec. II, after briefly reviewing some prominent features of multicanonical methods, we explain how we combine a multicanonical weighting with collective updates, with special emphasis given to the detailed balance equation. Section III addresses optimizations dedicated to long-range models. Numerical results regarding the dynamic characteristics of our method are presented in Sec. IV. In Sec. V, we compare our results for the two-dimensional Potts model with nearest-neighbor interactions, with exactly known results, and Sec. VI is devoted to the investigation of the precision of our method in the context of the long-range Potts chain with power-law decaying interactions. Overall, we pay particular attention to comparison with other algorithms operating in the multicanonical ensemble, especially in terms of tunneling rates, dynamical

exponents, and estimates of thermodynamical averages. Finally, we discuss several procedures aimed at estimating the microcanonical temperature, and in particular, how we can efficiently combine our method with the transition matrix approach.

II. A METHOD TO EMBED CLUSTER UPDATES IN A FLAT-HISTOGRAM ALGORITHM

Monte Carlo simulations are based on the generation of a Markov chain of configurations $\{\sigma_i\}$, where each configuration is assigned a weight $w[E(\sigma_i)]$ corresponding to the probability distribution one wishes to sample. In canonical simulations, i.e., carried out at a fixed inverse temperature β, one chooses a Boltzmann weight $w[E(\sigma_i)]=e^{-\beta E(\sigma_i)}$, thus thermodynamical averages are straightforwardly obtained by computing the appropriate moments of the data accumulated at the given temperature. On the other hand, reweighting methods based on multihistogramming [46] are hampered at large lattice sizes by the narrowness of the energy window that is sampled, let alone additional supercritical slowing down. In the multicanonical ensemble, one allows the dynamics to jump across free energy barriers and, from a more general standpoint, to sample wide energy windows, by producing a flat energy distribution over the energy range of interest for the problem at hand. This is formally carried out by setting $w(E)=e^{-S(E)}\propto 1/n(E)$, where $n(E)$ is the density of states and $S(E)$ is the microcanonical entropy. This in effect leads to $N(E)\propto n(E)w(E)=$const for the number of visits to energy E. Since the density of states is obviously *a priori* unknown, $w(E)$ is estimated using an iterative procedure initially fed from, e.g., a canonical guess $w(E)=e^{-\beta_0 E}$ at a carefully chosen inverse temperature β_0, a flat guess $w(E)=1$, or—whenever feasible—a properly scaled estimate obtained at a smaller lattice size. Thermodynamic quantities that depend solely on the energy, like the specific heat or Binder cumulants, can then be computed directly from the estimated density of states. Other quantities, e.g., those depending on the order parameter, are obtained through a reweighting procedure based on data gathered during an additional production run.

Historically, Berg's recursion scheme [47,48] was the first iteration procedure specifically dedicated to multicanonical simulations. It consists in accumulating histogram entries of the energy during each iteration run, and updating $w(E)$ from the histogram of the energy obtained in a previous iteration run, until eventually the histogram becomes flat up to a given tolerance. Entropic sampling [49] more or less boils down to the same key principle. Both methods suffer, however, from poor scalability. Looking at this issue from a slightly different angle, the recently proposed Wang-Landau acceleration method [27,28] updates $w(E)$ in real time during the course of the simulation, performing independent random walks in distinct energy ranges. Since modifying the weight of the Markov chain during a simulation breaks detailed balance, the amount by which $w(E)$ is modified during a given iteration is decreased from one iteration to the other until it reaches a negligible value; hence detailed balance is approximately restored in the last step of the iteration scheme. In this respect, an original approach aimed at reducing the statistical error in the density of states was recently proposed by Yan and de Pablo [50], whereby the density of states is obtained by integrating an instantaneous temperature computed from configurational information or from a so-called multi-microcanonical ensemble. Finally, a large class of iteration schemes have been proposed that are based on matrices of transition probabilities [25,26,51–53] or a combination thereof with Wang-Landau's algorithm [54,55]. Here, the density of states is computed through a so-called broad histogram equation involving infinite temperature transition matrices, where transition matrices keep track of the microcanonical average of the number of potential moves from one energy level to another (Sec. VII gives more details on how our method can efficiently capitalize on transition matrices). Historically, procedures based on transition matrices were termed *flat-histogram methods* in order to distinguish them from Berg's multicanonical method, although both approaches in effect yield a flat, broad histogram. To sum up, the main benefit of multicanonical methods is twofold: first, a wide energy range is sampled, irrespective of the presence of free energy barriers; second, the methods yield a direct estimate of the density of states.

A local-update implementation of a multicanonical algorithm may consist in updating a single spin at a time and accepting the attempted move from state a to state b with a probability given by $W(a\rightarrow b)=\min[1,e^{S(E_a)-S(E_b)}]$. We now show that the microcanonical temperature $\beta(E)$ defined as $dS(E)/dE$ is a relevant quantity for the acceptance rate of this process. Denoting $E_b=E_a+\epsilon$, we expand the probability $W(a\rightarrow b)$ for small ϵ, and obtain $W(a\rightarrow b)\sim\min[1,e^{-\beta(E_a)\epsilon}]$. This shows that, for small enough energy changes, the dynamics is equivalent to that of a canonical simulation at an inverse temperature $\beta(E)$. Our departure point for a collective-update implementation in the multicanonical ensemble is thus to build clusters of spins with the same bond probabilities as would be given by a canonical simulation at inverse temperature $\beta(E)$.

Although our algorithm may be equally well applied to other spin models, e.g., models incorporating disorder or exhibiting a continuous symmetry, we now consider, for the sake of clarity, a generalized ferromagnetic spin model with a $\mathbb{Z}_q$ symmetry, whose Hamiltonian reads $H=-\Sigma_{i<j}J_{ij}\delta_{\sigma_i,\sigma_j}$. Here $J_{ij}>0$ and the σ_i variables can take on integer values between 1 and q. Taking guidance from Swendsen-Wang's cluster algorithm [56], we start from an empty bond set, consider each pair of spins $\{\sigma_i,\sigma_j\}$ in turn, and activate a bond between them with a bond probability given by $\pi_{ij}(E_a)=\delta_{\sigma_i,\sigma_j}[1-e^{-J_{ij}\beta(E_a)}]$, where E_a is the current lattice energy and $\beta(E_a)$ the inverse microcanonical temperature at energy E_a. Efficient ways of estimating $\beta(E)$ are considered later on in Sec. VII. Then, we identify clusters of connected spins using, e.g., a multiple-labeling scheme [57], draw a new spin value at random for each cluster, and accept the attempted move with an acceptance probability $A_{\text{flip}}(a\rightarrow b)$ which ensures that detailed balance is satisfied. The derivation of this probability may be carried out in the following way. First, the total acceptance probability $W(a\rightarrow b)$, i.e., the quantity

that enters detailed balance in such a way that $e^{-S(E_a)}W(a \rightarrow b) = e^{-S(E_b)}W(b \rightarrow a)$, is split into two terms $P(a \rightarrow b)$ and $A_{\text{flip}}(a \rightarrow b)$ representing a *proposed update probability* and an *acceptance probability* for the proposed update, respectively. It is straightforward to show that the choice $A_{\text{flip}}(a \rightarrow b) = \min[1, [P(b \rightarrow a)/P(a \rightarrow b)]e^{S(E_a)-S(E_b)}]$ satisfies the detailed balance equation. Let us denote $\mathcal{B}$ the set of active bonds over the complete graph $\mathcal{G}$ engendered by all possible interactions: the *proposed update probability* is given by the probability to construct a given set $\mathcal{B}$ from an empty bond set, i.e.,

$$P(a \rightarrow b) = \prod_{b_{ij} \in \mathcal{B}} \pi_{ij}(E_a) \prod_{b_{ij} \in \mathcal{G} \setminus \mathcal{B}} [1 - \pi_{ij}(E_a)].$$

After simplification, we obtain for $P(b \rightarrow a)/P(a \rightarrow b)$

$$e^{\beta(E_b)E_b - \beta(E_a)E_a} \prod_{b_{ij} \in \mathcal{B}} \frac{e^{J_{ij}\beta(E_b)} - 1}{e^{J_{ij}\beta(E_a)} - 1};$$

where

$$A_{\text{flip}}(a \rightarrow b) = \min\left[1, \frac{e^{\alpha(E_a)}}{e^{\alpha(E_b)}} \prod_{b_{ij} \in \mathcal{B}} \frac{p_{ij}(E_b)}{p_{ij}(E_a)}\right], \quad (1)$$

where $\alpha(E) = S(E) - \beta(E)E$ and $p_{ij}(E) = e^{J_{ij}\beta(E)} - 1$. This expression can be further simplified if we consider long-range models whose coupling constant depends only on the distance between spins, i.e., $J_{ij} = J(l)$, where $l = \text{dist}(i, j)$. We have for $A_{\text{flip}}(a \rightarrow b)$:

$$A_{\text{flip}}(a \rightarrow b) = \min\left[1, \frac{e^{\alpha(E_a)}}{e^{\alpha(E_b)}} \prod_{l>0} \left(\frac{p_l(E_b)}{p_l(E_a)}\right)^{B(l)}\right], \quad (2)$$

where $B(l)$ stands for the number of bonds of length l. It is worth mentioning that, if one looks at this equation from the standpoint of canonical simulations at inverse temperature β_0, we have $w(E) = e^{-\beta_0 E}$; where $\beta(E) = \beta_0$ and $\alpha(E)$ no longer depends on E. As a result, the acceptance rate $A_{\text{flip}}(a \rightarrow b)$ is equal to 1 and we are back to the original Swendsen-Wang algorithm.

It is also crucial to underline that it is the microcanonical temperature, i.e., the lattice energy in the first place, which entirely governs the construction of clusters; indeed, for a given lattice configuration at energy E, bonds are placed as if the model were simulated at its microcanonical temperature using a Swendsen-Wang algorithm. As a result, cluster bond probabilities change continuously as the lattice configuration walks along the available energy range of the random walk, so that, e.g., small clusters are built in the upper energy range and conversely large clusters in the lower energy range.

III. OPTIMIZATION FOR LONG-RANGE MODELS

A. Computing the lattice energy through FFT acceleration

As is apparent in Eq. (1), determining the acceptance rate of a cluster flip demands that we compute the energy of the new (attempted) lattice configuration, which for long-range models is an $O(L^{2D})$ operation. This is similar to the local-update case, where performing one MC step, i.e., updating L^D spins subsequently, takes a CPU time proportional to the square of the number of spins, seeing that L^D operations are needed after each single spin update to compute the new partial energy between the updated spin and the rest of the lattice. Recently, Krech and Luijten proposed an algorithm that is able to compute the energy of a model governed by translation invariant interactions in $O(L^D \ln L^D)$ operations [40]. This method leans on the convolution theorem and the fast Fourier transform (FFT), for which numerous efficient radix-based implementations are available. As a result, updating the lattice configuration *globally* rather than a single spin at a time allows us to cut the $O(L^{2D})$ complexity down to an $O(L^D \ln L^D)$ one. A crucial point to be noted here is that this reduction is absolutely intractable with single-spin updates, owing to the very reason that the energy would have to be computed anew after each single-spin update; this requires L^D operations, and an FFT algorithm would output no gain at all.

Let us assume that we can write down the model Hamiltonian as a sum of dot products, i.e., $H = -\frac{1}{2}\Sigma_{i \neq j}J_{ij}\vec{S}(i) \cdot \vec{S}(j)$, with J_{ij} invariant by translation. This is straightforwardly done when $q = 2$, since in this case the dot product reduces to a product of scalar Ising spins. As we will witness in a moment, the presence of a delta Kronecker symbol in the Hamiltonian whenever $q > 2$ requires, however, a minor transformation of the Hamiltonian. For simplicity, we consider hereafter a one-dimensional lattice with an interaction $J(l)$ depending on the distance l between spins. The line argument is similar in higher dimensions, with the sole exception that multidimensional Fourier transforms are then performed. The discrete Fourier transform (DFT) of the spin sequence $\{\vec{S}(l)\}_{l=1\ldots L}$ reads

$$\tilde{\vec{S}}(k) = \sum_{l=0}^{l=L-1} \vec{S}(l)e^{-i2\pi kl/L},$$

and reciprocally,

$$\vec{S}(l) = \frac{1}{L}\sum_{k=0}^{k=L-1} \tilde{\vec{S}}(k)e^{i2\pi kl/L}.$$

Similarly, we define the DFT of the sequence of coupling constants $\{J(l)\}$ as

$$\tilde{J}(k) = \sum_{l=0}^{l=L-1} J_{pbc}(l)e^{-i2\pi kl/L},$$

where $J_{pbc}(l)$ incorporates the effect of infinite image periodic boundary conditions [11], that is, $J_{pbc}(l) = \Sigma_{m=-\infty}^{+\infty}J(l+mL)$; for algebraically decaying interactions, this sum can be exactly expressed in terms of Hurwitz functions [7]. We diagonalize the original Hamiltonian H by rewriting it in terms of the $\tilde{J}(k)$ and $\tilde{\vec{S}}(k)$,

$$H = -\frac{1}{2L}\sum_{k=0}^{k=L-1}\widetilde{J}(k)\widetilde{\vec{S}}(k)\cdot\widetilde{\vec{S}}(-k),$$

where it should be emphasized that $\widetilde{\vec{S}}(-k)$ and $\widetilde{\vec{S}}(k)$ are complex conjugates, since the original vectors $\vec{S}(l)$ have real coordinates. By relying on an FFT radix-2 algorithm, the task of computing the lattice energy is consequently reduced to $O(L\ln L)$ operations.

For $q > 2$, the Kronecker delta symbol in the Hamiltonian unfortunately rules out the previous diagonalization. One way to resolve this issue is to map the q-state Potts model onto a $(q-1)$-dimensional vector model, so that the Kronecker delta function in the original Hamiltonian is turned into a dot product. We define a one-to-one mapping between each Potts spin $\sigma = 1\cdots q$ and a unit-length vector $\vec{S}^{(\sigma)}$ belonging to a $(q-1)$-dimensional hypersphere, so that $\vec{S}^{(\sigma)}\cdot\vec{S}^{(\sigma')} = q\delta_{\sigma,\sigma'} - 1/(q-1)$. It is straightforward to prove that $\sum_\sigma \vec{S}^{(\sigma)} = 0$, and that

$$H = \frac{q-1}{2q}\sum_{i\neq j}J(i-j)\vec{S}^{(\sigma_i)}\cdot\vec{S}^{(\sigma_j)} + \frac{1}{q}\sum_{i<j}J(i-j).$$

In the case of the three-state model, this transformation is equivalent to mapping Potts variables onto the complex plane, i.e., $\sigma \rightarrow S^{(\sigma)} = e^{i2\pi(\sigma-1)/3}$, and writing the dot product $\vec{S}^{(\sigma_i)}\cdot\vec{S}^{(\sigma_j)}$ as $\mathrm{Re}\{S^{(\sigma_i)}S^{(\sigma_j)*}\}$. In this case, the term $\widetilde{\vec{S}}(k)\cdot\widetilde{\vec{S}}(-k)$ becomes $|S(k)|^2$, where $S(k)$ is the DFT of the sequence of (complex) variables $\{S^{(\sigma)}\}$. This reduces by 1 the number of $O(L)$ operations required, since computing a dot product is no longer required.

For $q > 3$, spin vectors on the $(q-1)$-dimensional hypersphere may be determined by using hyperspherical coordinates in $D = q-1$ dimensions, i.e., for the ith vector $\vec{S}^{(i)}$ (with $1 \leq i \leq q$),

$$x_1^{(i)} = \sin\theta_1^{(i)}\sin\theta_2^{(i)}\cdots\sin\theta_{q-3}^{(i)}\sin\theta_{q-2}^{(i)},$$

$$x_2^{(i)} = \sin\theta_1^{(i)}\sin\theta_2^{(i)}\cdots\sin\theta_{q-3}^{(i)}\cos\theta_{q-2}^{(i)},$$

$$x_3^{(i)} = \sin\theta_1^{(i)}\sin\theta_2^{(i)}\cdots\cos\theta_{q-3}^{(i)},$$

$$\cdots$$

$$x_{q-2}^{(i)} = \sin\theta_1^{(i)}\cos\theta_2^{(i)},$$

$$x_{q-1}^{(i)} = \cos\theta_1^{(i)}.$$

We initially set $\theta_i^{(i)} = 0$ for $1 \leq i \leq q-2$, $\theta_j^{(i)} = \alpha_j$ for $j < i \leq q$ and $1 \leq j \leq q-3$, and $\theta_{q-2}^{(q-1)} = -\theta_{q-2}^{(q)} = \alpha_{q-2}$. There remains $q-2$ angles α_j to be determined from $q-2$ equations $\vec{S}^{(i)}\cdot\vec{S}^{(i+1)} = -1/(q-1)$ with $1 \leq i \leq q-2$, from where we obtain $\alpha_1 = \arccos{-1/(q-1)}$, $\cos\alpha_{j+1} = \cos\alpha_j/(1+\cos\alpha_j)$, and thus by induction $\cos\alpha_j = -1/(q-j)$. After a bit of algebra, we find $\vec{S}^{(1)} = (0,\ldots,0,1)$, and

$$\vec{S}^{(i)} = \left(\underbrace{0,\ldots,0}_{q-1-i\text{ terms}},\sqrt{\frac{q(q-i)}{(q-1)(q-i+1)}},\{x_{q-1-i+j}^{(i)}\}_{1<j<i},\frac{-1}{q-1}\right)$$

for $1 < i < q$, where the $(q-1-i+j)$th coordinate reads

$$x_{q-1-i+j}^{(i)} = -\sqrt{\frac{q}{(q-1)(q-1-i+j)(q-i+j)}}.$$

$\vec{S}^{(q)}$ and $\vec{S}^{(q-1)}$ differ only in the sign of their first coordinate x_1. Once these vectors have been computed for a given q, which may be done on startup, determining the lattice energy requires first computing the DFT $\widetilde{S}_j(k)$ of each sequence of coordinates $\{\vec{S}(l)\cdot\vec{S}^{(j)}\}_{l=1,\ldots,L}$, and then evaluating the double sum $\sum_{k=0}^{k=L-1}\sum_{j=1}^{q}\widetilde{J}(k)|\widetilde{S}_j(k)|^2$. As a result, the whole operation is associated with a $O(qL\ln L)$ complexity—or in general $O(qL^D\ln L^D)$, provided the implementation relies on a FFT radix algorithm. As a by-product, it should be noted that once the Fourier components have been computed, it is straightforward to derive the Fourier transform of the spin-spin correlation functions at any inverse temperature β from $\widetilde{g}_\beta(k) = 1/L\langle\sum_{j=1}^{q}|\widetilde{S}_j(k)|^2\rangle_\beta$, where the mean value is obtained from a reweighting procedure. At large lattice sizes, the requirement that L Fourier components be stored at each MCS may constitute a significant challenge in terms of computer memory; in this case, a practical workaround consists in computing microcanonical averages for each energy level visited during the simulation, and then to perform the reweighting procedure directly from these microcanonical averages. In the case of long-range interactions, careful attention must be paid, however, to the influence of the discretization of the energy axis in terms of systematic error.

B. Efficient cluster construction for long-range interactions decaying with the distance

For long-range spin models, building a new cluster at each MCS takes of order L^{2D} operations, since $L^D(L^D-1)/2$ pairs of spin are considered in turn for bond activation. When interactions decay with distance, the probability of adding a bond between two spins falls off quite rapidly as the distance between them increases. A significant amount of time during the construction of the cluster is thus wasted because an overwhelming number of bonds are considered for activation which have only a negligible probability to be activated. Even in the case of interactions decaying as $1/|i-j|^{1+\sigma}$ with σ close to 0, does the bond count never exceed a few percent of the whole number of available bonds. In this respect, switching from a local- to a global-update scheme might well be an ill-fated choice as the gain in terms of autocorrelation time is spoiled by the exceedingly time consuming construction of the cluster. However, an efficient construction method was proposed by Luijten and Blöte in the recent past [12], with an efficiency that is independent of the number of interactions per spin, and a CPU demand that scales roughly as L^D. The rationale behind this method is to use cumulative probabilities, whereby instead of considering each spin in turn for addition to a given cluster, it is the index of the next spin to be added which is drawn at random. We

now give a sketchy outline of the method in the context of long-range chains. Extensive details may otherwise be found in Refs. [12,58]. First of all, the probability to add a bond is split up into two parts, namely, (i) a provisional probability $\pi_l(E)$ (hereafter simply denoted π_l) depending on the distance $l=|i-j|$ between spins and on the lattice energy E, and (ii) a factor $f(\sigma_i, \sigma_j)$ controlled by the spin values, e.g., a Kronecker delta symbol in the case of a Potts model. If 0 denotes the index of the current spin to which we are adding bonds (i.e., spin indices are considered to be relative to the current spin), then the provisional probability of skipping k −1 spins and bonding the current spin with a spin at position $k>0$ is given by $P_0(k)=\Pi_{m=1}^{k-1}(1-\pi_m)\pi_k$. From there, one builds a table of cumulative probabilities $C_0(j_1)=\Sigma_{k=1}^{j_1}P_0(k)$ for all $j_1>0$, so that the index j_1 of the spin to be bound with current spin 0 is obtained by first drawing a random number $0<r<1$ and then reading out j_1 from the table, i.e., j_1 is such that $C_0(j_1-1)<r<C_0(j_1)$. Standard binary-search algorithms may be used for this purpose. Last, a bond is activated between spins 0 and j_1 with a probability $f(\sigma_0, \sigma_{j_1})$, and we proceed further with the computation of the index $j_2>j_1$ of the next spin to be bound with current spin 0. The corresponding provisional probability thus becomes $P_{j_1}(k)=\Pi_{m=j_1+1}^{k-1}(1-\pi_m)\pi_k$, and the cumulative probabilities read $C_{j_1}(j_2)=\Sigma_{k=j_1+1}^{j_2}P_{j_1}(k)$. The same procedure is repeated for $\{j_3, j_4, \ldots\}$ until we draw a $j_\alpha>L$, in which case we jump to the next current spin, which in a one-dimensional model is the nearest-neighbor of the previous current spin. In addition, there are two formulas which make it easier to compute cumulative probabilities: first, one can show that $C_0(j)=1-\exp[-\beta(E)\Sigma_{k=1}^{j}J(k)]$, where E is the energy of the current configuration, and second, the cumulative probabilities $C_{j_\alpha}(j_{\alpha+1})$ can be straightforwardly derived from the $C_0(j)$ coefficients through the relation $C_{j_\alpha}(j_{\alpha+1})=[C_0(j_{\alpha+1})-C_0(j_\alpha)]/[1-C_0(j_\alpha)]$. It follows from the last relation that, instead of building a look-up table for each $C_{j_\alpha}(j_{\alpha+1})$, we may as well draw a random number $0<r<1$, transform it to $r'=r[1-C_0(j_\alpha)]+C_0(j_\alpha)$, and choose the next spin to be added from the relation $C_0(j_{\alpha+1}-1)<r'<C_0(j_{\alpha+1})$. In practice, we thus simply need to compute a single look-up table filled with $\Sigma_{k=1}^{j}J(k)$ for each j at the beginning of the simulation, from where we will derive the $C_0(j)$ coefficients at each new MCS corresponding to a lattice configuration with a given energy E. This last task requires of order L^D operations. To sum up, the construction of each cluster thus consists in choosing a "current" spin amongst $L-1$ possible spins in turn, e.g., starting from the leftmost one, and then activating bonds between the current spin and other spins located to its right by drawing a random number, scaling it, and selecting the bond indices from a look-up table containing the $C_0(j)$ coefficients at energy E. Once each spin has been considered as a current spin, a cluster multiple labeling technique can eventually be used to identify every set of spins actually belonging to the same cluster [57].

IV. NUMERICAL TESTS OF ALGORITHM PERFORMANCE

In this section, we address the performance of our algorithm in terms of dynamical behavior. Since our work fo-

cuses mainly on long-range spin models, we decided to perform intensive numerical tests on the one-dimensional q-state Potts chain with LR interactions $1/|i-j|^{1+\sigma}$ decaying as a power law of the distance between spins. The rich phase diagram of this model, and the fact that several numerical studies have been carried out on this model in the recent past, makes it a perfect test case. For the sake of comparison with other numerical methods, and in order to ensure that our algorithm did not produce systematic errors, we also performed several tests on the two-dimensional model with nearest-neighbor (NN) interactions, for which exact results are known (see Refs. [59,60]; also references in Ref. [61]). Both models are known to exhibit a first-order transition for an appropriate set of parameters, namely, $q>4$ for the NN model [62], and $\sigma<\sigma_c(q)$ for the LR one, with for instance $\sigma(3)=0.72(1)$ [7]. We chose a set of parameters that would allow us to observe both weak and strong first-order transitions, and concentrated on several indicators of performance, reliability, and scalability: these include tunneling, equilibrium, and effective autocorrelation times, and mean acceptance rates. These indicators inform us on the efficiency with which the Markov chain reaches the equilibrium distribution and explores the phase space. They also tell us at what rate successive measurements decorrelate from each other; hence what amount of resources is needed to obtain reliable statistics. Overall, they are therefore good indicators as to whether CPU resources are efficiently utilized or not. As regards scalability, we also computed the dynamical exponents associated with tunneling and equilibrium times; these indicate how fast needs in CPU time grow with the lattice size.

All densities of states were calculated by means of the Wang-Landau algorithm, whereby, starting from an initial guess of the density of states $n(E)$, we update $n(E)$ after each visit to energy level E according to the rule $\ln n(E) \leftarrow \ln n(E)+\ln f$, where $\ln f$ is hereafter termed *Wang-Landau modification factor*. In the case of LR models, the unequal spacing of energy levels and the existence of energy gaps in the vicinity of the ground state required that we introduced a few changes over the original version. In particular, using an interpolator for $\ln n(E)$ turned out to be mandatory in order to compensate for the finite width of histogram bins—as would also be required for models having a continuous symmetry; indeed, we observed that using large bins tends to strongly reduce the acceptance rate if no interpolator is used. Bezier splines provide good interpolators, although a linear interpolation with a slope given by the microcanonical temperature $\beta(E)$ also proved to be particularly efficient whenever this last quantity was made available by other means, e.g., the transition matrix.

For small and medium lattice sizes, we systematically performed all simulations twice, first with standard single-spin updates (SSUs) and then with our method embedding cluster updates (CUs). We give an estimate of the error in the density of states obtained from both types of update schemes. For the largest lattice sizes we studied, however, the SSU implementation simply turned out to be impracticable, due to either exceedingly high tunneling times, and—for LR models—excessive CPU demands, and we present results for the CU algorithm only.

A. Phase space exploration and mean acceptance rates

As opposed to the (canonical) Swendsen-Wang cluster algorithm, the acceptance rate of our algorithm—Eq. (1)—is not trivially equal to unity. Still, it is tightly related to the efficiency with which the Markov chain wanders about the phase space, since a low acceptance rate would lead to very repetitive dynamics. In this view, it is instructive to compute an approximate analytical expression of this acceptance rate when the initial and the final energies E_a and E_b differ only by a small amount. Writing $E_b = E_a + \epsilon$, and carrying out a series development to first order in ϵ, one obtains $A_{\text{flip}} = \min[1, 1 + \Delta(E_a)dE]$, where

$$\Delta(E_a) = \beta'(E_a)\left(\sum_{b_{ij} \in \mathcal{B}} J_{ij}\frac{1 + p_{ij}(E_a)}{p_{ij}(E_a)} - |E_a|\right),$$

with the same notation as in Sec. II. We wish to obtain an estimate of the first statistical moments of $\Delta(E)$. We hereafter consider the case of a model with nearest-neighbor interactions ($J=1$), for which we can carry out an exact derivation. The last expression simplifies to

$$\Delta(E_a) = \beta'(E_a)\left(B\frac{1 + p}{p} - |E_a|\right),$$

where B stands for the total number of bonds and $p = p(E_a) = e^{\beta(E_a)} - 1$. At a given energy E_a, at most $|E_a|$ bonds may be activated. The probability to activate a bond is $\pi(E_a) = p(E_a)/[1 + p(E_a)] = 1 - e^{-\beta(E_a)}$; hence the probability to create a bond set $\mathcal{B}$ containing B bonds writes $P(B) = \binom{|E_a|}{B}[\pi(E_a)]^B[1 - \pi(E_a)]^{|E_a| - B}$, which may be reexpressed as

$$P(B) = \binom{|E_a|}{B}\frac{p^B}{(1 + p)^{|E_a|}}.$$

From there, we can derive the average bond count, $\langle B \rangle = |E_a|[p/(1 + p)]$. This allows us to rewrite $\Delta(E_a)$ as

$$\Delta(E_a) = \beta'(E_a)\frac{1 + p}{p}(B - \langle B \rangle);$$

hence $\langle \Delta(E_a) \rangle = 0$. The variance of $\Delta(E_a)$ is thus proportional to the variance of the bond count distribution, i.e., $\langle B^2 \rangle - \langle B \rangle^2 = |E_a|[p/(1 + p)^2]$, which yields

$$\sqrt{\langle \Delta(E_a)^2 \rangle} = \delta\Delta(E_a) = |\beta'(E_a)|\sqrt{\frac{|E_a|}{\exp\beta(E_a) - 1}}.$$

For a given $\epsilon > 0$, one half of all attempted cluster flips thus leads to an acceptance rate which is lower than 1, the other half saturating at unity. Assuming a Gaussian distribution for $\Delta(E)$, with the standard deviation computed above (which is valid for large enough lattice sizes), the mean acceptance rate is readily obtained from the mean value of a Gaussian distribution centered at unity and truncated above 1, which yields
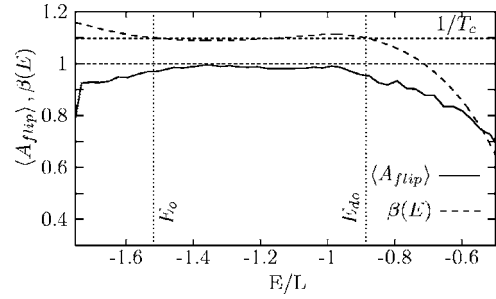

FIG. 1. Mean acceptance rate as a function of the energy per spin for the six-state long-range Potts with $\sigma = 0.7$, and $L = 1024$ spins. The dashed line shows the estimated inverse microcanonical temperature. The vertical dotted lines indicate the position of the histogram peaks corresponding to the ordered and disordered phases.

$$\langle A_{\text{flip}} \rangle(E_a) = 1 - \frac{\delta\Delta(E_a)}{2\sqrt{2\pi}}\epsilon.$$

In the case of interactions depending on the distance l between spins, one may observe that the average energy is related to the average number of bonds of length l by $-\langle E \rangle = \Sigma_{l>0}J(l)[(1 + p_l)/p_l]\langle B(l)\rangle$, which shows that $\langle \Delta(E) \rangle = 0$ also in this case.

At a first-order transition, $\beta(E)$ varies smoothly between the energy peaks of the ordered and disordered phases, which ensures that $\Delta(E)$ remains small. The mean acceptance rate for the six-state LR Potts chain with $\sigma = 0.7$ and $L = 1024$ spins is sketched in Fig. 1. While the acceptance rate is close to 1 inside the range of phase coexistence, the variance of $\Delta(E)$ increases when E lies outside the range of phase coexistence, and therefore leads to a reduction in the acceptance rate. We observe that this diminution is less marked at low-energy levels, for the energy cost associated with flipping a small number of big clusters is lower than that associated with randomly updating a great deal of small clusters, and $E_b - E_a$ is consequently lower in the last case. It is worth underlining, however, that the energy range of interest in the analysis of first-order phase transitions spans an interval which is only moderately larger than that corresponding to phase coexistence, the only requirement being that metastability plateaus [7] and histogram peaks must be clearly visible. As a result, the fact that the mean acceptance rate for cluster flips remains well above 90% inside this range of energy represents already an improvement of a factor 3 with respect to the standard multicanonical approach, where we obtained acceptance rates oscillating around 30%.

B. Dynamic properties

Where performance measurements at first-order transitions are concerned, tunneling times have thus far been regarded as one of the most meaningful measurement parameters [32,61,63]. They are defined as one half of the average number of MCSs needed for the walk to travel from one peak of the energy histogram to the other—where peaks are defined with respect to the finite-size transition temperature—and turn out to represent a fairly good indica-
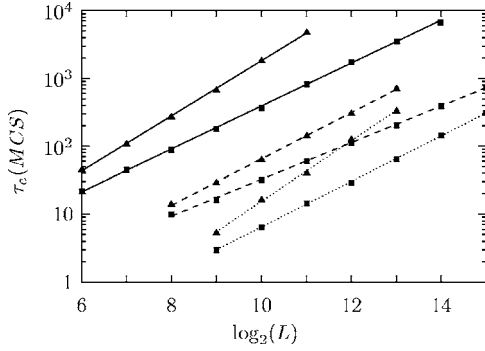
FIG. 2. Tunneling times for the long-range Potts chain with $q=3$, $\sigma=0.4$ (dashed lines) and 0.6 (dotted lines), and $q=6$, $\sigma=0.7$ (solid lines). Triangles refer to the SSU implementation, while squares indicate estimates for our method (CU). Dynamic exponents z were determined from a fit to the power law $\tau_e \sim L^z$.

tor of the interval between roughly independent samples.

Results for the LR chain with $q=3$ and 6 are shown in Fig. 2. Dynamic exponents z were determined from a fit to the power law $\tau_e \sim L^z$, and are summarized in Table I. We can witness a substantial reduction for both the LR and the NN models, with exponents close to and sometimes even below the ideal random-walk value $z=D$. As regards the NN model, our values compare extremely well with those obtained with the multibond method [32] and with Rummukainen's hybridlike two-step algorithm [39], although these approaches and ours differ markedly in the way clusters are constructed.

It should be mentioned that the distance $E_d - E_o$ the random walker must travel, i.e., the energy gap between the peaks of the histogram, does not scale linearly with the number of spins. This feature is especially apparent for long-range interactions, where $E_d - E_o$ grows all the more faster with increasing lattice size that σ comes closer to 0. As a result, the power law $\tau_e \sim L^z$ yields dynamical exponents which are underestimated with respect to the value given by a power law of the form $\tau_e \sim (E_d - E_o)^z$ (up to a dimensional factor 2 for the NN model). For instance, we would obtain $z=1.40(3)$ instead of $z=1.35(3)$ for $q=6$ and $\sigma=0.7$, and $z=1.10(1)$ instead of $z=1.05(1)$. Where the performance in

TABLE I. Dynamic exponents z for the q-state Potts chain. z(SSU) and z(CU) refer to single-spin and cluster updates, respectively, while z_{muBo} and z_{muClus} make reference to the multibond method [32] and Rummukainen's multimicrocanonical cluster method [39] applied to the NN model.

q	σ	z(SSU)	z(CU)	z_{muBo}	z_{muClus}
6[a]	0.7	1.35(3)	1.05(1)		
3[a]	0.6	1.48(2)	1.11(1)		
3[a]	0.4	1.13(2)	0.89(1)		
7[b]		2.60(4)	1.82(2)	1.84	1.82(3)
10[b]		2.87(4)	2.23(1)	2.1	

[a]Power-law decaying interactions.
[b]Two-dimensional counterpart with nearest-neighbor interactions.

terms of CPU demands is concerned (and in particular if one is interested in how it grows with the size of the system), we think, however, that the traditional definition $\tau_e \sim L^z$ is more meaningful.

While tunneling times represent a practical way to estimate the efficiency with which the random walker drifts along the energy landscape, they are subject to two limitations. First, they cannot be properly defined in the case of second-order phase transitions, since the histogram of the energy no longer exhibits two peaks. Second, there is no direct connection between tunneling times and autocorrelation times, which makes it difficult to estimate the optimum interval between measurements that will yield perfectly uncorrelated data, and thus minimum statistical error in estimates of thermodynamic data. It is worth mentioning here that computing integrated autocorrelation times naively from the set of measurements, i.e., just as is usually done in the canonical case, simply makes no sense when simulating in the multicanonical ensemble, because the quantities we are interested in are, in the first place, reweighted averages of thermodynamical data [61].

Therefore alternate definitions have been proposed, which try to circumvent these limitations. One approach is to compute the so-called *round-trip times* [64], which are computed from the number of MCS needed to get across the whole energy axis, that is, from the ground state to the upper energy level. Although round-trip times may be determined for any order of phase transition, they present unfortunately no more connection with statistical errors than do tunneling times. On the contrary, multicanonical *effective autocorrelation times*, which were first introduced in the framework of the multibond algorithm [32], offer a direct comparison with exponential or integrated autocorrelation times of traditional use in canonical simulations. Mimicking the canonical case, the effective autocorrelation time τ_{eff} can be defined for any thermodynamic variable θ by inverting the standard error formula $\epsilon_\theta^2 = \sigma_\theta^2 2\tau_{eff}/N$, where N stands for the total number of (possibly correlated) measurements, σ_θ^2 denotes the variance of the (reweighted) thermodynamic variable θ, e.g., $\langle E^2 \rangle - \langle E \rangle^2$, and ϵ_θ^2 is the squared statistical error in the same variable. The error may be estimated either from resampling or (jackknife) blocking procedures, or by performing multiple independent runs. Since both the variance and the error depend on the reweighting temperature, the previous definition obviously yields an effective autocorrelation time which also depends on the temperature.

We now discuss our results for effective autocorrelation times obtained for the six-state LR Potts chain with $\sigma=0.9$ and $128 \leq L \leq 1024$ spins. For this value of σ, the model exhibits a very weak first-order transition with no clearly visible histogram peaks for sizes below $L \sim 2000$. The choice of medium lattice sizes was dictated by the fact that we computed the error from multiple independent runs (around 20 runs of 10^6 MCSs each), which we found a more reliable way of computing the statistical error than using a blocking procedure. Figure 3 shows the dependence of τ_{eff} on the temperature for $L=512$. For both algorithms, τ_{eff} exhibits a peak in the vicinity of the effective transition temperature $T_c(C_v) = 0.7163(2)$. As expected, the reduction brought about by

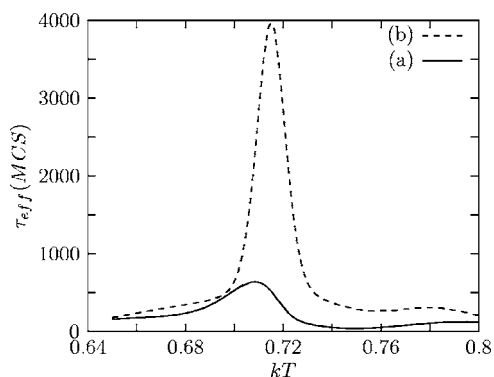FIG. 3. Effective autocorrelation time τ_{eff} for $q=6$, $\sigma=0.9$, and $L=512$ with (a) cluster and (b) single-spin updates. The effective transition temperature defined from the peak of the specific heat is $T_c(C_v)=0.7163(2)$.
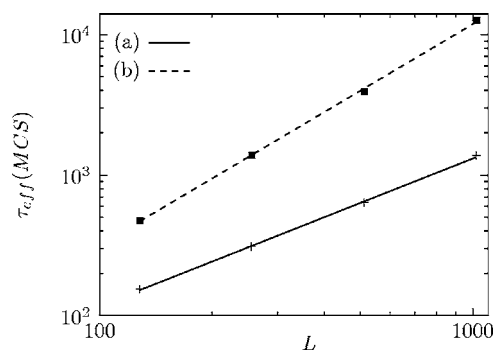

FIG. 4. Fit of effective autocorrelation times τ_{eff} to the power law $\tau_{eff} \propto L^\alpha$ for the six-state Potts chain ($\sigma=0.9$ and $L=512$) with (a) cluster updates and (b) single-spin updates.

cluster updates in terms of correlation between measurements is marked, especially in the transition region, where single-spin update leads to a critical slowing down similar to the one encountered in canonical simulations. This behavior is consistent with the very general observation reported recently in Ref. [30] in the framework of the optimal ensemble, and also in Ref. [65] in the context of equilibration time for multicanonical algorithms (see also the next paragraph for more details on this issue), whereby the random walker diffuses at a slower pace in the critical region. In this respect, cluster updates optimize the diffusive current of the random walker in the critical region in much the same way as do the optimal ensemble weighting proposed in Ref. [30], yet with a distinct strategy: in the latter, the error is reduced by allowing the walker to spend more time in the critical region than in the rest of the energy axis; in our approach, it is the decorrelating capability of the move update itself which reduces the statistical error in the transition region. As is well known, however, cluster updates are especially efficient at the percolating threshold, and the reduction in terms of correlation is large *because* bond probabilities are governed by the microcanonical temperature. This interpretation is clearly underpinned by our investigation of the effect of poor estimates of $\beta(E)$ on tunneling times, presented later in Sec. VII. Finally, we focus on the scaling behavior of autocorrelation times. Table II reports our results for L ranging from 128 to 1024 spins, where τ_{eff} is evaluated at the effective transition temperature determined from the peak of the specific heat. Our method gives smaller autocorrelation times

TABLE II. Effective autocorrelation times at the transition temperature defined from the location of the peak of the specific heat, for the six-state LR Potts chain with $\sigma=0.9$.

L	τ_{eff}(SSU)	τ_{eff}(CU)
128	475	155
256	1390	310
512	3960	635
1024	12700	1370
z	1.6(1)	1.0(1)

already for $L=128$ spins. From these values, we also determined the associated scaling exponents by a fit to the power law $\tau_{eff} \propto L^z$ (Fig. 4), and obtained a highly satisfying value of $z \sim 1.0(1)$ with cluster updates.

We conclude the discussion on the dynamic characteristics of our algorithm with an investigation of equilibrating properties. As opposed to canonical simulation, estimating equilibrium times has been much less common in the context of multicanonical simulations; the nonlinear relaxation function, while very informative when the equilibrium distribution is driven by a Boltzman weight [66], is of limited use indeed if the engendered distribution is flat. Recently, however, an efficient procedure aimed at estimating equilibrium times for any equilibrium distribution was proposed by Guerra and Muñoz [65]. This procedure relies on a χ^2 regression with respect to the (expected) flat equilibrium distribution $\mathcal{P}(E)$. Starting from the same initial lattice configuration, n Markov processes are run with distinct random seeds, and at each MC step t, a histogram of the energy $V_t(E)$ is filled with the value of the energy of each process. Asymptotically, $V_t(E)$ should approximate the expected flat distribution $\mathcal{P}(E) \propto n(E)w(E)$. In order to estimate the equilibrium time in a more quantitative way, a $\chi^2(t)$ deviation of $V_t(E)$ with respect to the flat distribution is carried out at each MC step t, i.e., $\chi^2(t)=\Sigma_E[V_t(E)-n\mathcal{P}(E)]^2/[n\mathcal{P}(E)]$, where the sum runs over histogram bins. For large n, and provided equilibrium has been reached, the distribution of $\chi^2(t)$ over m experiments obeys a χ^2 law with a number r of degrees of freedom given by the number of histogram bins minus 1, that is, with a mean equal to r and a standard deviation given by $\sqrt{2r/m}$. Due to the intensive demand in CPU required by this procedure, we restricted our estimation of equilibrium times to the single case $q=6$ and $\sigma=0.9$. We performed $n=1000$ Markov processes for sizes between $L=128$ and $L=512$, and estimated the equilibrium time from a single experiment (that is, $m=1$) by simply monitoring the time needed for $\chi^2(t)/r$ to reach unity and then stay within the interval $[1-2\sigma/r, 1+2\sigma/r]$. As illustrated in Fig. 5, relying on a single experiment leads to quite large error bars, yet this is sufficient for our purpose. From the graphs of $\chi^2(t)$ we read $\tau_{eq}=4500\pm500$ MCSs and $\tau_{eq}=23000\pm2000$ MCSs for the cluster- and single-spin updates, respectively; in spite of the large uncertainty, the reduction in terms of equilibrium time
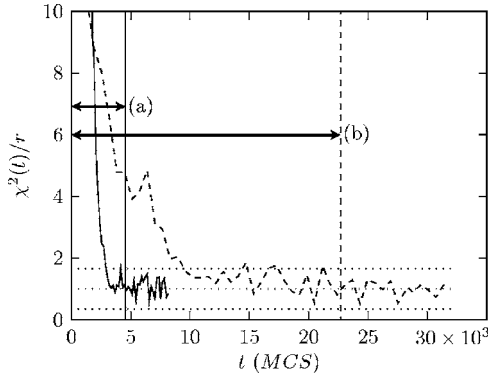
FIG. 5. Plot of $\chi^2(t)/r$ for the six-state Potts chain ($\sigma=0.9$, $L=512$) using (a) cluster updates and (b) single-spin updates. The regression was carried out over a histogram containing 20 bins populated from 1000 runs, all starting in ground state configuration but with distinct random seeds.

brought about by our method is clearly visible. Results for other lattice sizes are summarized in Table III. A fit to the power law $\tau_{eq} \propto L_{eq}^z$ (see Fig. 6) yields the scaling exponents $z_{eq}=1.96(5)$ and $z_{eq}=1.21(3)$ for the single-spin and the cluster updates, respectively. Here again, we think that lower diffusion currents in the latest case account for the higher pace at which the random walker reaches the equilibrium distribution.

C. Overall CPU demand for LR models

We now discuss CPU demand in the case of LR models, and concentrate on the gain in CPU resources brought about by the optimization schemes proposed in Sec. III. Assuming a decently efficient algorithm implementation, this indicator yields a rough account of the real algorithm complexity, although it should be mentioned that it is usually an elaborate task to estimate this quantity rigorously, partly because its value hinges heavily on a variety of implementation, CPU architecture, and compiler dependent properties. We decided to measure CPU times over a series of 1-h-long simulation runs on a handful of distinct CPU architectures, including Intel Pentium and Xeon at 2.4 and 3.2 GHz. Figure 7 sketches averages of the CPU (user) time per MCS and per spin, where small fluctuations might be attributed to the effect of varying CPU cache sizes amidst our clutch of CPU's. While for the local-update implementation the demand in CPU per spin grows linearly with the number of spin, it is roughly constant over a fairly large range of lattice sizes in

TABLE III. Equilibrium times for the six-state LR Potts chain with $\sigma=0.9$ obtained by monitoring the graph of $\chi^2(t)/r$.

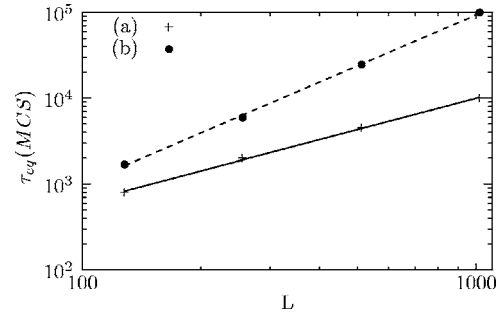| L	τ_{eq}(SSU)	τ_{eq}(CU)
128	1700 (100)	800 (120)
256	6000 (750)	2000 (200)
512	23000 (2000)	4500 (500)
1024	101000 (8000)	10000 (800)

FIG. 6. Fit of equilibrium times to the power law $\tau_{eq} \propto L^\alpha$ for the six-state Potts chain ($\sigma=0.9$): (a) updates and (b) single-spin updates.

the case of our cluster-update algorithm. Moreover, our method already outperforms the local-update scheme starting from several hundred spins, with nonetheless an increased footprint for higher q values which is accounted for by the correspondingly higher number of FFT's to be computed. This, however, clearly demonstrates the breakthrough that our method brings about for the study of long-range models, paving the way for precise tests of finite-size scaling.

V. TWO-DIMENSIONAL NN POTTS MODEL: COMPARISON WITH EXACT RESULTS

In order to check that our algorithm did not produce systematic errors, we computed transition temperatures and interface tensions for the two-dimensional q-state Potts model ($q=7$, 10) with nearest-neighbor (NN) interactions and helical boundary conditions. Results regarding the dynamic characteristics of our algorithm for this model were reported in Sec. IV; we will concentrate here on precision matters. For $q=10$, we obtained $T_c(L)=0.706\,99(5)$, $0.704\,91(5)$, $0.703\,00(2)$, $0.702\,78(1)$, $0.701\,64(1)$, $0.701\,328(4)$, and $0.701\,249(2)$ for $L=16$, 20, 30, 32, 64, 128, and 256, where T_c was determined from the location of peaks of the specific heat. C_v was computed directly from the estimated density of
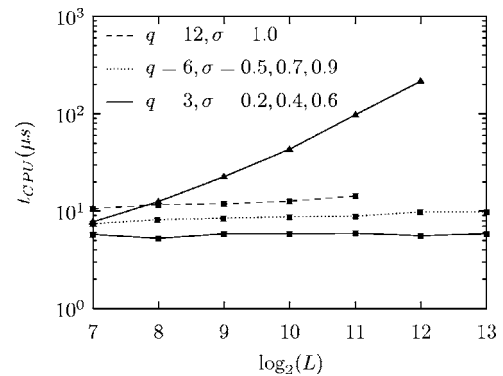

FIG. 7. CPU time per MCS and per spin for the long-range Potts chain. Triangles indicate typical CPU times for the local-update algorithm (SSU), irrespective of q and σ. Squares refer to our algorithm (CU) with LR specific optimizations included; for $q=3$ and $q=6$, estimates were determined by averaging over the indicated σ values.

states, and then refined from an additional production run of length 10^7 MCS. The error was estimated by means of the jackknife method. Following standard finite-size scaling theory at first-order phase transitions, we collapsed $C_v(T)/L^2$ vs $(T-T_c)L^2$ over the five highest lattice sizes and found an infinite size temperature $T_c(\infty)=0.701\ 236(3)$ in very good agreement with the exact value $0.701\ 231\ 57\ldots$. The same procedure applied to $q=7$ and $L=64$, 128 and 256 yielded $T_c(\infty)=0.773\ 059(1)$ which again matches perfectly the exact value $0.773\ 058\ 9\ldots$. We estimated the interface tension (i.e., per unit length of the interface) f^s between the ordered and disordered phases from the histogram of the energy, reweighted at a temperature where energy peaks have the same height [67]: $2f^s=-L^{-1}\ln P_{min}$, where the factor 2 arises from the use of periodic boundary conditions. Here, P_{min} denotes the minimum of the histogram between the two energy peaks; peak heights are normalized to unity. We computed f^s directly from the density of states, and estimated the error from the additional production run. In this respect, it should be noted that estimating interface tensions directly from the density of states generally yields values that lie below those computed from histograms collected during production runs. Our algorithm allowed us to determine f^s with a four-digit precision for sizes up to $L=256$ and nonetheless rather modest statistics. For the seven-state model, we obtained $2f^s=0.0336(6)$, 0.0294(1), 0.026 31(8), and 0.023 84(9) for $L=32$, 64, 128, and 256, a linear fit of the form $\Sigma\sim\Sigma(\infty)+c/L$ [68] performed over the three largest sizes (i.e., for L above the disordered phase correlation length $\xi\sim48$ [60]) yielded the infinite size value 0.022 30(11), still above the exact value 0.0207 92, yet closer to it than estimates reported in several previous studies [32,39,61].

VI. LR POTTS CHAIN: ERROR ESTIMATES IN THERMODYNAMICAL DATA

In this section, we discuss the precision of our results for the q-state Potts chain with algebraically decaying interactions, i.e., $J(r)=1/r^{1+\sigma}$. Our purpose is twofold. First, we estimate the error in the density of states $n(E)$ obtained from the Wang-Landau algorithm, so that we can obtain a better insight into the benefit of our method with regards to the iterative calculation of $n(E)$. Second, we determine confidence intervals on reweighted averages computed from an additional production run. Since computing thermodynamic quantities from a production run does not require that the histogram be perfectly flat, nor that the estimate of the density of states be perfectly accurate, this reduces to estimating the gain in precision brought about by lower autocorrelation times.

A. Statistical error of the density of states

In order to compare the error in the density of states produced by the single-spin update implementation and our method, we performed for each method a series of 12 independent simulations with the Wang-Landau algorithm, all starting with the same initial guess of the density of states. The model parameters were set to $q=6$, $\sigma=0.9$, and $L=512$.
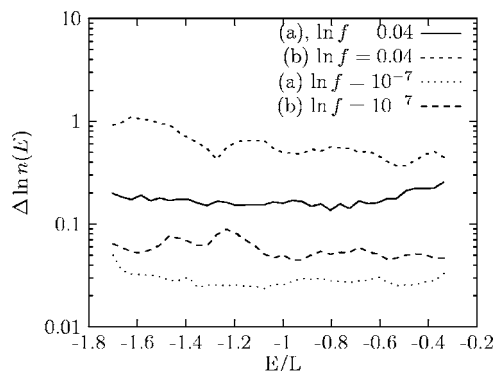

FIG. 8. Statistical error in the density of states of the six-state Potts chain for two distinct modification factors $\ln f$ of the Wang-Landau algorithm. The statistical errors were obtained from 12 independent runs. The parameters of the model are $\sigma=0.9$ and $L=512$. (a) and (b) correspond to our method and the local-update algorithm, respectively.

This choice of parameters guarantees that, in spite of the modest lattice size we consider, autocorrelation times differ by a sufficient amount between the single-spin updates method and our method, so that the benefit may be clearly interpreted in terms of decorrelating capabilities. The initial guess of $S(E)$ was scaled up from an estimate obtained at $L=256$, and the updating factor of the Wang-Landau algorithm was initially set to $\ln f=5$. We did not make use of all improvements to the original Wang-Landau algorithm, as proposed by Zhou and Bhatt in Ref. [34], since these would have partly overshadowed the gain produced solely by lower autocorrelation times. Indeed, we mainly focused on the systematic error (rather than the whole statistical error) that may show up during the first iterations. It was shown in Ref. [34] that this systematic error results from the combination of a large $\ln f$ coefficient with the presence of strong correlations between adjacent binning. We thus simply relied on the original histogram flatness criterion to switch from one iteration to another, and divided $\ln f$ by the same amount (namely, 5) after each iteration which passed the flatness check. We found out, however, that using the criterion in Ref. [34] instead, that is, averaging $n(E)$ on multiple independent runs after each iteration, and switching to the next iteration only after a given number of entries was recorded in the histogram [see Eq. (12) in Ref. [34]], led to markedly lower statistical errors. As illustrated in Fig. 8, the statistical error in the density of states is clearly improved by our method. In particular, cluster updates lead to a spread of the error over the whole energy axis. In this respect, and as already mentioned in Sec. IV, the lower diffusion rates associated with collective updates in the critical energy region offer a clear benefit. As expected from the arguments of Zhou and Bhatt, the reduction is also more marked for $\ln f=0.04$ than for $\ln f=10^{-7}$, and the systematic error brought about by correlations between successive binning is indeed partly tamed by a lower Wang-Landau modification factor. Finally, we show in Fig. 9 the resulting statistical error in the specific heat, since thermodynamical averages are the relevant quantities in the first place. C_v was computed directly from the estimated density of states $n(E)$, i.e., according to
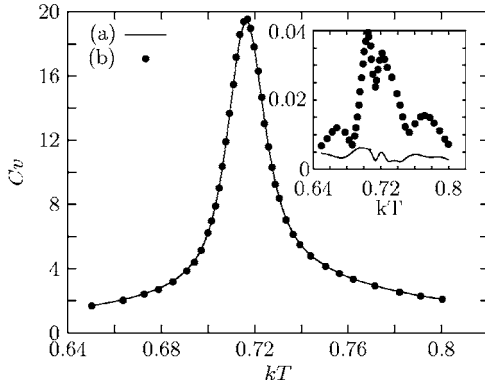
FIG. 9. Graph of the specific heat for the six-state Potts chain ($\sigma=0.9$ and $L=512$) obtained directly the final estimate of the density of states with (a) our method and (b) the local-update algorithm. The inset shows the relative error $\mathrm{err}(C_v)/C_v$.

the formula $C_v(kT)=(\langle E^2\rangle_{kT}-\langle E\rangle^2_{kT})/(L\,kT^2)$, where $\langle E^n\rangle = [\Sigma_E E^n n(E)e^{-E/kT}]/[\Sigma_E n(E)e^{-E/kT}]$. For long-range models, energy levels are not equally spaced, and it should be noted that too large histogram bins may cause a systematic deviation of the averages as well. We paid attention to this by comparing our results for several bin widths, and made sure that the systematic deviation engendered was always lower than the statistical error itself. As shown in the inset of Fig. 9, the accuracy of the estimation of C_v is larger by nearly an order of magnitude at the transition temperature. Incidentally, we observe that this is comparable to the gain in terms of autocorrelation times, as already presented in Fig. 3.

B. Transition temperatures, Binder cumulants, and interface free energies

We now discuss some of our results for the three-state Potts chain, for which we performed extensive simulations for sizes ranging from $L=128$ to $L=65536$. As opposed to higher values of q, there exists indeed a large collection of numerical studies for $q=3$, so that comparison with previous estimates is easier. Table IV reports our values for transition temperatures and peaks of response functions for $\sigma=0.5$. Both the specific heat C_v and the susceptibility χ were computed from production runs whose length varied between 10^6 and 10^8 MCSs depending on the lattice size, and error bars were computed by means of the jackknife blocking method. We performed these production runs twice, first using single-spin updates, and then using our method, yet in both cases with the same estimate of the density of states. Figure 10 shows the graph of C_v as obtained with cluster updates. We mention that, for $L>4096$, the local-update implementation was simply intractable as a result of excessive computation times. For all sizes, our results match within error bars for both methods, and it should be noticed that, for the two largest sizes, we obtain estimates of finite-size transition temperatures accurate up to the fifth digit.

In order to determine infinite-size transition temperatures, we performed a fit of the finite-size temperatures $T_c(L)$ reported in Table IV, to the power law $T_c(L)-T_c(\infty)=a/L$. This is illustrated in Fig. 11 where for the sake of clarity, only

TABLE IV. Estimates of peaks of the specific heat C_v and the susceptibility χ, and corresponding effective transition temperatures for the three-state LR Potts chain with $\sigma=0.5$. Error calculations were carried out by means of the jackknife method applied to a single production run. The number of MCSs per production run is the same for both methods, yet varies between 10^6 and 10^7 from the smaller to the larger lattice sizes.

| | $T_c(C_v)$ | | C_v^{max} | |
L	(CU)	(SSU)	(CU)	(SSU)
128	1.6450(18)	1.645(3)	3.55(2)	3.55(3)
256	1.6607(2)	1.6607(13)	4.86(2)	4.88(5)
512	1.6741(9)	1.675(1)	6.54(3)	6.47(6)
1024	1.6815(2)	1.6815(17)	9.14(8)	9.10(24)
2048	1.6856(3)	1.685(1)	13.63(15)	13.73(80)
4096	1.68742(9)	1.6875(10)	22.21(34)	21.9(2.2)
8192	1.68801(7)		40.28(44)	
16384	1.688031(34)		79.46(35)	
32768	1.687851(12)		164.1(4)	
65536	1.687749(09)		332.8(6)	

| | $T_c(\chi)$ | | χ^{max} | |
L	(CU)	(SSU)	(CU)	(SSU)
128	1.6793(14)	1.679(4)	3.44(3)	3.46(3)
256	1.6837(2)	1.6837(15)	6.09(3)	6.13(5)
512	1.6864(8)	1.6877(15)	10.81(7)	10.86(13)
1024	1.6882(3)	1.688(2)	19.73(28)	19.8(5)
2048	1.6887(2)	1.6882(15)	37.6(5)	37.5(1.8)
4096	1.68869(9)	1.6887(11)	75.4(1.2)	74.6(6.7)
8192	1.68842(7)		165.2(1.8)	
16384	1.688148(35)		369.8(1.1)	
32768	1.687870(20)		827(2)	
65536	1.687773(16)		1754(3)	

data obtained using cluster updates are reported. An important effect we noticed is the presence of a crossover around $L=16384$ for $T_c(C_v)$, where the finite-size transition temperature reaches a maximum; a similar behavior can be witnessed for $T_c(\chi)$ with a maximum occuring around $L=4096$. In view of this, we first performed a fit over lattice sizes ranging from $L=256$ to $L=4096$: for this range of lattice sizes, data points lie neatly on a straight line within error bars for both $T_c(\chi)$ and $T_c(C_v)$. As regards the peaks of the specific heat, a fit performed over the range $L=[256-8192]$ yielded the same estimate of the infinite size temperature, which seems to indicate that $L=8192$ still belongs to the region where a linear fit is reliable. Infinite size temperatures are reported in Table V: the temperatures computed from both methods compare very well with each other and with the value of 1.691(3) reported in Ref. [7] using a multicanonical approach and medium lattice sizes. The value obtained from a previous MC study based on a canonical version of the Luijten-Blöte algorithm [69] lies clearly above our estimate, while estimates obtained from a transfer matrix method [70] and a real-space renormalization group ap-
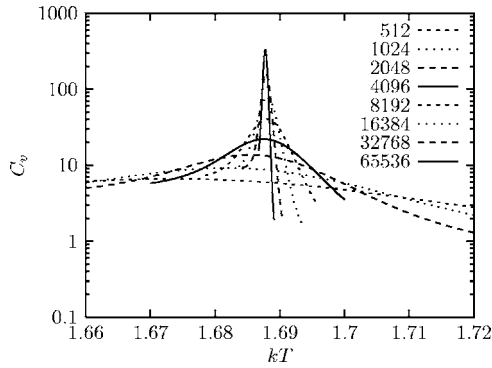
FIG. 10. Specific heat for the three-state Potts chain with σ =0.5 as obtained with our method.

proach [71] fall markedly below our values. Finally, the best estimate determined so far (to the best of our knowledge) with a numerical approach, namely, the value of T_c =1.685 42 obtained in Ref. [72] with the cluster mean-field method, lies slightly below ours. The occurence of a crossover for both response functions led us to perform a distinct linear fit restricted to the three largest lattice sizes (see the inset of Fig. 11 and results in Table V): this yielded $T_c(\infty)$ =1.687 64(1) and $T_c(\infty)$=1.687 65(2) for $T_c(C_v)$ and $T_c(\chi)$, respectively. In terms of the number of digits of precision, these estimates are comparable to the value reported in Ref. [72], although they still lie above it by around 0.13%. This is all the more surprising that, even if the accuracy of the cluster mean-field method decreases with increasing σ (with σ
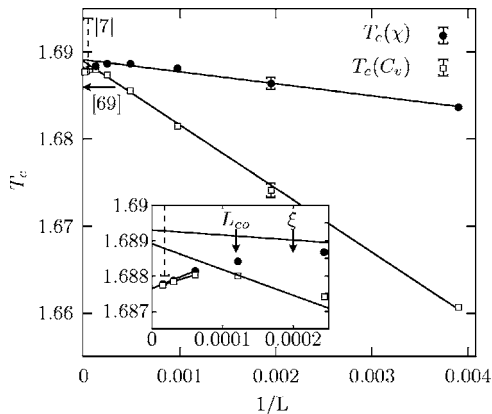

FIG. 11. Solid lines show linear fits of finite-size transition temperatures vs $1/L$ for the three-state LR Potts chain with σ=0.5, performed over sizes ranging from L=256 to L=4096. $T_c(C_v)$ and $T_c(\chi)$ are defined from the peaks of the specific heat C_v and the susceptibility χ, respectively; all data were obtained using cluster updates (CU). The inset shows a detailed view at larger lattice sizes, along with a linear fit carried out over the three largest sizes L =16 384, L=32 768, and L=65 536. Error bars are smaller than the size of symbols, except where explicitly indicated. The horizontal arrow shows the infinite-size estimate T_c=1.685 42 obtained by Monroe using the cluster mean-field method [72], while the vertical dashed error bar refers to the infinite-size estimate T_c=1.691(3) obtained in Ref. [7]. Finally, L_{co} indicates the lattice size at which the graph of the peaks of U_L vs the lattice size changes slope (see Fig. 12), and ξ refers to our estimate of the correlation length.

TABLE V. Infinite size transition temperatures computed from a fit of finite-size temperatures to a power law of the form $T_c(L)$ $-T_c(\infty)=a/L$. The first column indicates the range of lattice sizes that were used in the fit. Results from previous studies are shown for comparison in the last column: Ref. [7], MC study based on Berg's multicanonical method and sizes up to L=400; Ref. [72], cluster mean-field method; Ref. [70], transfer matrix method; Ref. [69], MC study based on the Lujiten-Blöte cluster algorithm and sizes up to L=3000; Ref. [71], real-space renormalization group approach.

$[L_1-L_2]$	$T_c(C_v)$		
	(CU)	(SSU)	
256–4096	1.6889(1)	1.6888(8)	
256–8192	1.6889(1)		
16384–65536	1.68764(1)		
	$T_c(\chi)$		
	(CU)	(SSU)	
256–4096	1.6893(1)	1.6892(6)	
16384–65536	1.68765(2)		
Ref. [7]			1.691(3)
Ref. [72]			1.68542
Ref. [70]			1.664
Ref. [69]			1.72
Ref. [71]			1.41

$\leqslant 0.5$ being quoted by the authors in Ref. [72] as the interval where the method produces its most accurate estimates), this method always produces—by construction—upper bounds for transition temperatures, irrespective of the size of the clusters. Noteworthy enough, a similar discrepancy was reported by the authors of an MC study of the LR Ising model based on the Luijten-Blöte algorithm [73], with the cluster mean-field approach [74] yielding values underestimated by 0.8% for σ=0.5. Since the authors in Refs. [72,74] performed an extrapolation procedure to compute the infinite-size estimate from temperatures obtained at finite cluster size, the discrepancy may thus be attributed to their specific extrapolation procedure rather than to the method used to produce finite-size estimates.

In order to shed light on the presence of a crossover, we compared the lattice size at which the crossover occurs with two characteristic lengths. One such length is the lattice size L_{co} where the peak of the reduced Binder cumulant of the energy, namely, $U_L^{max}=\max(\langle E^4\rangle/\langle E^2\rangle^2)$, experiences a minimum. At a second-order phase transition, U_L^{max} tends trivially to 1 in the thermodynamic limit, while it tends to a distinct value if the transition is of the first order. As illustrated in Fig. 12, the graph of U_L^{max} with respect to the lattice size is clearly characteristic of a first-order transition, with an infinite-size value lying close to 1.033. A minimum occurs at $L_{co} \sim 8192$ and—as shown in the inset of Fig. 11—the crossover occurs slightly above L_{co} for $T_c(C_v)$, and slightly below for $T_c(\chi)$. For σ=0.2, the same correlation is witnessed by our results, with the change of slope of U_L^{max} taking place
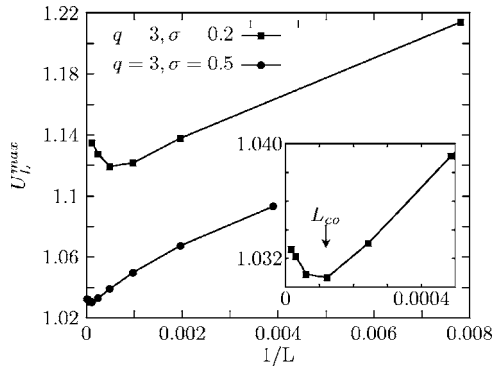
FIG. 12. Peak of the cumulant of the energy $U_L = \langle E^4 \rangle / \langle E^2 \rangle^2$ as a function of the lattice size for the three-state LR Potts chain with $\sigma = 0.2$ and $\sigma = 0.5$. The inset shows a magnification near the origin for $\sigma = 0.5$; L_{co} indicates the approximate lattice size at which the curve changes slope, crossing over to a first-order transition behavior.

FIG. 13. Graph of the free energy $F(T,E) = -\ln N(T,E)$ for the three-state LR Potts chain with $\sigma = 0.5$. $N(T,E)$ is the histogram of the energy reweighted at a pseudotransition temperature T_{eqh} where both peaks have the same height. For the four lattice sizes shown here, lattice configurations corresponding to phase coexistence are suppressed by a factor ranging from 0.1 to 10^{-6} with respect to pure phase configurations; for the three largest sizes, we note that a canonical simulation is clearly intractable.

around $L_{co} = 2048$, and a change of behavior $T_c(C_v)$ occuring near $L = 4096$. Another characteristic length is the correlation length of the disordered phase, which we compute in a subsequent part of this section: for $\sigma = 0.5$, our estimate of $\xi_d \sim 5000$ lies, here again, very close to the change of behavior of $T_c(C_v)$ and $T_c(\chi)$ (see the inset of Fig. 11); for $\sigma = 0.2, \xi_d \sim 600$, and $T_c(C_v)$ and $T_c(\chi)$ depart from the straight line at $L \sim 3000$ and $L \sim 1000$, respectively. For both values of σ, we thus observe the same qualitative behavior, namely, the crossover occurs at sizes slightly above L_{co} and ξ. We think that this may be attributed for one part to the fact that, as long as the lattice size is smaller than the correlation length, the finite-size scaling behavior of the model is that of a continuous transition, whereas it is first-order-like at larger sizes. Another important finite-size effect may be specifically linked to the long-range nature of the interaction: in Ref. [7], it was suggested that, irrespective of the type of periodic boundary conditions implemented, the finite size of the lattice *shifts* the (effective) decay parameter σ towards the mean-field regime, in a way that is more pronounced at small lattice size; this effect may thus compete with the previous effect in a nontrivial manner. It is therefore apparent that the power law used in our fits is not sufficient to model the finite-size scaling behavior in an accurate way over the entire range of lattice sizes; the derivation of an appropriate finite-size scaling law including corrections to scaling is left, however, to a subsequent work.

We also note in passing that, for $\sigma = 0.5$, relying on the Binder cumulant to assess the first-order nature of the transition requires simulating the system up to sizes that are far beyond the capabilities of single-spin update implementations. In particular, performing a power-law fit of U_L restricted to sizes below $L \sim 3000$ would yield underestimated values. Our results in Fig. 12 show that the infinite size value lies around 1.033, and thus that the transition is stronger than suggested, for instance, in Ref. [69].

Although a precise determination of correlation lengths for the LR Potts chain is beyond the scope of this work, we tried to obtain rough estimates of them from the finite-size scaling behavior of the interface free energy F^s. As in Sec. V,
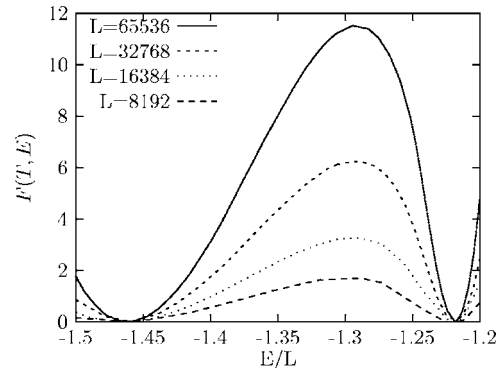
we computed a histogram $N(T,E)$ of the energy reweighted at the pseudotransition temperature T_{eqh} where both peaks of the histogram have equal height (see Fig. 13). After normalization of the peak heights to unity, $2F^s$ is given by $-\ln P_{min}$ where P_{min} stands for the minimum of the histogram between the two energy peaks. With this definition, which is analog to that used in models with nearest-neighbor interactions, $2F^s$ corresponds to the free energy cost associated—when periodic boundary conditions are used—with the creation of mixed ordered-disordered configurations from pure phase configurations. As we will witness, there is strong evidence that the dimension of the interface is no longer an integer. By fitting the interface free energy to the power law $F^s \propto L^\alpha$, we obtain a very good fit for sizes ranging from $L = 256$ to $L = 65\,536$, yielding $\alpha = 0.91(2)$ and $2F^s/L^\alpha = 0.0004$ in the thermodynamic limit. This is illustrated in Fig. 14. For $\sigma = 0.2$, we obtain $\alpha = 0.74(3)$ from a fit performed over sizes ranging from $L = 512$ to $L = 8192$. In view of the expected behavior for nearest-neighbor interactions, namely, F^s scales to leading order as a power of the lattice size with an expo-
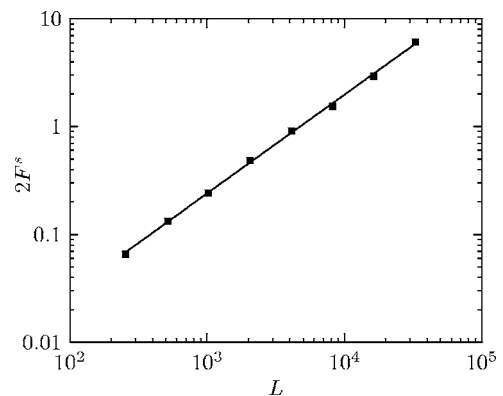

FIG. 14. Fit of the interface free energy F^s to a power law of the lattice size for the three-state LR Potts chain with $\sigma = 0.5$. All estimates of F^s were obtained with our method.

nent given by the dimension of the interface [75], this suggests that the effective dimension of the interface lies between 0 and 1 for long-range chains. This assumption is also supported by the fact that the fits of F^s/L in Ref. [69] exhibit important finite-size corrections, while our fit with a noninteger exponent does not suggest such corrections.

Finally, we can estimate the correlation length by transposing the argument that was proposed in Ref. [59] to relate the interface tension f^s of the nearest-neighbor Potts model to the correlation length ξ_d of its disordered phase. This argument relies on two ingredients: (i) the correspondence between the order-order interface tension and the correlation length of the disordered phase, which results from duality; (ii) the assumption of complete wetting, which implies a simple relation between the order-order and order-disorder interface tensions, where the latter is the quantity f^s we measure from the reweighted histogram of the energy. The second assumption is rigorous if q is large enough, although the authors in Ref. [59] suggest that it should stay valid for all q for which the transition is of the first order. The argument leads to the relation $\xi_d^{-1}=2f^s=2F^s/L$ for the nearest-neighbor model: in view of the finite-size scaling behavior of F^s reported above for the LR chain, it seems natural to transpose this relation by defining the interface tension as $f^s=F^s/L^\alpha$ for LR chains, and dimensional analysis thus suggests to use the relation $\xi_d^{-\alpha}=2f^s=2F^s/L^\alpha$ in the LR case. This yields an estimate of $\xi_d \sim 5000$ for the LR chain at $\sigma=0.5$, a value that seems at least consistent with the fact that the change of slope of U_L^{max} occurs at a lattice size slightly above this size (see Fig. 12). For $\sigma=0.2$, the same relation yields $\xi_d \sim 600$, and here again this estimate is consistent with the behavior of U_L^{max}; in addition, and as expected, ξ_d decreases with σ, i.e., as the transition becomes stronger. Although there is no rigorous proof that the ingredients invoked in Ref. [59] retain their validity in the case of LR interactions, it seems that a straightforward transposition to LR interactions yields reasonable results, even though the topology of the interface between the ordered and disordered phases is certainly far more complex than in nearest-neighbor models. In addition, we make use of infinite image periodic boundary conditions [11], so that the factor 2 in the definition of the interface tension might be questionable in LR models. Altogether our estimates should thus be taken as very rough ones.

VII. COMBINATION WITH THE TRANSITION MATRIX METHOD

In this section, we examine how our method can be efficiently combined with the transition matrix method [25]. We show in particular that transition matrices represent a very efficient way of estimating the microcanonical temperature $\beta(E)$ used to compute cluster bond probabilities when nothing is known initially about the density of states. We also discuss how the estimated $\beta(E)$ can then be used as an efficient predictor to speed up the convergence towards the ground state during the early iterations of the Wang-Landau algorithm.

A. Efficient estimation of $\beta(E)$ and bootstrapping

As seems obvious from the scheme presented in Sec. II, one of the basic requirements of our algorithm is to have an
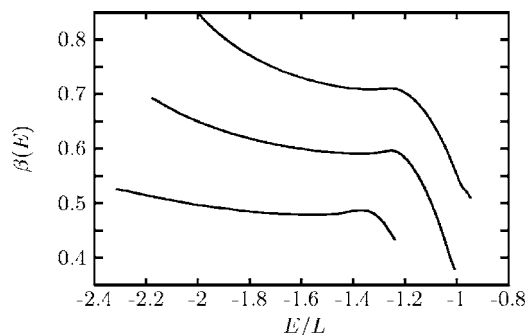

FIG. 15. Microcanonical inverse temperature $\beta(E)=dS(E)/dE$ computed from the estimated density of states using a spline interpolation, for the three-state long-range chain with $\sigma=0.4$, 0.5, and 0.6 from bottom to top.

estimate of $\beta(E)$ at our disposal over the whole energy axis in order to compute cluster bond probabilities. One rather simple way of estimating $\beta(E)$ is to compute it from the current estimate of the density of states $n(E)$ using a finite-difference scheme, i.e., in real-time in the course of the iteration scheme. This is the most tractable approach if one decides to rely solely on Wang-Landau's algorithm to estimate $n(E)$. During early iterations, however, the estimate of $n(E)$ is somewhat rough and it is necessary to resort to a spline interpolation in order to obtain a sufficiently smooth estimate of $\beta(E)$. Since the unequal spacing of energy levels in long-range models renders an interpolation scheme for $n(E)$ absolutely mandatory [7], $\beta(E)$ is already available to us for free. Figure 15 shows estimates obtained with this approach for the three-state long-range chain with various interaction ranges, computed after ten iteration steps of 10 000 measurements each. We note in passing that the presence of a clearly visible minimum in the three cases results from the first-order nature of the transition. For sufficiently short-range interactions, and when no random disorder is present, the microcanonical entropy $S(E)$ scales quite gently with the lattice size, and it is also perfectly feasible to use the value of $\beta(E)$ obtained at a smaller lattice size as an initial guess.

In any case, it is crucial for the performance of our algorithm that we should compute $\beta(E)$ to sufficient accuracy. Indeed, we have found that any departure from the ideal line results in poorer performance, as illustrated in Fig. 16. The curve (a) in the figure shows the mean acceptance rate as a function of the energy for an estimate of $\beta(E)$ obtained after the ultimate Wang-Landau iteration and a modification factor $\ln f = 10^{-7}$. Curves (b) and (c) show the same quantity for microcanonical temperatures that were under- and overestimated by 10%. The poor estimate of $\beta(E)$ causes a marked decrease of the acceptance rate in the transition region (around $E/L \sim -1.5$), from around 100% to nearly 40%. Tunneling times obviously experience a corresponding increase, from 243 for the under- and overestimated temperatures, respectively. This can be easily explained, if one considers that the efficiency of cluster updates reaches a maximum at the percolation threshold. Any departure of the estimate of $\beta(E)$ from the ideal line results in a shift between the temperature
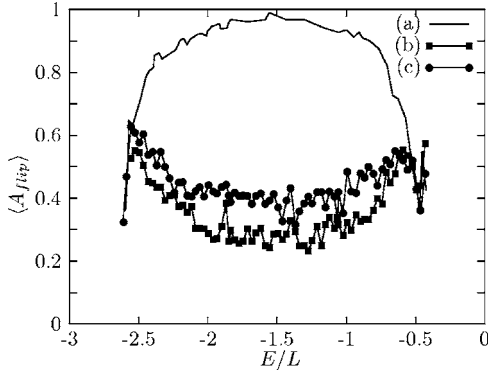
FIG. 16. Mean acceptance rate as a function of the energy per spin for the six-state long-range Potts chain with $\sigma=0.5$ and $L=512$ spins (strong first-order regime) for three distinct estimates of $\beta(E)$. (a) Best estimate, as given by the ultimate iteration of the Wang-Landau algorithm; (b) $\beta(E)$ scaled by 0:9; (c) $\beta(E)$ scaled by 1:1.
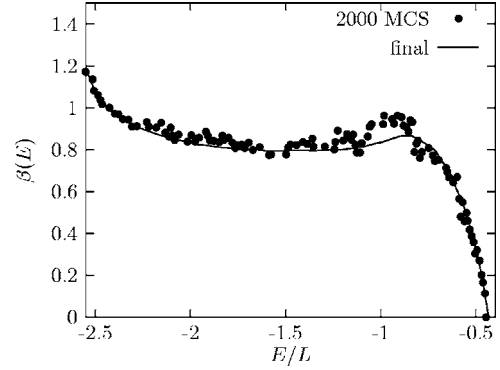

FIG. 17. Symbols show the microcanonical inverse temperature $\beta(E)$ computed from the transition matrix accumulated over 2000 MCSs, for the six-state LR model ($\sigma=0.5$) containing 512 spins. The estimate obtained from an interpolation scheme after the ultimate iteration is shown as a solid line for comparison.

at which clusters percolate [which depends on $\beta(E)$] and the *effective* temperature of the system [which is given by $dS(E)/dE$]. This behavior has been observed in the context of canonical simulations of disordered systems, e.g., the random field Ising model [76], where the presence of randomness depresses the critical temperature. In this case, using the (canonical) simulation temperature to compute the bond probabilities simply results in a growing shift between the critical temperature and the percolation threshold as the randomness is increased.

In view of the previously mentioned requirements on the estimation of $\beta(E)$, it is clear that, if one does not have a reliable guess of $\beta(E)$ at hand before the simulation starts, an efficient scheme must be devised in order to compute $\beta(E)$ in the early stage of the Wang-Landau algorithm. This is vital at this stage, because the exceedingly noisy estimate of the density of states makes it more likely to obtain under- or over-estimated values for $\beta(E)$. An efficient approach in this regards relies on transition matrices [25,77]. This method produces highly precise estimates of $\beta(E)$, although it has an inherently higher cost in terms of computer load. The starting point is the broad histogram equation [53,78]:

$$n(E)T_\infty(E \rightarrow E') = n(E')T_\infty(E' \rightarrow E),$$

where $T_\infty(E \rightarrow E')$ is the transition matrix element between energy levels E and E' [also denoted as $\langle N(\sigma, E'-E)\rangle_E$ in Ref. [78]]. This quantity contains the microcanonical average at energy E of the number of potential single-spin moves from a state σ of energy E to a state σ' of energy E'. It is estimated by accumulating a double-entry histogram $h(E, \Delta E)$ containing the number of potential moves from E to $E+\Delta E$ each time the energy level E is visited. Long-range interactions lead to energy levels which are irregularly spaced, with in particular a few gaps in the vicinity of the ground state [7], and it is necessary to choose an axis bin small enough to minimize discretization errors, and at the same time sufficiently large to contain at least a handful of entries. In this case, $T_\infty(E \rightarrow E')$ varies sufficiently smoothly

for the following approximation scheme to be valid:

$$\beta(E) = \frac{1}{\Delta E}\ln\frac{T_\infty(E \rightarrow E + \Delta E)}{T_\infty(E \rightarrow E - \Delta E)},$$

where the actual estimate is obtained by weighted-averaging over several values of ΔE. As illustrated in Fig. 17 for the six-state LR chain, the estimation of $\beta(E)$ from the transition matrix elements is reliable already after 2000 MCSs, which roughly corresponds to 50 round trips between the upper and the lower energy range. For long-range models, each estimation of the number of potential moves requires of order L^{2D} operations (as opposed to L^D for nearest-neighbor interactions). However, we have shown in Sec. III that a single cluster update can demand as little as $O(L^D\ln L^D)$ operations when long-range specific optimizations are carried out; hence estimation schemes based on transition matrices partly scupper the benefits of these optimizations, and should therefore be employed only as a bootstrap procedure when nothing is known yet about the microcanonical temperature. Conversely, models with nearest-neighbor interactions do not undergo such a drawback, and make the transition matrix approach a perfectly transparent one from the viewpoint of algorithm complexity.

B. Efficient predictors for the Wang-Landau algorithm

Finally, we discuss how $\beta(E)$ can be used as an efficient predictor during the early stage of the Wang-Landau algorithm when nothing is known about the density of states. In the original implementation of this algorithm, we start with $S(E)=0$ for all energy levels, and simply increment $S(E)$ by the modification factor $\ln f$ each time the corresponding energy level is visited. One of the main drawbacks of this approach is that the Markov chain tends to wander around a fairly long time in the upper energy range, until eventually enough visits have been recorded in the histogram for the system to start exploring low-energy levels. This point has already been mentioned in Ref. [34], where it was suggested that starting with a good initial guess of $S(E)$ was more efficient in terms of the number of histogram entries required
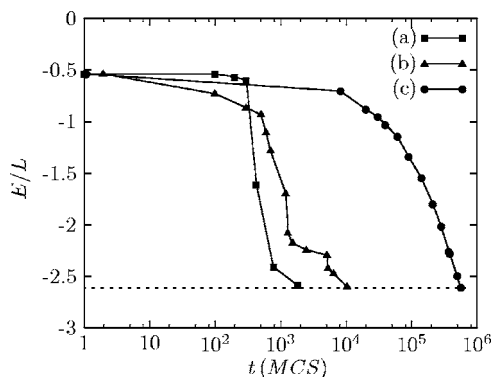
FIG. 18. The graph shows the number of MCSs needed to reach the ground-state (dashed horizontal line) of the six-state Potts chain ($\sigma=0.5$ and $L=512$) for an initially unknown density of states, using three distinct schemes: (a) and (b) predictor based on $\beta(E)$, local- and collective-update algorithms, respectively; (c) no predictor [$S(E)=0, \forall E$], local-update algorithm.

to reach the final estimate, than performing a multirange run with no initial guess at all. To circumvent this drawback when no initial guess is available, we therefore propose to use $\beta(E)$ to predict $S(E)$ for energy levels that have not been visited yet during the simulation, and thus for which $S(E)$ is not available [i.e., it is set to $S(E)=0$ in the original implementation of the Wang-Landau algorithm]. A linear prediction scheme turned out to be sufficiently efficient for our purpose. As illustrated in Fig. 18, using a predictor brings about a gain of three orders of magnitude in the time needed to reach the ground state. Our method and the single-spin update method lead similar performances, with however a slightly better behavior when cluster updates are used. We note that the Markov chain stays initially somewhat longer in the upper energy range when cluster updates are used, since a good estimate of $\beta(E)$ is needed to build the clusters with the correct bond probabilities. We think that this approach would prove particularly useful when the characteristics of the model makes it impossible to obtain an initial guess of $S(E)$ from simulations at smaller lattice sizes, e.g., in the presence of disorder or when the long-range interaction experiences a slow decay.

VIII. CONCLUSION

In conclusion, we have developed a new Monte Carlo method which combines in an efficient and straightforward way the benefits of flat histogram algorithms with the fast-decorrelating capabilities of cluster updates. It is suited for spin models with any number of interaction between spins. Our formulation is versatile, and the method can be applied to a variety of density of states estimation schemes, including the Wang-Landau algorithm, Berg's recursion scheme, or the transition matrix method. We have shown that using the microcanonical temperature to compute cluster bond probabilities leads to a drastic reduction in effective autocorrelation times, tunneling times, and equilibration times. In the context of the Wang-Landau implementation, the reduced correlation between successive binning of the energy histogram yields a lower error in the estimation of the density of states, and as a result more reliable estimates of thermodynamic averages. Several schemes for the estimation of the microcanonical temperature were proposed, among which an efficient procedure which harnesses the power of the transition matrix method, and allows us to bootstrap the algorithm even if nothing is known initially about the density of states. Finally, we carefully examined the precision of our method in the case of spin models with power-law decaying interactions. Here, our method proves all the more powerful that it is able to reduce the algorithm complexity to that of a short-range model having the same number of spins. This allowed us to study several finite-size effects at large lattice sizes, otherwise largely out of reach of conventional local-update implementations. In particular, we found out that the interface free energy scales perfectly well with a power of the lattice size, yet with a noninteger exponent which lies between 0 and 1. This, we think, is accounted for by the complex topology of the phases in coexistence in long-range models. A more detailed study, including a deeper insight into the topological properties of the generated clusters and the estimation of correlation lengths at large lattice sizes, would be very promising. We think that our method clearly draws this challenge within computation range.

[1] G. Giacomin and J. L. Lebowitz, Phys. Rev. Lett. **76**, 1094 (1996).

[2] P. J. Ford, Contemp. Phys. **23**, 141 (1982).

[3] D. J. Amit, *Modeling Brain Functions* (Cambridge University Press, Cambridge, England, 1989).

[4] E. Luijten, Phys. Rev. E **60**, 7558 (1999).

[5] W. K. Theumann and M. A. Gusmao, Phys. Rev. B **31**, 379 (1985).

[6] E. Luijten and H. W. J. Blöte, Phys. Rev. Lett. **89**, 025703 (2002).

[7] S. Reynal and H. T. Diep, Phys. Rev. E **69**, 026109 (2004).

[8] M. Biskup, L. Chayes, and N. Crawford, math-ph/0501067 (unpublished).

[9] S. A. Cannas and F. A. Tamarit, Phys. Rev. B **54**, R12661 (1996).

[10] B. P. Vollmayr-Lee and E. Luijten, Phys. Rev. E **63**, 031108 (2001).

[11] S. A. Cannas, C. M. Lapilli, and D. A. Stariolo, Int. J. Mod. Phys. C **15**, 115 (2004).

[12] E. Luijten and H. W. J. Blöte, Int. J. Mod. Phys. C **6**, 359 (1995).

[13] K. Uzelac and Z. Glumac, Fiz. B **6**, 133 (1997).

[14] E. Bayong, H. T. Diep, and V. Dotsenko, Phys. Rev. Lett. **83**, 14 (1999).

[15] K. Binder and E. Luijten, Comput. Phys. Commun. **127**, 126 (2000).

[16] K. Binder, E. Luijten, M. Müller, N. B. Wilding, and H. W. J. Blöte, Physica A **281**, 112 (2000).

[17] K. Binder and E. Luijten, Phys. Rep. **344**, 179 (2001).

[18] K. Uzelac, Z. Glumac, and A. Anicic, Phys. Rev. E **63**, 037101 (2001).

[19] S. Reynal and H. T. Diep, Comput. Phys. Commun. **169**, 243 (2005).

[20] B. A. Berg, Int. J. Mod. Phys. C **3**, 1083 (1992).

[21] Y. Iba, Int. J. Mod. Phys. C **12**, 623 (2001).

[22] B. A. Berg, Comput. Phys. Commun. **147**, 52 (2002).

[23] B. A. Berg and T. Neuhaus, Phys. Lett. B **267**, 249 (1991).

[24] B. A. Berg and T. Neuhaus, Phys. Rev. Lett. **68**, 9 (1992).

[25] J.-S. Wang, T. K. Tay, and R. H. Swendsen, Phys. Rev. Lett. **82**, 476 (1999).

[26] G. R. Smith and A. D. Bruce, J. Phys. A **28**, 6623 (1995).

[27] F. Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001).

[28] F. Wang and D. P. Landau, Phys. Rev. E **64**, 056101 (2001).

[29] B. Hesselbo and R. B. Stinchcombe, Phys. Rev. Lett. **74**, 2151 (1995).

[30] S. Trebst, D. A. Huse, and M. Troyer, Phys. Rev. E **70**, 046701 (2004).

[31] Y. Wu, M. Koerner, L. Colonna-Romano, S. Trebst, H. Gould, J. Machta, and M. Troyer, Phys. Rev. E **72**, 046704 (2005).

[32] W. Janke and S. Kappler, Phys. Rev. Lett. **74**, 212 (1995).

[33] P. Dayal, S. Trebst, S. Wessel, D. Würtz, M. Troyer, S. Sabhapandit, and S. N. Coppersmith, Phys. Rev. Lett. **92**, 097201 (2004).

[34] C. Zhou and R. N. Bhatt, Phys. Rev. E **72**, 025701 (2005).

[35] W. Janke and S. Kappler, Nucl. Phys. B (Proc. Suppl.) **42**, 876 (1995).

[36] W. Janke, Physica A **254**, 164 (1998).

[37] M. S. Carroll, W. Janke, and S. Kappler, J. Stat. Phys. **90**, 1277 (1998).

[38] C. Yamaguchi and N. Kawashima, Phys. Rev. E **65**, 056710 (2003).

[39] K. Rummukainen, Nucl. Phys. B (Proc. Suppl.) **390**, 621 (1993).

[40] M. Krech and E. Luijten, Phys. Rev. E **61**, 2058 (2000).

[41] K. Uzelac and Z. Glumac, Physica A **314**, 448 (2002).

[42] F. Niedermayer, Phys. Rev. Lett. **61**, 2026 (1988).

[43] F. Niedermayer, in *Advances in Computer Simulation, Lectures Held at the Eötvös Summer School in Budapest, Hungary, 1996*, edited by J. Kertész and I. Kondor (Springer, Berlin, 1998).

[44] J. Asikainen, A. Aharony, B. Mandelbrot, E. Rausch, and J.-P. Hov, Eur. Phys. J. B **34**, 479 (2003).

[45] W. Janke and A. M. J. Schakel, Phys. Rev. E **71**, 036703 (2005).

[46] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **63**, 1195 (1989).

[47] B. A. Berg, J. Stat. Phys. **82**, 323 (1996).

[48] B. A. Berg, Fields Inst. Commun. **26**, 1 (2000).

[49] J. Lee, Phys. Rev. Lett. **71**, 211 (1993).

[50] Q. Yan and J. J. de Pablo, Phys. Rev. Lett. **90**, 035701 (2003).

[51] H. S. Chan and K. A. Dill, J. Chem. Phys. **99**, 2116 (1993).

[52] J.-S. Wang, Physica A **281**, 147 (2000).

[53] J.-S. Wang and R. H. Swendsen, J. Stat. Phys. **106**, 245 (2002).

[54] M. S. Shell, P. G. Debenedetti, and A. Z. Panagiotopoulos, J. Chem. Phys. **119**, 9406 (2003).

[55] T. Surungan, Y. Okabe, and Y. Tomita, J. Phys. A **37**, 4219 (2004).

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

[57] J. Hoshen and R. Kopelman, Phys. Rev. B **14**, 3438 (1974).

[58] E. Luijten, *Computer Simulation Studies in Condensed-Matter Physics* (Springer, Heidelberg, 2000), Vol. XII, pp. 86–99.

[59] C. Borgs and W. Janke, J. Phys. I **2**, 2011 (1992).

[60] E. Buddenoir and S. Wallon, J. Phys. A **26**, 3045 (1993).

[61] W. Janke, in *Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms*, edited by J. Grotendorst, D. Marx, and A. Muramatsu (John Von Neumann Institute for Computing, Jühlich, Proceedings of the Euro Winter School, 2002), Vol. 10 of NIC Series, pp. 423–445.

[62] F. Y. Wu, Rev. Mod. Phys. **54**, 235 (1982).

[63] W. Janke and T. Sauer, Phys. Rev. E **49**, 3475 (1994).

[64] S. Alder, S. Trebst, A. K. Hartmann, and M. Troyer, J. Stat. Mech.: Theory Exp. (2004) P07008.

[65] M. L. Guerra and J. D. Munoz, Int. J. Mod. Phys. C **15**, 471 (2004).

[66] D. P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, Cambridge, England, 2000).

[67] W. Janke, B. A. Berg, and M. Katoot, Nucl. Phys. B **382**, 649 (1992).

[68] J. Lee and J. M. Kosterlitz, Phys. Rev. B **43**, 3265 (1991).

[69] Z. Glumac and K. Uzelac, Phys. Rev. E **58**, 4372 (1998).

[70] Z. Glumac and K. Uzelac, J. Phys. A **26**, 5267 (1993).

[71] S. A. Cannas and A. C. N. de Magalhães, J. Phys. A **30**, 3345 (1997).

[72] J. L. Monroe, J. Phys. A **32**, 7803 (1999).

[73] E. Luijten and H. W. J. Blöte, Phys. Rev. B **56**, 8945 (1997).

[74] J. L. Monroe, J. Phys. A **31**, 9809 (1998).

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

[76] M. E. J. Newman and G. T. Barkema, Phys. Rev. E **53**, 393 (1996).

[77] R. H. Swendsen (private communication).

[78] P. M. C. de Oliveira, T. J. P. Penna, and H. J. Herrmann, Braz. J. Phys. **26**, 677 (1996).